



HAL
open science

Organisational awareness : mapping human capital for enhancing collaboration in organisations

Dor Avraham Garbash

► **To cite this version:**

Dor Avraham Garbash. Organisational awareness : mapping human capital for enhancing collaboration in organisations. Library and information sciences. Université Sorbonne Paris Cité, 2016. English. NNT : 2016USPCB134 . tel-02141742

HAL Id: tel-02141742

<https://theses.hal.science/tel-02141742>

Submitted on 28 May 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Université Paris Descartes

Frontières du vivant

Systems Engineering and Evolution Dynamics - SEED Team INSERM U1001

Organisational Awareness: Mapping Human Capital for Enhancing Collaboration in Organisations

Par Dor Avraham Garbash

Thèse de doctorat d'Informatique Sociale (Social computing)

Dirigée par Ariel Lindner

Présentée et soutenue publiquement le 29.9.2016

Devant un jury composé de :

Fourcade Francois, Maître de conférences, H.D.R, ESCP - Rapporteur

Piece Chris, Phd and Lecturer, Stanford university - Rapporteur

Tesniere Antoine, PU-PH, Université Paris Descartes - Membre du jury

Cherel Eric, Directeur du Numérique, Université Paris Descartes - Membre du jury

Danos Vincent, Directeur de Recherches, École normale supérieure - Membre du jury

Prendre soin de l'organisation: cartographier le capital humain pour le renforcement de l'organisation.

Comment peut-on devenir plus conscients des sources de connaissance au sein des organisations humaines?

Les changements économiques et technologiques rapides forcent les organisations à devenir plus souples, agiles et interdisciplinaires. Pour cela, les organisations cherchent des alternatives aux structures de communication hiérarchiques traditionnelles qui entravent les pratiques de collaboration ascendantes.

Pour que les méthodes ascendantes soient efficaces, il est nécessaire d'offrir aux membres l'accès à l'information et à l'expertise dont ils ont besoin pour prendre des décisions qualifiées. Ceci est un défi complexe qui implique une culture organisationnelle, et des pratiques de travail et d'usage de l'informatique. Un défaut au niveau de l'application de ce système peut ralentir les processus de travail, entraver l'innovation et conduit souvent à un travail suboptimal et redondant. Par exemple, une enquête 2014 de 152 dirigeants de Campus IT aux Etats-Unis, estime que 19% des systèmes informatiques du campus sont redondants, ce qui coûte aux universités des états-uniennes 3.8B\$ par an. Dans l'ensemble, les travailleurs intellectuels trouvent l'information dont ils ont besoin seulement 56% du temps. Avec un quart du temps total des travailleurs intellectuels consacré à la recherche et l'analyse des informations. Ce gaspillage de temps coûte 7K\$ pour chaque employé par an. Un autre exemple du gaspillage est celui des nouveaux arrivants et des employés promus qui peuvent prendre jusqu'à 2 ans pour s'intégrer pleinement au sein de leur département.

En outre et selon des enquêtes étendues, seulement 28% des apprenants estiment que leurs organisations actuelles «utilisent pleinement» les compétences qu'ils ont actuellement à offrir et 66% prévoient de quitter leur organisation en 2020. Répondre à ce défi avec succès peut motiver les membres de l'organisation, ainsi qu'y améliorer l'innovation et l'apprentissage.

L'ambition de notre travail est de comprendre ce problème en étudiant les défis que rencontre le département informatique d'une université et d'une centre de recherche interdisciplinaire. Deuxièmement, co-développer et mettre en œuvre une solution avec ces institutions. Je décris leur utilisation des logiciels que nous avons développés, les résultats et la valeur obtenus avec ces pilotes.

Troisièmement, tester l'efficacité de la solution, et explorer de nouvelles applications et le potentiel d'un tel système utilisé à une plus grande échelle.

Pour mieux comprendre le problème je me suis engagé dans une discussion avec les membres et les dirigeants des deux organisations. Une conclusion importante des discussions est que les membres de ces organisations souffrent souvent d'un manque de sensibilisation à propos des compétences et des connaissances des processus et des relations sociales de leurs collègues dans l'organisation. A cause de cette situation, les idées novatrices, les opportunités et les intérêts communs des pairs sont sévèrement limités. Cela provoque des retards inutiles dans les projets inter-équipes, des goulots d'étranglement, et un manque de sensibilisation sur les possibilités de stages. Aussi, j'ai analysé le problème plus avant et l'ai défini un problème de fragmentation de l'information. Différentes informations sont stockées dans des bases de données disparates ou dans la tête des gens, exigeant un effort et de savoir-faire pour l'obtenir. Suite aux conclusions de cette analyse et l'examen des connaissances, nous avons mis l'ensemble des résultats afin de créer une base de données visuelle de collaboration pour cartographier les personnes, les projets, les compétences et les institutions pour le département informatique de l'Université Descartes, et en plus, les gens, les intérêts et les possibilités de stages au sein du CRI, un centre de recherche et de formation interdisciplinaire. Nous avons également mené des interviews, des sondages et des questionnaires qui montraient que les gens avaient des difficultés à identifier des experts en dehors de leurs équipes de base.

Au cours de cette thèse, j'ai progressivement surmonté ce défi en développant deux applications de web collaboratives appelées Rhizi et Knownodes. Knownodes est un graphique collaboratif de connaissances qui a utilisé des bords riches en informations pour décrire les relations entre les ressources. Rhizi est une plateforme pour cartographier la connaissance collaborative du capital dans le temps réel. Une caractéristique unique de la plateforme de Rhizi est qu'il fournit une interface d'utilisateur qui transforme des affirmations basées sur des textes faits par les utilisateurs dans un graphe visuel des connaissances. Les assertions sont stockées sous forme de données structurées qui sont simples à interroger, à explorer et à mettre à jour. Le produit final du processus est un ensemble de cartes de ressources transparentes. Rhizi a évolué à travers plusieurs projets pilotes réalisés dans dix contextes différents, créant ainsi les cartographies des individus, des projets, des compétences au sein des cartes visuelles distinctes qui décrivent les relations entre ces entités. Parmi nos réalisations, on trouve la création d'un éditeur graphique collaboratif dans le temps-réel, et une interface conviviale pour la saisie et l'exploration des informations au sein d'une base de données graphique.

A la suite des projets pilotes, j'ai mené plusieurs enquêtes, des entretiens semi-structurés et des tests qui m'ont permis d'évaluer la valeur du logiciel, ainsi que la collecte de commentaires. Les principales conclusions sont les suivantes:

1. L'identification d'expert au sein du département informatique de l'université Descartes était significativement plus efficace à l'aide du logiciel Rhizi.
2. La cartographie d'experts participative basée sur plusieurs sources est possible sur une base volontaire.
3. La cartographie des compétences a déclenché une collaboration inattendue entre les étudiants et même avec les équipes en pleine concurrence.
4. L'utilisation des logiciels a accru la motivation pour partager l'expertise et collaborer avec les autres.

Pour synthétiser les connaissances recueillies tout au long de cette recherche, je déclare que le coût perçu élevé et le manque d'incitations sont les principaux points qui bloquent la collaboration inter-équipe.

Je termine cette thèse avec quelques observations à propos des moyens praticables pour simplifier une collaboration à grande échelle au sein des organisations et attacher une proposition visant à construire une nouvelle application logicielle basée sur les conclusions des projets pilotes qui ont utilisé Rhizi.

Organisational Awareness: Mapping Human Capital for Enhancing Collaboration in Organisations

Abstract

How can we become more aware of the sources of insight within human organisations? Rapid economical and technological changes force organisations to become more adaptive, agile and interdisciplinary. In light of this, organisations are seeking alternatives for traditional hierarchical communication structures that hinder bottom-up collaboration practices.

Effective bottom-up methods require empowering members with access to the information and expertise they need to take qualified decisions. This is a complex challenge that involves organisational culture, IT and work practices. Failing to address it creates bottlenecks that can slow down business processes, hinder innovation and often lead to suboptimal and redundant work. For example, a 2014 survey of 152 Campus IT leaders in the US, estimated that 19% of the campus IT systems are redundant, costing US universities 3.8B\$ per year. In aggregate, knowledge workers find the information they need only 56% of the time. With a quarter of knowledge workers total work time spent in finding and analyzing information. This time waste alone costs 7K\$ per employee annually. Another example of the waste created is that newcomers and remote employees may take up to 2 years to fully integrate within their department.

Furthermore according to extended surveys, only 28% of millennials feel that their current organizations are making 'full use' of the skills they currently have to offer and 66% expect to leave their organisation by 2020. Successfully resolving this challenge holds the potential to motivate organisation members, as well as enhance innovation and learning within it.

The focus of this thesis is to better understand this problem by exploring the challenges faced by a university IT department and an interdisciplinary research center. Second, co-develop and implement a solution with these institutions, I describe their usage of the software tool we developed, outcomes and value obtained in these pilots. Third, test the effectiveness of the solution, and explore further applications and potential for a similar system to be used in a wider scale.

To better understand the problem I engaged in discussion with members and leaders of both organisations. An important conclusion from the discussions is that members of these organizations often suffer from lack of awareness about their organisation's knowledge capital—the competencies, knowledge of processes and social connections of their colleagues. Due to this exposure to innovative ideas, opportunities and common interests of peers is severely limited. This causes unnecessary delays in inter-team projects, bottlenecks, and lack of awareness about internship opportunities. I further broke down the problem, and defined it as one of information fragmentation: Different information is stored in disparate databases or inside people's heads, requiring effort and know-how in order to obtain it. Following the conclusions of this analysis and state-of-the-art review, we have set together the goal to create a collaborative visual database to map the people, projects, skills and institutions for the IT department of Descartes University, and in addition, people, interests and internship opportunities within the CRI, an interdisciplinary research and education center. We have also conducted interviews, surveys and quizzes that ascertain that people had difficulties identifying experts outside their core teams.

During the course of this thesis, I progressively addressed this challenge by developing two collaborative web applications called Rhizi and Knownodes. Knownodes is a collaborative knowledge graph which utilized information-rich edges to describe relationships between resources. Rhizi is a real-time and collaborative knowledge capital mapping interface. A prominent unique feature of Rhizi is that it provides a UI that turns text-based assertions made by users into a visual knowledge graph. The assertions are stored as structured data that is simple to query, explore and update. The final product of the process is a set of transparent *resource maps*. Rhizi evolved through multiple pilot projects made in ten different contexts, creating mappings of individuals, projects, skills within distinct visual maps that describe the relationships between these entities. Among our achievements was the creation of a real-time collaborative graph editor, and a user-friendly interface for inputting and exploring information within a graph database.

Following the pilot projects, I have conducted several surveys, semi-structured interviews and tests that helped evaluate the value of the software, as well as collection of feedback. The principal findings were:

1. Expert identification within the IT department of Descartes university was significantly more effective with the help of the Rhizi software.
2. Crowd-sourced and participatory expert mapping is possible on a voluntary basis.
3. Skill mapping has triggered unexpected collaboration between students and even *competing teams*.
4. Usage of software Increased motivation for sharing expertise and collaborating with others.

Synthesizing the insights gathered throughout this research, I conclude that high perceived cost and lack of incentives are the main blocking points of effective inter-team collaboration.

I finish the thesis with some observations about practical ways to simplify large scale collaboration within organisations and attach a proposal to build a new software application based on the conclusions from the pilot projects that utilized Rhizi.

Mots clés (français) : expertise du management, intelligence collective, cartographie des ressources, logiciel de collaboration, interface graphique, informatique sociale.

Keywords : expertise management, collective intelligence, resource mapping, collaboration software, graph interface, social computing.

Dedicated to my mother, Naama, for her wise and loving support

Acknowledgements

This work was supported in part by Les Laboratoires Servier as a PhD fellowship to DG. We wish to thank the director and team members of the IT department under study and specifically for the following people:

Alon Levy for amazing generosity, persistence and help developing Rhizi. **Owen Cornec** For prototyping Rhizi while being bombarded by rockets **Amir Sagie** For his work on developing Rhizi's backend. **Ofer Lehr** for work on Rhizi's graphics and user research. **Liad Magen** and **Mikael Couzic** For the development of Knownodes. **Alexandre Lejeune**, **Dmitry Paranyushkin** and **Armella Leung** for their visionary work on Knownodes interface and graphics. **Yael Ben-Dov** for many hours of sage consultation regarding UX for both Rhizi and Knownodes. **Eyal Rotbart** and **Emilie Laffray** for product management and consulting. **Petr Johanes**, **Francois Taddei** and **Francois Fourcade** for consulting and ideation. **Eric Cherel** for his wise guidance. **Shimon Amar** and **Stephen Friend** for their confidence in the project. **Emily Schneider** for her love and support and special acknowledgement for **Ariel Lindner** for his mentorship, vision and relentless support throughout the twists and turns of this project.

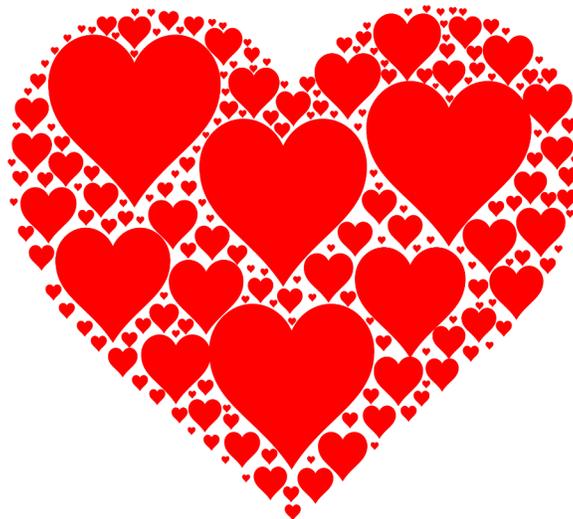


Table of contents

Acknowledgements	6
Table of contents	7
List of figures	9
Thesis Structure	12
1. Introduction	13
1.1 Problems and Challenges	15
1.1.1 Why human capital awareness?	16
1.1.2 What stands in the way of organisational awareness?	20
1.1.3 Human factors limiting organisational self-awareness	22
1.1.4 Structural factors limiting organisational self-awareness	26
1.1.5 Technical factors limiting organisational self-awareness	29
1.2 Background	32
1.2.1 Human Capital Within Organisations	33
1.2.2 Collective Intelligence	36
1.2.3 The Network Paradigm Within Organisations	41
1.2.4 Knowledge Management	46
1.3 Review of existing solutions	51
1.3.1 Review of academic literature on knowledge sharing	52
1.3.2 Expert identification software and methods	66
1.4 Challenges summary	85
2. Development and design methodology	87
2.1 On usage of Lean startup methodologies within a Phd	99
2.2 From Human-centric design to Activity-centric design	108
2.3 Technical approach	116
3. Results and Pilot projects	118
3.1 Knownodes Technology and data structure description	120
3.2 Rhizi Technology and data structure description	126
3.3 Knownodes Bibsyn pilot	131
3.4 CRI - Opportunity mapping	141
3.5 Descartes IT	156
3.6 IGEM	172
3.7 Other pilot projects	178
4. Conclusion: Met and unmet challenges analysis	191
4.1 Proposal for a future project	198
Bibliography	199
Appendix 1 Instructions for AIV students	220
Appendix 2 Rhizi manual	222
Appendix 3 Manuscript submitted to Human Computation Journal	229
Appendix 4 Product scenarios	254

[Appendix 5 Digital organisation - example project page](#)255

[Index](#)257

List of figures

- 1: Percentage of millennials expecting to leave the organisation (Deloitte, 2016).
- 2: Gaps in job satisfaction issues between millennials who are leaving or staying in their workplace (Deloitte, 2016).
- 3: Synthesis of the different barriers towards diffusion of best practices
- 4: Simple breakdown of perceived cost and lack of incentives for organisation wide collaboration.
- 5: Adapted from Expanding capital for competitive advantage. Business horizons 47/1 January–February 2004 (45–50)
- 6: Predictors for the collective intelligence factor c . Suggested by Woolley et Al. 2015.
- 7: Exponential growth of publications indexed by Sociological Abstracts containing “social network” in the abstract of title.
- 8: Ego network and overall network
- 9: A knowledge management framework and the place of knowledge identification within it. (Probst et al. 2000)
- 10: the T shaped model
- 11: The paint-drip model.
- 12: The results of factor analysis on group characteristics variables (Sawng, Kim, & Han, 2006).
- 13: Components of expertise seeking: Each one of the components might affect the preferred selection criteria (Hertzum 2014).
- 14: Expert selection interface taking into account the top criteria for selection: Overall expertise, case based expertise and crowdsourced rating (Paul 2016).
- 15 - formal and informal communication networks(What is ONA? 2016)
- 16: List of leading ONA providers(Organizational network analysis, 2016)
- 17: The who of knowledge mapping(Wexler 2001)
- 18: The main steps of the ExTra process(Weber et al., 2007)
- 19: Different software projects during thesis.
- 20: PPG’s Framework for Responding to Wicked Issues(Camillus 2008)
- 21 - Presentation slide of Knownodes project 24.10.2013
- 22 - Knownodes vision: Create an open parallel universe on top of the world wide, where people can connect and debate about content.
- 23: The lean startup feedback loop.
- 24: Venn diagram of an MVP
- 25: Main development hypotheses for each product.
- 26: Example of user archetypes used in Rhizi V2
- 27: The human centered design pyramid (Giacomin, 2014)
- 28: Mock-up of connection centric quizzing platform
- 29: Mock-up of a browser plug-in to connect between webpages.
- 30: Flip animation of browser plug-in for Knownodes.
- 31: Mockup of an annotation system for MOOCS
- 32: Mockup of Knownodes page interface

- 33: Mock up of a connection between concepts in Knownodes
- 34: Knownodes map interface.(Credit: Alexandre Lejeune)
- 35: Knownodes query interface mockup(Credit: David Bikard)
- 36: Rhizi search results mockup
- 37: Rhizi Gantt chart view mode
- 38: Timeline of the different pilots during the course of the thesis.
- 39: Knownodes features information-rich connections between entities.
- 40: How To use Knownodes
- 41: Advanced social features implemented in Knownodes V2.
- 42: Knownodes data structure
- 43: (a) Syntax of a simple phrase to input in the graph. (b) Syntax and text-to-network preview in interface.
- 44: The different visualization layouts in Rhizi: Force layout, ring layout and custom layout.
- 45: Rhizi data model
- 46: Node types and their properties as defined for the IT department use-case
- 47 : (a) Query result for Shibboleth (b) Shortest path between person104 to Wifi skill (c) Project overview - node size depict total amount of work hours invested in project (d) project centered view with time-allocation visualization per staff member.
- 48: Students using Knownodes during a Bibsyn session
- 49: Knownodes page displays all connections related to “Translation”, a concept in biology from the course.
- 50: Mockup Knownodes step by step form
- 51: Mockup Create new Problem/Question dialogue box
- 52: Mockup following first iteration
- 53: Final input form that was implemented
- 54: Visualization of Knownodes database including rapid prototypes and workshops
- 55: Rhizi CRI landing page
- 56: CRI M2 2015 map. View is arranged in custom layout as columns from left to right: organisation (purple), non students people (blue), students (blue), internships (red), skills (yellow) and keyword (green).
- 57: Ewen Corre M2 2015 Ego network.
- 58: Clicking on the modeling skill and the resulting view.
- 59: (left)Internship centric view (right) Infocard revealing additional information regarding the internship
- 60: Shortest path visualization between a student and prospective institution.
- 61: Types of nodes inputted by students for each map throughout the pilot
- 62: Mapping learning intentions
- 63: Mapping learning Intentions to learning opportunities
- 64: Survey results on questions related to needs related to expertise and project awareness
- 65: Results of the expert identification quiz.
- 66: Node types and their properties as defined for the IT department use-case
- 67: Number of entities in each of the resulting maps, according to Entity type. (True for 04.15.2016)
- 68: Survey results on questions related to experience with Rhizi.
- 69: Survey results on questions related to network visualization interface.
- 70: Team-centric view of CSU Fort Collins
- 71: Organism-centric view on “Saccharomyces cerevisiae” organism
- 72: The CRI IGem team and the nodes connected to them.

- 73: Card view of the Paris Bettencourt team
- 74: Circular layout of the Igem map.
- 75: Rhizi Igem perspectives
- 76: Knownodes concept example.
- 77: Mockup of the final result.
- 78: Data schema for nano-publications
- 79: Mock-up of input interface for Nanopub platform
- 80: Mock-up of input interface for Nanopub platform
- 81: Query interface for Nanopub platform
- 82: Example of a map interface containing existing collaboration(in blue) and potential collaboration ideas(in orange)
- 83: A list view of the most recent connections created by the system.
- 84: Illustration of usage of Knownodes within a MOOC

Thesis structure

This thesis contains four main chapters: Introduction, Technology and Design of Rhizi, Results and pilot projects and Conclusion: Met and unmet challenges.

The Introduction provides the necessary background for the premise of the thesis. It is divided into three parts: The first part defines the problems and challenges this thesis deals with. The second part provides background to basic concepts dealt within it: Human capital, Collective intelligence, the Network Paradigm in organisations and Knowledge Management. The third part is a review of how academic literature and current software solutions deal with the challenges posed in the first part of the chapter.

The Design and Technology of Rhizi chapter describes the methodologies used for managing the different software projects and pilots during the thesis.

The results and pilot projects chapter describes the Software developed during this thesis, Rhizi and Knownodes. Afterwards, I describe the different pilot projects we did to test the software in a real organisational setting and provide lessons learned from each pilot iteration. This chapter includes an in-depth review of the use-case within the IT department of Descartes university, which has been submitted for publication.

The Met and unmet challenges chapter analyses the results obtained compared to the problems and challenges section in the introduction. It also offers a perspective how to address some of the unmet challenges in a short proposal for a future project.

Introduction

"The tension of our times is that we want our organizations to behave as living systems, but we only know how to treat them as machines."

- Margaret J. Wheatley & Myron Kellner-Rogers (1996)

My fascination with knowledge sharing technologies started from a personal place and at an early age. Growing up, I suffered in classrooms. The world is full of curious and wonderful things to learn and do, why must we all, students and teachers, be coerced to learn and teach in rigid ways that take the fun out of education?

As Ivan Illich said "A good educational system should have three purposes: it should provide all who want to learn with access to available resources at any time in their lives; empower all who want to share what they know to find those who want to learn it from them; and, finally, furnish all who want to present an issue to the public with the opportunity to make their challenge known."

To figure out how I can contribute effectively, I wanted to understand the challenges within the field of education. I wanted to know the state of the art - what were the questions people were currently asking? Which paths in research are worth following, and which lead to dead-ends?

I quickly realized how difficult it is to both understand and communicate the state-of-the-art. It is difficult to learn because despite the vast amount of knowledge online, published research and Wikipedia are about things we already know, and hardly what questions we should be asking. Despite the vast amount of knowledge stored in our global brain, the most insightful parts of my journey happened through a direct interaction with a human.

These great human interactions happened in research centers or hackerspaces. With people who are parts of teams that are committed to work on long-term projects. I got several important things that we the online world cannot yet replace: (i) *Knowledge and skills*: People sharing with me books, research and projects of note. Becoming aware of the research questions people in several scientific domains were asking. Criticising old ideas, as well as developing new ones. And finally pointing me to scientific and project management methods I can use for my own project. (ii) *institutional knowledge*: Helping me find good internships and scholarship opportunities, preparation for the various selections stages, finding good interns, employees, thesis advisors and co-founders. (iii) *identity*: The most important of all. When directly interacting with people, you get a real sense of them. There's a feeling of community, of working together towards a shared goal that is hard to emulate otherwise.

Is there a way to *scale* these kind of interactions so that others, that are not as bold or privileged as me, can also go on their own journey? Is there a way for us humans to share our understanding with the world in a way that is not lost in jargon, or hidden like a needle in a haystack of information overload?

These questions form the base of my thesis. My goal is to help jumpstart the unique emotionally engaging and context sensitive interactions that cannot yet be “eaten” by software. To help people become more acutely aware of the experience, capabilities and intentions of other humans around them.

This aim of this project is to explore ways to transform organisations to become more adaptive and innovative by using human capital mapping software to:

- Empower each member to become aware of the human capital and opportunities available to him throughout their organisation.
- Provide those with the desire to innovate and learn with the means to do so.
- Lower the energy barrier of collaboration across teams.
- Help people feel more comfortable to ask for help

Spoiler - I did not manage to solve all these lofty problems. I did, however, write an interesting report about my discoveries during the journey: (i) A review and analysis of the state of the art. (ii) Sharing the methods I used to incorporate lean startup methodologies within the context of a thesis project. (iii) Reporting on the result of prototyping Knownodes and Rhizi as tools for mapping knowledge, projects and human capital within organisations. (iv) A proposal for a ticketing system for people to directly help each other, integrating within it some of the lessons learned from the previous parts.

Welcome aboard!

Problems and Challenges

If we wish to empower individuals within organisations to become aware of the resources available to them, we need to understand the reasons why this is not the case in the first place.

In the following chapter I will first review some of the obstacles related to the challenge of human capital awareness across large organisations. Segment [1.1.1 Why human capital awareness?](#) deals with the question why human capital awareness is important and how it is related to the well-being and efficiency of organisational work and its' members. Segment [1.1.2 What stands in the way of organisational awareness?](#) presents a more conceptual overview of the barriers as identified by the literature. Segments [1.1.3 Human factors limiting organisational self-awareness](#), [1.1.4 Structural factors limiting organisational self-awareness](#), [1.1.5 Technical factors limiting organisational self-awareness](#) looks on the challenge from different perspectives.

Why Human Capital Awareness?

While there are many different types of benefits I think might be associated with human capital awareness, I will focus on two that I think are pertinent to almost any large organisation: Work satisfaction and productivity.

Work satisfaction

Rigid organisational structures may limit the formation of collaborative relationships between its' members. The generation which grew up in the internet age, and educated based on 21st century education values such as critical thinking, entrepreneurship and system-level thinking are at the start of their careers. They work within organisations that for the most part, do not provide them with the sense of agency they value.

According to Deloitte survey of Millennials (People born between 1980 and 1995) 66% expect to leave their organisations by 2020, 71% of those likely to leave within the next two years are unhappy with how their leadership skills are being developed, and only 28% of Millennials feel that their current organizations are making 'full use' of the skills they currently have to offer.

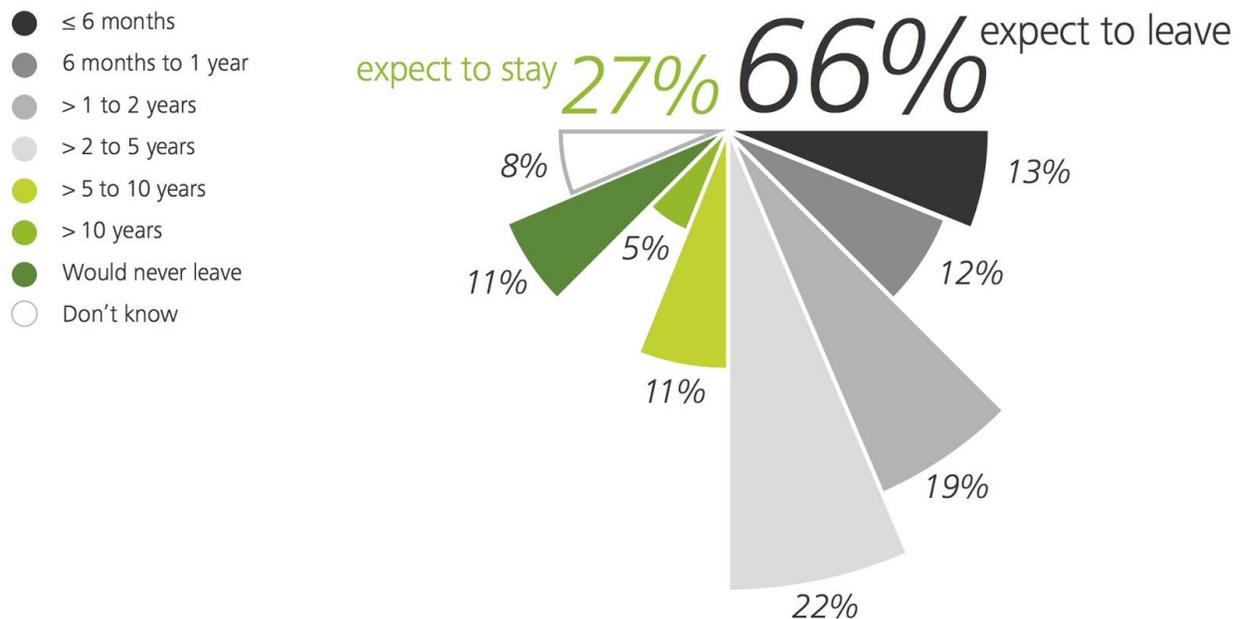
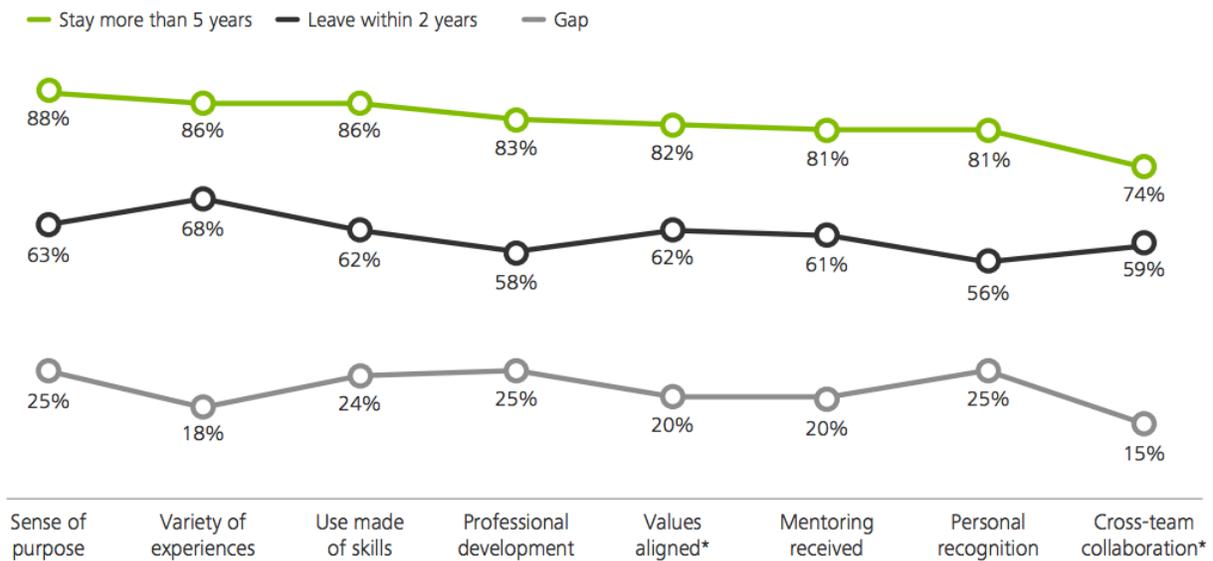


Figure 1: Percentage of millennials expecting to leave the organisation (Deloitte, 2016).

They are more likely to report high levels of satisfaction where there is a creative, inclusive working culture (76 percent) rather than a more authoritarian, rules-based approach (49 percent). More

specifically, in organizations with high levels of employee satisfaction, Millennials have a much greater tendency to report: Open and free-flowing communication (47 percent versus 26 percent where employee satisfaction is low); A culture of mutual support and tolerance (42 percent versus 25 percent); A strong sense of purpose beyond financial success (40 percent versus 22 percent); The active encouragement of ideas among all employees (38 percent versus 21 percent); A strong commitment to equality and inclusiveness (36 percent versus 17 percent); and support and understanding of the ambitions of younger employees (34 percent versus 15 percent) (Deloitte, 2016).



Q. How satisfied are you with the following aspects of your current working life?

Figure 2: Gaps in job satisfaction issues between millennials who are leaving or staying in their workplace (Deloitte, 2016).

Schools, military, universities, governmental institutions, corporations and SMBs (Small and medium businesses) are slow to evolve. A brain-drain is created as those who seek to innovate either seek out opportunities with the few organisations that do provide these conditions such as start-ups, resource-rich companies who have invested tremendously in collaboration practices and technologies, or just adapted to the situation and simply gave up on fulfilling their creative potential within their work life. This brain-drain, besides the individual dissatisfaction that it breeds, ultimately causes a flawed allocation of resources on a societal level. The organisations with the highest needs of skilled innovative workers such as education, science, health, energy and governance stagnate while workers seek to be employed elsewhere. While higher compensation could certainly be a strong motivator to attract talent to these areas, how much of it is caused by the frustration of innovators with inefficient and politicized hierarchies?

How can we gradually transform existing organisations to become more innovative and start a benevolent loop and attract the right talent?

Productivity

“If HP knew what HP knows, we would be three times as profitable” - Former HP CEO Lew Platt

Lack of awareness of human-capital creates bottlenecks that can slow down business processes, hinder innovation and often lead to suboptimal and redundant work.

It has been estimated that at least \$31.5 billion are lost per year by Fortune 500 companies as a result of failing to share knowledge (Babcock, 2004).

Another example for the tremendous costs of blocked collaboration is expressed in a 2014 survey of 152 Campus IT leaders in the US. The leaders estimated that 19% of the campus IT systems are redundant, costing US universities 3.8B\$ per year(Cloud Campus, 2015). This waste can be substantially reduced if IT employees were aware of the existing solution already implemented and the people who have experience with them, especially since many of IT solutions today are open-source.

In aggregate, knowledge workers find the information they need only 56% of the time. People spend as much as 56– 65% of their working time communicating to obtain and supply information (Pinelli, Kennedy, & Barclay, 1991; Robinson, 2010). thereby making source selection important to spending this time effectively, or to reducing it by removing barriers to expertise seeking. This time waste alone amount to an approximate cost of \$7K per employee annually in enterprise (Schubmehl & Vesset 2014). Large proportions of this time could be saved if knowledge seekers had access to an available expert with the relevant knowledge, and, from the other end, the knowledge provider would be incentivized to dedicate time to provide the help needed.

Gabriel Szulanski found that good practices could linger unrecognized for years within companies. Even when being recognized, in-house best practices took an average of 27 months to find their way from one part of the organization to another (Szulanski 1994). Hansen and Noharia (2004) report how through using inter-unit collaboration for the transfer of best practices, British Petroleum reduced costs: A business-unit head in the United States sought to improve the inventory turns of service stations. Tapping the expertise of her peer group, she obtained knowledge of best practices from operations in the United Kingdom and the Netherlands, leading to a 20% decrease in working capital needed by U.S. service stations. Buckman Laboratories’ transfer of knowledge and best practice system helped push new product-related revenues up 10%, a 50% increase since 1992.

Efforts in an oil rig saved \$5M just from mapping experts who could advise on the spot when a technical issue occurred (Cross et al, 2006).

Another example of waste created is how newcomers and remote employees may take up to two years to fully integrate in their department, not only they miss out on knowing who to consult, their own talents and expertise is slow to be exposed to the rest of the department. This process can become faster if there people were collectively aware knew who knew what and who works on what project.

O'Dell and Grayson (1998) enumerate some gains made by enterprises adopting internal best-practices: "Texas Instruments generated \$1.5 billion in annual free wafer fabrication capacity by comparing and transferring best practices among its existing 13 fabrication plants. Kaiser Permanente's benchmarking of internal best practices helped drastically cut the time it took to open a new Women's Health Clinic and it opened smoothly, with no costly start-up problems. By comparing practices on the operation of gas compressors in fields in California, the Rockies, and offshore Louisiana, a Chevron team learned that it could save at least \$20 million a year just by adopting practices already being used in the company's best-managed fields. Chevron's network of 100 people who share ideas on energy-use management has generated an initial \$150 million savings in Chevron's annual power and fuel expense by sharing and implementing ideas to reduce company wide energy costs. By 1996, Chevron could credit its best-practice transfer teams with generating over \$650 million in savings. Most large consulting firms have built huge systems for capturing and transferring internal engagement information and practices to consultants so they can sell projects and help clients design new approaches built on best practices.

Beyond quantifiable costs to be saved, there is a great opportunity for self-aware organisations. The potential to motivate organisation members, attract and retain talent and enhance innovation and learning. According to one study, half of performance gains in a business come from collaboration (The future of corporate IT 2013). In their literature review, Wang and Noe conclude that "Research shows that knowledge sharing and combination is positively related to reductions in production costs, faster completion of new product development projects, team performance, firm innovation capabilities, and firm performance including sales growth and revenue from new products and services (Arthur & Huntley, 2005; Collins & Smith, 2006; Cummings, 2004; Hansen, 2002; Lin, 2007d; Mesmer-Magnus & DeChurch, 2009)." Hansen and Nohria mention five categories of benefit in which multinational corporations can benefit from inter-unit collaboration. "Cost savings through the transfer of best practices; Better decision making as a result of advice obtained from colleagues in other subsidiaries; Increased revenue through the sharing of expertise and products among subsidiaries; Innovation through the combination and cross-pollination of ideas; and Enhanced capacity for collective action that involves dispersed units."

What stands in the way of organisational awareness?

According to research, large organisations are bad at identifying and leveraging internal organisational knowledge (Alavi and Leidner 2001; Davenport and Prusak 2000; Evans and Ali 2013; Hibbard 1997, Hinds and Pfeffer 2002; Nevo, Benbasat, and Wand 2009, 2012; O'Dell and Grayson 1998).

Best practices are professional procedures that are accepted as being correct or most effective. They serve as a good example for leveraging internal knowledge because many companies are motivated to spread those around (See [1.1.1 Why human capital awareness?](#)). Gabriel Szulanski, in his work with several large firms, found several barriers to the transfer of best practices (Szulanski, 1994).

The first is ignorance. At most companies, particularly large ones the ignorance went in both directions: The employee possessing the knowledge was not aware that someone is interested in the knowledge they had while the employee receiving the knowledge was not aware that someone possessed the knowledge they required. The most common response from employees was either "I did not know that you needed this" or "I did not know that you had it" .

The second was the absorptive capacity of the recipient: Even if a manager knew about the better practice, he or she may have had neither the resources (time or money) nor enough practical detail to implement it.

The third barrier was the lack of a relationship between the source and the recipient of knowledge - i.e., the absence of a personal tie, credible and strong enough to invest in listening or helping each other.

In figure 3 below, I've synthesized the different barriers as a simple model which reflects how these barriers are encountered when trying to start a typical collaboration. First, the collaboration could not happen if you are ignorant of it can do not even know there is someone to collaborate with. Second, if you do not trust or relate to the other person, you will not collaborate with him. Third, if you find that the cost benefit of the transfer does not make sense, you will avoid it.

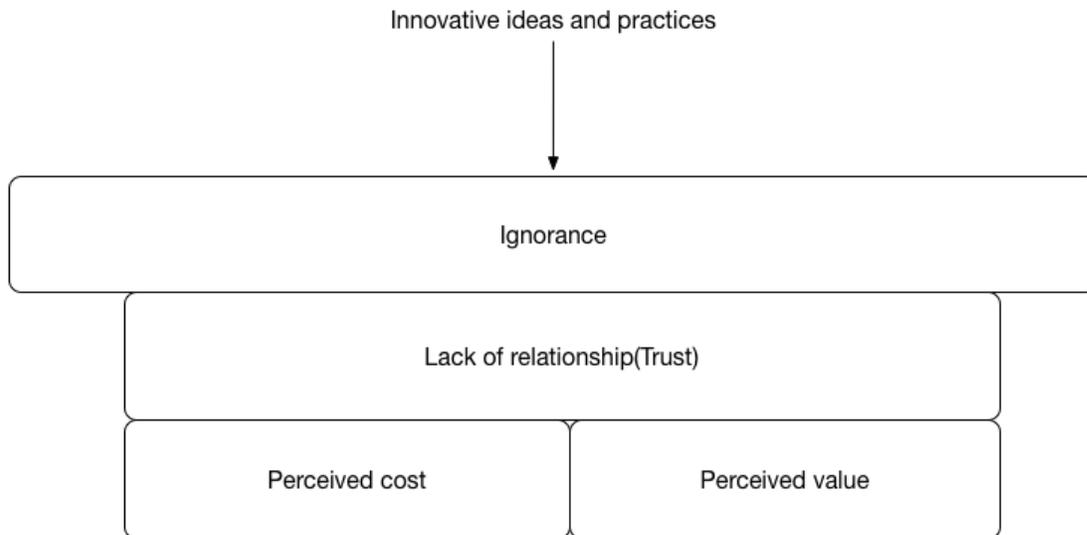


Figure 3: Synthesis of the different barriers towards diffusion of best practices

Manual Castells (2010) coined the term “self-programmable labor”: “Self-programmable labor has the autonomous capacity to focus on the goal assigned to it in the process of production, find the relevant information, recombine it into knowledge, using the available knowledge stock, and apply it in the form of tasks oriented towards the goals of the process. The more our information systems are complex, and interactively connected to databases and information sources, the more what is required from labor is to be able of this searching and recombining capacity. This demands the appropriate training, not in terms of skills, but in terms of creative capacity, and ability to evolve with organizations and with the addition of knowledge in society. On the other hand, tasks that are not valued are assigned to generic labor, eventually replaced by machines, or decentralized to low cost production sites, depending on a dynamic, cost-benefit analysis.”

The barriers described above can also be considered as barriers facing the self-programmable workers to effectively collaborate.

Human Factors Limiting Organisational Awareness

Collective Stupidity

There are certain human traits that have been identified by research as limiting in our efforts to become aware of others. Such biases that are referred to as “Collective stupidity”. For example, the “hidden profile” phenomena, revealed by Stasser, shows that humans do a poor job when trying to synthesize different facts from people in a group discussion. The decision quality was improved when the facts and expertise of the group members were explicitly shared, but that did not happen without an outside intervention (Stasser & Titus 2003). We would like to believe that a team of knowledgeable individuals would share and absorb different perspectives before deciding on complex issues, but this is not a natural occurring process¹.

Another factor raising some inherent issues in human cognition is Dunbar’s number. Dunbar theorizes that there is a social limit to the awareness that one can have about others is limited to approximately 150 individuals ([Hill & Dunbar, 2003](#)). Lack of organizational awareness becomes even more acute when considering the increase number of those that work remotely or at separate locations other than the central office.

Other factors related to collective stupidity are:

- The hunger for confirming what we already believe(Confirmation bias)
- Distortion of facts to serve self-interests
- Control and optimism bias
- Suspension of moral responsibilities or over-confidence.
- Groupthink - When members of a group prioritize getting along together over critically assessing ideas.

¹ Numerous variations of the classic Stasser experiment tried to negate the “hidden profile” effect, there were only two variations that proved to be extremely effective, both leading to a 41% increase in the team making the right decision, and both connected to priming the subjects with a preliminary task before the main decision making task. Postmes et al. (2001) replicated the Stasser experiment but (a) The unique information was highlighted and explicitly discussed (b) Two optional priming tasks were given to the decision making team, one was based on a consensus norm of “getting along” and the other based on “critical thinking”. 22% of the groups primed with a consensus norm selected the best option in contrast to 67% of groups primed with the critical thinking norm. Galinsky and Kray (2004) used a counterfactual mindset to achieve similar results. Where they primed groups with a scenario that promoted “considering alternative possibilities” and a similar scenario without the “considering alternative possibilities” priming. The former groups solved the task correctly 67% of the time vs 23% by the latter group.

Human factors within organisations

When the perceived cost of generating and maintaining organisational awareness outweighs the perceived benefits and rewards obtained from it, the process is most likely to fail. Furthermore, the cost-benefit equation must also make sense for each specific group of actors within the organisation: from the “rank and file” employees, to middle-managers and executives.

So what qualities inherent in human behavior increase the perceived cost and reduce the perceived benefit? Hinds and Pfeffer(2003) identify cognitive limitation and motivational limitation to the transfer of expertise.

Cognitive limitations

Research shows that experts demonstrate a rather poor ability to share their expertise. In their analysis Hinds and Pfeffer explain that such poor ability is due to the several cognitive limitations.

The first limitation is that as expertise deepens, experts tend to represent their understanding in more simplified and abstract ways. For example, Gitomer(1984) found that when skilled experts viewed an electronical device, they used conceptual models of the way the device worked, while the less skilled described the device as a collection of unrelated components and spend more time using trial-and-error procedures. Other examples such as using higher level abstract concepts in programming(Adelson 1984), physics(Chi, Glaser, & Rees, 1982) and other research(i.e., Ceci, & Liker, 1986; Gobet & Simon, 4 1998; Johnson, 1988; Lamberti & Newsome, 1989; Chase & Simon, 1973; McKeithen, Reitman, Rueter, & Hirtle, 1981) suggest that expertise is characterized by conceptual and abstract representation of the subject matter. In other words, expert terminology tends to phrase and represent information in a way that makes it inaccessible for the non-expert, thus limiting knowledge propagation. An experiment done by Langer and Imber (1979) uncovered that when examining lists of task components those made by experts contained significantly fewer and less specific steps than did the lists of those with less expertise. From the expert’s point of view, optimizing and removing “obvious” steps makes sense, but as a side-effect, this complexifies the transferral of expertise. Making the problem even harder, experts tend to under-estimate the performance time of novices(Hinds 1999). “The curse of knowledge”, is a term coined by Camerer and his colleagues (Camerer, Weber, & Loewenstein, 1989) and describes a bias experts have in estimating the point-of-view of the non-expert. In their experiments, when those informed with the economic state of a company showed a bias towards their own knowledge when trying to predict how the less-informed would assess the valuation of the company. Even following feedback and debiasing methods, the experts failed to adjust their predictions appropriately.

The research done suggests that translation from “expert-speak” to “novice-speak” requires establishing common ground between the audiences. “The curse of knowledge” makes the process of finding common ground an expensive multi-iteration processes. I can imagine the reader might have encountered the rather common frustration caused by under-valuing the amount of attention needed for successful knowledge transfer.

A second cognitive limitation is the challenge of articulating tacit knowledge. Tacit knowledge is learned through experience and held at the unconscious or semiconscious level (Polanyi, 1966; Leonard & Sensiper, 1998). Experts, in addition to assessing the competency of others, also need to expend additional effort to understand what is the missing tacit knowledge and how to articulate it. Another issue is that a lot of the tacit knowledge is often environment and context-specific. Codifying knowledge in a way that makes it useful within multiple contexts and environments is hard work, often requiring a different set of skills that the expert possess. The risk faced by the organisation, is taking away the expert’s attention from resolving tangible tasks at hand into an intangible task of translating his knowledge to others (and often making a poor job at that).

Motivational limitations

A crucial element of successful knowledge sharing is that employees actually want to contribute to these processes (Cabrera et al., 2006; Wang and Noe, 2010).

The inherent cost of transferring expertise via documentation of conversation must be balanced by an incentive system, unfortunately, few organizations provide the time required for knowledge transfer, believing that “conversations” are not real work. In general, companies want to see a return on their investment in transferring skill and knowledge but are not willing to adequately compensating employees for their time doing it.

Hinds and Pfeffer raise the issue that the compensation structure of companies often pits people and departments against each other. The promotion and pay raise compensation model, as well as rewards systems such as “employee of the month”, have the side effect of inducing competition as individual performance is measured relative to the performance of colleagues.

The human tendency to identify with your own group and develop an in-group bias (Abrams & Hogg, 1990) results in higher levels of inter-group conflict and reduces cooperation across the organisation (Kramer, 1991).

Performance measurement is often related to tasks and interactions within the core-team. Helping others outside the scope of a given set of tasks is seldom measured. This generates a clear incentive structure that promotes siloed teams.

The “Knowledge is power” equation is a cultural belief that being the source of knowledge, or becoming a middleman or broker between parties increases an individual’s power and positioning within an organisation. Monopolizing knowledge becomes a political leverage point as those who possess the knowledge can accelerate or slow a specific initiative. An exaggerated yet illustrative way to describe this is to imagine a plane crashing on a deserted island, and the survivors needing to negotiate and decide how to organise themselves. If some of the group don’t speak the same language, the multilingual within the group would become the language brokers and gain the power to decide which information is passed to the other sub-group and in what style.

Beyond mistrust in individuals, lack of trust towards the organisation can inhibit knowledge sharing. Would the information you share with the organisation be used against you or other employees? There is evidence that organizational actions that destroy trust, such as downsizing, induce fear and make the transfer of expertise and experience less likely (Davenport and Prusak, 1998; Pan & Scarbrough, 1999; Pfeffer and Sutton, 2000).

The norm of reciprocity describes the tacit exchange that happens when you request help from another. When you ask for a favor, you also “owe” the other person a favor in the future. The bookkeeping of the network of who owes what to whom is informal and imprecise. Many prefer to avoid procuring this “debt”, especially since it is implicit and without clarity of what would be the favor asked in return (Hinds et al 2001, Hollingshead, Fulk, & Monge).

Another issue worth addressing can be phrased as “the tragedy of the help-seeker”. People think that they will be regarded as less capable if they ask for help, when in fact, within reason, seeking help actually improves how colleagues perceive you (Brooks et al. 2015)

Formal and rule-based methods for knowledge-sharing have various technical aspects that are specified in 1.1.5 Technical factors limiting organisational self-awareness, but there are also human motivational factors in play. Filling forms and following strict sets of rules can make knowledge sharing a lot less satisfying than informal ways. In addition, Reactance theory (Brehm, 1966) suggests that forcing people to do something may produce exactly the opposite result, as people rebel against the constraints imposed on them.

To synthesize this part, figure 4 is a simplification of the different barriers and lack of incentives existing for sharing knowledge about human capital.

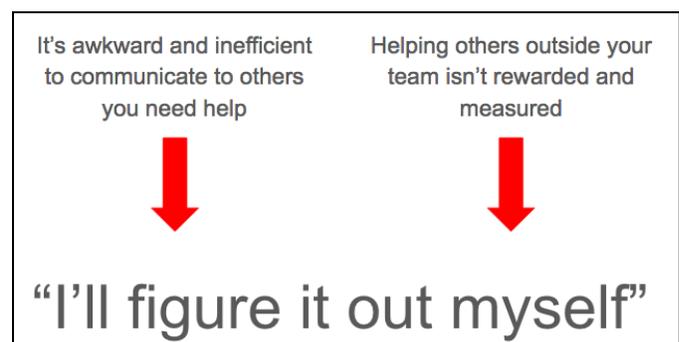


Figure 4: Simple breakdown of perceived cost and lack of incentives for organisation wide collaboration.

Structural Factors Limiting Organisational Awareness

Structural factors inherent in organisations themselves may hinder organisational awareness. Structural factors refer to intra-organisational structural factors such as governance systems, and legal constraints.

Hierarchies are an organisational pattern found in many complex systems including biological, ecological, information systems and social structures. The paper “The Evolutionary Origins of Hierarchy” (Mengisty et al 2016) describes how the cost of connections leads to the evolution of systems to optimise towards hierarchies.

It is no surprise this structure is prevalent in human societies, as the cost of creating and maintaining trust, communication lines, and understanding between individuals is indeed a costly process. “In organizations such as militaries and companies, a hierarchical communication model has been shown to be an ideal configuration when there is a cost for communication links between organization members(Guimera et al, 2001).” Hinds and Pfeffer assert that “Formal hierarchies have traditionally served the purpose of coordinating and making more efficient the flow of information in organizations (Aldrich, 1979, Cyert & March, 1963, Simon, 1962). This is accomplished through a division of labor in which functionally specialized units and unity of command constrain communication flows to those defined by the chain of command (Galbraith, 1973). By constraining communication so that instructions flow downward and information flows upward, organizations are made more efficient and predictable”

The thriving of collaboration practices in open-source collaboration, Wikipedia, social networks and real-time collaboration software such as Dropbox and Google Drive, are driving down the maintenance cost of connections with the aid of technology as well the perceived cost of adopting new practices as more and more people are becoming familiar with practices around non-hierarchical collaboration.

Few studies have investigated how organizational structure impacts knowledge sharing in public and private sector organizations(Kim and Lee 2006), but there is evidence that even if the structure of the organization is hierarchical, but it permits the people to access each other when they require desired knowledge, the hierarchical structure does not hinder the transfer of knowledge (Fahey & Prusak 1998).

In this segment I'll mention some of the “baggage” that traditional organisational hierarchies have. It's useful to reflect which are inherent to hierarchical structures, and which are coupled to habits, cultural artifacts, technical limitation or human practices that can be improved on.

Reluctance and sanctions against bottom-up information and norm sharing - Leonard and Sensiper (1998) describe how nurses become reluctant to suggest patient treatments because of the higher-status of the physicians. Similarly, in a study of operating room teams, Edmonson (2000) describes how low status employees were often reluctant to share their expertise with surgeons because of their negative response to advice from lower status team members. This lead to errors in the operating room. The lower status team members frequently had a better understanding of some aspects of the new technologies and procedures. Beyond status, rigid roles in hierarchies can sometime lead to collaborative work becoming a challenge to another employee's authority. Defined roles may obscure interdisciplinary abilities of members.

Preplanned, static division of labor - Following rigid roles prevents people from dedicating time to help others, even when their expertise can save considerable amount of time. In his analysis of the challenges of collaboration, Nielsen(2010) says that creative work often requires being open for the unplanned and unexpected insights that matter the most. It's hard to expect in advance what creative output people can have on a project. Unfortunately, once roles are defined, it takes a lot of efforts to redefine things in rigid structures. To illustrate his argument, Nielsen provides the example of the Kasparov vs the world project - in which a mass of people collaboratively competed against chess master Kasparov. Had predefined roles were assigned to the crowd, a lot of great ideas would have been lost.

Incentive systems lack collaboration measurements outside scope of defined role or task - People in organisations are usually assessed by their superiors and peers based on individual performance or the performance of their subordinates. Leaving little incentives to help others outside their command chain. In hierarchies there is less motivation for non-managers to synthesize their point of view with those of others, and the responsibility for synthesis is on the managers themselves.

In organizations where individual competition is emphasized, employees will not be likely to share knowledge with others at work (Schepers and van den Berg, 2007; Willem and Scarbrough, 2006)

Cultural resistance - Another point raised by Mintzberg et al(1996) is that practices that enabled people to climb to the top of hierarchies such as experience in imposing their will across command-chains, may not be as willing to engage in, or might even be suspect of horizontal and informal collaborations.

Sharratt and Usoro (2003), found that “organizations with a centralized, bureaucratic management style can stifle the creation of new knowledge, whereas a flexible decentralized organizational structure encourages knowledge-sharing, particularly of knowledge that is more tacit in nature” (p. 189). Tsai (2002) discovered that centralization could reduce individual's' interest in sharing knowledge with other units within an organization. When formalization is low in the organization structure, knowledge sharing will increase (Lin, 2008)

Security and privacy concerns - Many organisations are reluctant to share information from the fear of it leaking out, hurting their competitive advantage. Raghatz [19] identifies the consequence of this risk as being a resistance within the organisation to sharing proprietary information with suppliers. But it can also be also be the fear of leaking strategic information about their priorities, or potential targets for hackers or corporate espionage.

Geographical distance - Nonaka [20] identified face-to-face communication as the most efficient means for knowledge transferral, but the geographical location of the organisations may mean that this is not possible. Geographical separation may also result in the organisations working in different cultural, legislative or linguistic environments.

High barriers for entry - Nielsen lists that high barriers for entry in collaboration are high because of the following principles:

- *Task scope* - Tasks are too broad making the initial things an individual needs to know to participate unreasonably high.
- *Conversational critical mass* - Having enough people participate so the info generated is useful.
- *Knowledge commons* - Having enough of the knowledge available is a shared repository to lower the cost of acquiring initial knowledge.

Time constraints - It takes time to transition from a self-serving mode into a collaborative interaction. In their pilot study of a virtual community, Hall and Graham(2004) found that the initial impetus for members to join the group was to discover information for personal benefit. Over time, however, individual desire to reciprocate the help received from the group developed out of the online interactions. This long-term commitment to nurture a collaborative culture can go against the chances of it being adopted, especially since the initial feedback received is of people using new systems for their individual needs rather than for collaboration.

Technical Challenges Limiting Organisational Awareness

Complementing the human and structural challenges, technology has its own set of challenges that limit its ability to facilitate organisational awareness. Below are some of the major challenges.

Unstructured information

For information to be useful, it must be brought to the people that can do something with it.

According to the International Data Corporation (IDC), 80-90% of the organisations information is unstructured (2012 IDC). With limited findability or appropriate context, the likelihood of it being reused is low. For information to be structured, it must adhere to a system of categorization.

Defining the categories and ontologies of these systems is hard, often domain-specific and subject to extensive research. As reality moves forward in time, new categories emerge, business needs change and restructuring the data to fit these changes can be a challenge.

The question what is an ontology can be a cause for disagreement. Martin Hepp(2008), raises three common disagreements:

Truth vs. consensus - Early ontology research was very much driven by the idea of producing models of reality that reflect the “true” structures and that are thus valid independent of subjective judgment and context. Other researchers, namely Fensel (Fensel, 2001), have stressed that it is not possible to produce such “true” models and that instead consensual, shared human judgments must be the core of ontologies.

Formal logic vs. other modalities - For a large fraction of ontology researchers, formal logic as a means (i.e., modality) for expressing the semantic account is a constituting characteristic of an ontology. For those researchers, neither a flat vocabulary with a set of attributes specified in natural language nor a conceptual model of a domain specified using an UML class diagram(Unified Modeling Language, a general-purpose, developmental, modeling language in the field of software engineering, that is intended to provide a standard way to visualize the design of a system) is an ontology. This is closely related to the question on whether the ontological commitment is only the logical account of the ontology or whether it also includes the additional account in textual definitions of its elements.

Specification vs. conceptual system - There is also some argument on whether an ontology is the conceptual system or its specification. For some researchers, an ontology is an abstraction over a domain of interest in terms of its conceptual entities and their relationships. For others, it is the explicit (approximate) specification of such an abstraction in some formalism, e.g. in OWL (Web Ontology Language), WSML (Web Service Modeling Language), or F-Logic (frame logic).

Furthermore, Hepp identifies some of the main challenges for ontology architects as:

Interaction with human minds - Ontologies are the glue between human perception of reality and models of that reality in computers. Therefore it is imperative that ontologies are easy to read and understand. Therefore, visualized and use of natural language is important.

Integration with existing knowledge organization systems - Pre-existing knowledge is stored using traditional systems of knowledge organisation, or in some cases, an improvised solution that was made without interoperability in mind. Therefore, ontologies must often first develop glue layers between these different systems.

Managing dynamic networks of formal meaning - As ontologies are not static conceptual models of “eternal” truth, but artifacts reflecting our gradual understanding of reality, we face the difficulty of managing such dynamic networks of meaning (Fensel, 2001). This is also linked with insuring ontologies are evolvable, insuring interoperability between different ontologies, and linking between the emergent need to modify an ontology (e.g. User-feedback, or modification provoked when using search UI) with the modification of the ontology.

The vocabulary problem - Furnas and colleagues have shown that the chance that two individuals choose the same word for the same thing in human-system communication is less than 20% (Furnas, Landauer, Gomez, & Dumais, 1987). They have basically proven that there is “no good access term for most objects” (Furnas, Landauer, Gomez, & Dumais, 1987, p. 967).

Fragmented information

Beyond unstructured information, there is the issue of fragmented information. This describes the situation in which information is stored in different formats (CSV files, Relational databases, custom text formats, etc.) on disconnected systems. Deriving insights requires manual and expensive human intervention. IDC study states that 61% of knowledge workers access 4 or more systems on a regular basis and 13% need to access 11 or more systems to get the information they need to do their job. This information is locked in a variety of formats in network file shares, content management systems, intranets, bespoke applications, and other repositories that don't talk to each other (Schubmehl & Vesset 2014).

Defragmenting information is a costly process because each of the operation necessary to do it is costly in itself: Storing information, dealing with redundancy, changing data models and having to keep information up to date.

Collaboration requires maintaining a shared state

Computers are very good at making copies, but sharing takes more work: History, Real-time communication, Synchronization. Adding this kind of functionality is a non-trivial technical challenge.

Information overflow

Any knowledge system must provide usable tools for filtering and sorting content according to a person's interest.

Information security and privacy

Both from a regulatory standpoint and practicalities of organisations. Any information system needs some sort of access control - A system to specify and enforce who has access to which data within the database. Furthermore, how much data should be accessible to outsiders?

Background

Following the review of challenges faced with organisational awareness. This chapter provides a review of some background to key concepts important to understand this thesis.

Chapter [1.2.1 Human Capital Within Organisations](#), describes different interpretation and context of the term human capital. Human capital, or in other words, people and their competencies, was discovered to be the central object of mapping through the Rhizi tool.

Many of the ideas for Knownodes and Rhizi stem from trying to integrate practices of collective intelligence - which in which groups can act in an intelligent way. Chapter [1.2.2 Collective Intelligence](#) provides a brief review of the term and some examples of the way it is being used in different contexts.

Network analysis, modelling data as networks and building user interfaces to make working with graph more intuitive for non-experts were one of the aims of the work done. Chapter [1.2.3 The Network Paradigm Within Organisations](#) provides a brief review of network analysis, and its usage within organisational context.

The last background chapter, [1.2.4 Knowledge Management](#), is a review of a very dense academic field that surfaces issues that occur when trying to reform and develop human and technological practices to collect, store and retrieve knowledge within organisations.

Human Capital within organisations

Definition(s)

The term Human Capital originates from the fields of social science and economy. According to economic professor Steve Pischke, these definitions can be divided into the following categories(Pischke, 2015):

The Becker view: The prevailing definition, which was popularized by Gary Becker and Jacob Mincer is that human capital is the human aspect in the economic value chain. It represents the human attributes that enable or enhance labor for the purpose of producing economic value. These attributes may be general or task-specific. Such attributes include: Knowledge, habits, social skills, talent, skill or experience(Becker, 1975). This classic definition sees human capital as a one-dimensional object, a means of production that while not transferable, is it substitutable.

The Gardener view: Originating from Howard Gardener, father of the multi-intelligence theory, this view adds more refinement to the definition as it differentiates between types of skills, each with it's own unique properties in relation to production. For example, physical and mental skills are different, and so are other sub-categories of mental and physical faculties.

The Schultz/Nelson-Phelps view: Human capital is defined by the ability to adapt to changing conditions. And so the more human capital an organisation has, the greater the capacity of its' employees to adapt their practices accordingly.

The Bowles-Gintis view: A critique to prevailing theories of Human Capital. It is defined as the capability to conform to work norms of organisations, obey orders and work in hierarchical/capitalist structures. According to this view, the main objective of schools is to instill these values and behavior within students.

The Spence view: Observable human skills are more signals of competency rather than directly useful to the production process.

I find the Schultz/Nelson-Phelps to be the most compelling in light of the pace of change in present times, it is also the most human-centric as it empowers the humans within the system to become adaptive agent themselves, and not overly dependent on the particular organisation they are part of. As a complementary viewpoint, and in light of interviews describes in chapter [3.5 Descartes II](#), I have found that the Spence view holds a lot of ground “in the trenches”. For the practical purpose

of collaboration, the act of consciously signalling, rather than merely indexing a person skills is a powerful practice to filter out the skills that might exist on paper but are unlikely to be used, as there is a lack of inner motivational power from the human they are attached to.

Different sources of human capital

Human capital can be separated into the following sources:

Innate ability: The Nature part in the Nature/Nurture dichotomy. Inherent physical strength, or, requiring a more nuanced debate, intelligence and other talents.

Schooling: Acquired skills and knowledge

Training: Similar to schooling, but often more domain-specific, related to specific technologies or industry.

Pre-labor market influences: The environment and peers you interact with influences the experience, outlook, knowledge and habits of individuals.

Context within organisational capital

In order to understand human capital within organisations, it is important to first place it within the context of the total capital an organisation has. See figure 5.

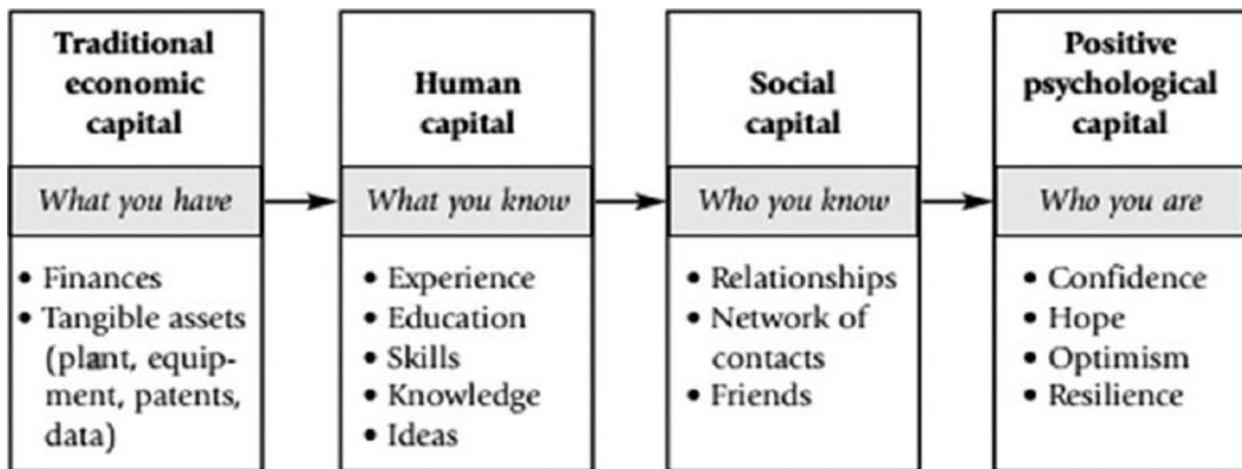


Figure 5: Adapted from *Expanding capital for competitive advantage*. *Business horizons* 47/1 January–February 2004 (45–50)

We can divide the total organisational capital into:

Human capital - As defined above.

Financial capital - Represents obligations, and is liquidated as money for trade, and owned by legal entities. It is in the form of capital assets, traded in financial markets.

Social capital - Partly captured as goodwill or brand value in enterprise, but is a more general concept of inter-relationships between human beings having money-like value that motivates actions in a similar fashion to paid compensation. Includes also relational capital and

Instructional capital - Aspect of teaching and knowledge transfer that is not inherent in individuals or social relationships but transferrable. Books, manuals, videos etc'.

The term *Intellectual capital* is also often used as a theoretical construct invented to account for all the assets not depicted in a company's balance sheets (Brooking 1996). This describes the sum total of Human capital, Structural capital, Relational capital and Strategic Alliance capital. For the purpose of this thesis, there is no need to elaborate on those.

Human capital + Collaboration protocol = Organisational Awareness

As the famous Sociologist, Robert s. Lynd, wrote: "People need to rebuild organisations so that knowledge flows freely to create opportunities and solve problems. "(Lynd, 1939). Adopting a more network-thinking perspective, human capital should not be looked at as a static resource, but as a network. What is the best protocol for managing the network of interactions so that more value is produced? Organisation Awareness can be the result of an organisation successfully leverage its' knowledge resources and people through a collaboration protocol.

Collective Intelligence

Definition

Collective intelligence can be defined as the ability of a group of agents to think and act intelligently, whether from collaboration, competition or joint effort. Used in sociology, business, computer science and mass communications. The word intelligence comes from the roots *inter*, meaning ‘in between’, and the word *legere*, meaning ‘choose’. This refers to our ability to decide between choices. The word collective comes from the word *colligere*, composed of *col*, meaning ‘together’, and *legere*, ‘choose’. This refers to our ability to choose who we want to be with. Joining the words together, means our ability to choose our companions and how we act together. Merging these two definitions together, Collective Intelligence is the ability to properly use the combined faculties of our intelligence to make choices within our chosen group.

A more narrow way to define collective intelligence is to refer to it as an emergent property between people and ways of processing information (Glenn 2008). This allows a wide range of activities to be included in the definition such as:

Wisdom of the Crowds - Information is aggregated from a group of people, for example, [Francis Galton](#)'s describes how the weight of an Ox can be more accurately guessed by averaging individual guesses made by a crowd compared to the individual estimates of most crowd members (Galton 1907).

Politics - How views of people are formed into policies, selection of candidates and run campaigns.

Collaborative editing - How platforms like Wikipedia, or systems of peer-review allow to enhance texts.

Search engines and social media - The human collaboration act of placing hyperlinks on HTML documents was the data on which Google applied its famous PageRank algorithm to sort search results (Brin & Lawrence 1998). How users activity on Facebook influenced sorting of information in their news feed.

Beyond the neutral definition, in order to understand how to increase collective intelligence, we must be mindful of the powerful evidence of “*collective stupidity*”, meaning the shortcomings humans have as a result of failures in communicating, deciding and acting together. A few of these shortcoming such as hunger for confirming what we already believe (confirmation bias) (Plous & Scott 1993), distortion of facts to serve self-interests, control and optimism bias (Shepperd et Al. 2002), suspension of moral responsibilities or over-confidence. These will be expounded in *part 1.2.1 Human factors for limiting organisational self-awareness*.

Collective intelligence has developed a lot thanks to technologies and practices that lowered the energy barrier for larger scale and more effective communication. The printing press, the internet, collaboration technologies such as the wiki, real-time collaborative writing and MOOCs and lingual translation technologies. Those are all means in which the production, spread and co-creation of information requires less energy. The changes go beyond improvements in efficiency - It had changed the very activity of thinking and made dialogue more available, promoting less violent interactions between individuals, groups and communities.

Collective intelligence has the potential to not only communicate “what is” but also - and consequently - knowing “what could be”, in other words, what is our collective intentions.

Measuring Collective Intelligence

MIT center for Collective Intelligence has defined their main research question as “How can people and computers be connected so that—collectively—they act more intelligently than any person, group, or computer has ever done before?” ([ret 2016](#)) In a paper published in Science in 2010, a team of researchers defined the collective analogue to an individual's *G factor* (greatly related to IQ), the “c factor”, intended to predict a group’s performance in a variety of tasks and concluded that “c is not strongly correlated with the average or maximum individual intelligence of group members but is correlated with the average social sensitivity of group members, the equality in distribution of conversational turn-taking, and the proportion of females in the group.” (Wooley et al. 2010)

The reasons for why one group has a different *c factor* than another is hypothesized to originate from bottom-up factors inherent to the team members (people’s ability to judge others’ emotions, individual intelligence, cognitive diversity) and top-down interactions related to structure (equality in participation, amount of communication). See figure 6.

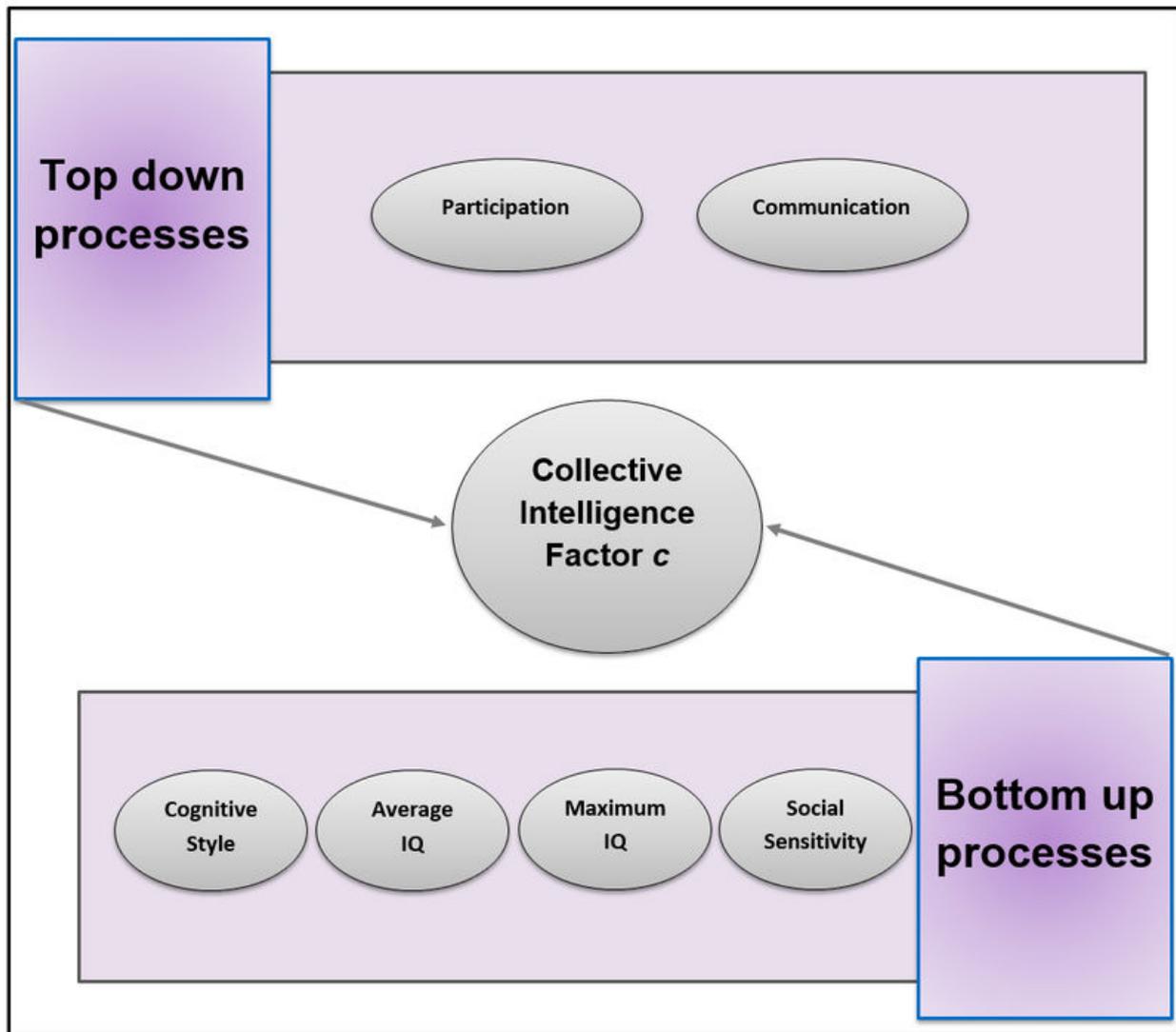


Figure 6: Predictors for the collective intelligence factor c . Suggested by Woolley et Al. 2015.

Applications of Collective Intelligence

Collective intelligence is a vast field in which uses different approaches and technologies, to solve different goals for different collectives. Below is a brief review of several applications that are relevant to this thesis:

Human Computation

Modern human organisations experience a high throughput of projects and of humans that make the cost of tracking available resources and attention difficult. Human computation defined as “information processing systems in which humans participate as computational elements” (HCJ, ret 2016) mostly focuses on the act of collecting information and mapping it so it can be processed by a

computer, leading to immense amounts of data, which in turn could be structured by humans and computers (Quinn & Bederson 2011).

Human computation approaches for solving information fragmentation problems have a rich history such as using semantic-games to structure meta-data (Siorpaes & Hepp 2008). Methods in human computation are often used to solve important scientific challenges as diverse as protein folding (Cooper et al. 2010) and training AI's (Lenat et al., 1990). Knowledge Collection from Volunteer Contributors (KCVC) set the goal to advance artificial intelligence research by using humans to build large databases of common sense facts. Several efforts have demonstrated ways of using volunteer contributors to provide such facts, either by using games such as Verbosity (Von Ahn et al. 2006), 1001 Paraphrases (Chklovski & Timothy, 2005), or plain volunteerism such as Learner (Chklovski, T. 2003) and Open Mind Common Sense (Singh et al., 2002).

Citizen Science

Citizen science is a method in which the general public help the scientific effort by their intellectual effort or their surrounding knowledge or their tools or resources (Socientize Project 2013).

In his policy for the Wilson Center, Pr. Muki Haklai lists the impact and current engagement of people in citizen science: “over one million people participated in classifying images of galaxies, listening to bat calls, transcribing World War I diaries, and identifying animals in the Serengeti in a Zooniverse project. In Germany, in 2012, scientists collaborated with 5000 people to capture over 17,000 samples of mosquito, resulting in the discovery of an invasive species (Asian bush mosquito) with implications to public health as they can carry certain diseases. And in the early months of 2014, a team of climate researchers at Oxford University, who wanted to suggest the degree to which recent floods could be attributed to climate change, was able to run over 33,000 models using the unused computing resources of over 60,000 volunteers.” (Haklai 2015)

Projects such as MathOverflow, the Polymath project and the rise of scientific discovery games such as Galaxy Zoo and Foldit have been heavily dependent on practices of collective intelligence, as they had to figure out how to organise the efforts of the group, how to engage the community around the project, and how to insure the flow of ideas to and from the community.

Science of Team Science

The organisational context inside academic institutions is very different, characterized by an extremely fragmented and mostly decentralized environment with rapid transitions of students and researchers. Projects start, pivot, and end as new discoveries and problems are unveiled. Science of team science (SciTS) encompasses strategies to understand and evaluate processes and outcomes of

collaborative research (Stokols, Hall, Taylor, & Moser 2008). For example, providing evidence that demonstrates that impactful research comes from inter-institutional teams (Börner et al. 2010).

Collective Intelligence Within Organisations

It seems that all institutions and large organisation struggle with the concept of collective intelligence. It is not easy for organisations to shift from viewing intelligence as something happening outside a single mind, but with today's social networks, sensors, virtual communities, it's become more apparent that intelligence happens in the space between individuals - not only inside them. Despite several avant-garde projects, few democracies and cities really collect insights from their populations, and as a result, usually don't seem to be as smart as the societies that they are supposed to govern.

The notion that just by the fact we are networked, some self-organising principle will make us smarter has some grain of truth in it, but it is also misleading. Getting the protocols of the network right, developing practices, changing habits, developing technology and investing time, resources and maintenance are an essential part of allowing the proper flow of intelligence within organisations. We can build a world where, by using technology, people can realize more of their potential, as well as recognize and benefit from the potential of others.

The Next Generation of Collective Intelligence Systems

Existing systems of decision-making such as representative democracy and the market, use rather crude mechanisms for capturing the will of collectives. Voting between candidates or the binary choice of purchasing or not purchasing using a single currency leaves a lot of power in the hands of few intermediaries. New systems for direct democracy and decentralized government are being piloted over the world. [Crowdsourcing and open innovation platforms](#), and [The DAO](#) are notable examples.

Organising Collective Intelligence

Geoff Mulgan lists five principles for organising collective intelligence:

Autonomous Commons - A culture which values and provides necessary space for free reign of developing ideas, using critical thinking and seeking alternatives as mechanisms of refinement , permission to change one's mind and less sense of ownership.

Balancing intelligence faculties - The ability to give proportionate attention and synergy between different types of intelligence such as observation, focus, memory and creativity.

Focus - The ability to differentiate between what matters and what doesn't.

Understanding granularity - What's important at different scales.

Reflexivity - Ability of a group to “know itself, to reflect about it’s own thought processes and patterns and make them explicit as part of a knowledge commons.

Integrative thinking - Ability to translate data and thinking methods into decisions and actions. In the words of Geoff Mulgan - We complicate to understand, but simplify to act.

Network Paradigms Within Organisation

Networks - Definition and Scope

Scott et al. (2015) Describe networks as “Possibly the defining metaphor of our time”. From neural networks to terrorist networks, and computer networks - networks consist of nodes (entities) and edges (relationships between entities).

Networks can be measured by the properties of their structure such as size or density or specific nodes in a network can be understood by its centrality in the network, for example, the amount of outgoing connections or connections from other nodes.

Social network analysis is concerned with examining the structure of relationships between social entities. A node can be a person, a concept, an idea, a physical place, etc. Relationships between entities (also referred to as edges, connections or links) can be of different types such as a physical, associative, or a social relationship. An organisation can be described as a network of people, projects, teams, documents, processes, materials and spaces. Between them, materials, energy and information can flow. Similar to biological networks, which exist in different levels of networks, such as, interactions inside cells, between cells, between whole organisms and between ecosystems, there are networks inside individual people’s heads, teams, departments, whole organisations and economical and political systems. Dell, the software company, was described by its’ founder as a “virtually integrated” company where many smaller teams of suppliers and clients are weaved together into a single organisation by information technologies (Child 2005: 197).

Social networks

The concept of applying network thinking in Sociology is often associated to Georg Simmel with his work about triads - that is relationship between three people, where one of them could act as a “broker”, a middle-man between two people, an arbitrator mediating conflict or a spoiler who tries to sabotage relationships. Since the advancement of network methodology and computing power in the 70’s, we have seen an exponential growth of papers that include the word “social network” in their abstract. In their review paper, Stephen and Pacey report about the rising popularity and diversity related to networks within organisational research. One of their findings is that we are seeing exponential growth of papers that include the word “social network” in their abstract (See figure 7). Social networks are no longer just an academic artifact but a popular metaphor used by anyone using websites such as Facebook and LinkedIn.

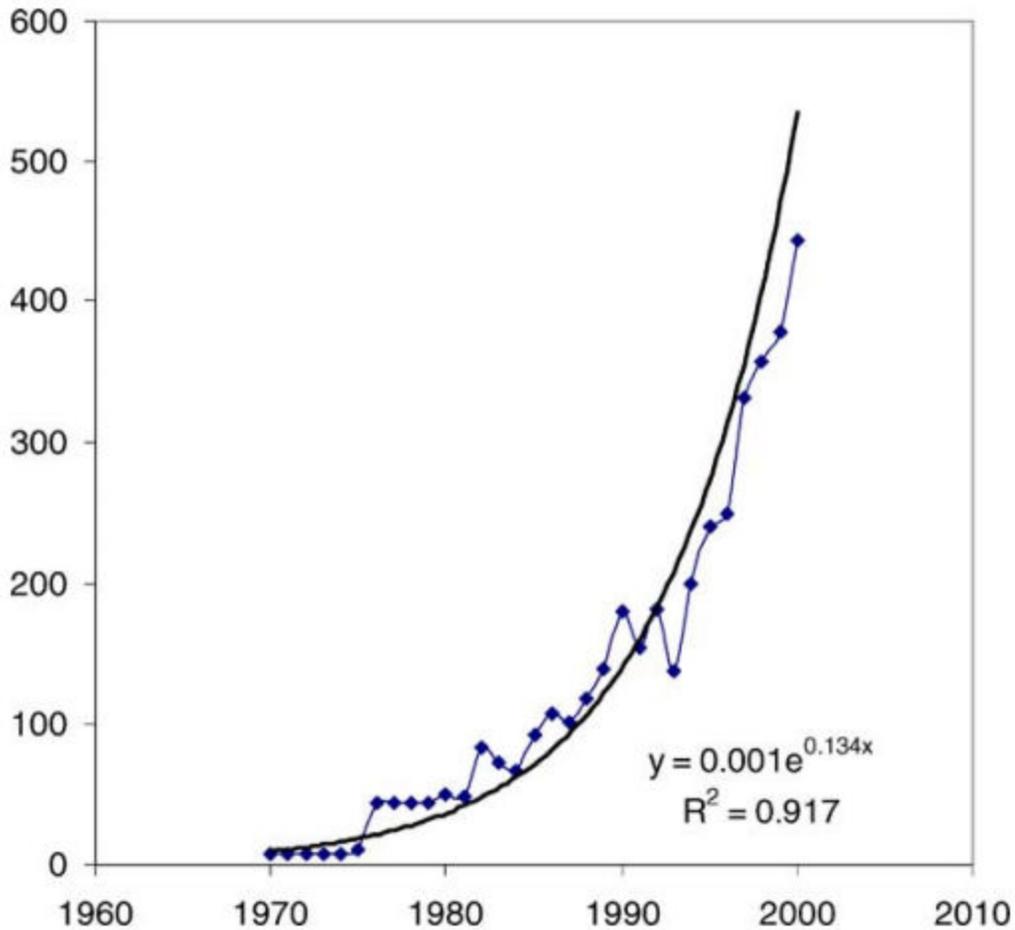


Figure 7: Exponential growth of publications indexed by Sociological Abstracts containing “social network” in the abstract of title.

Key concepts in measuring networks

Within the body of organisational research, Scott et al identify three analytical levels of looking at networks: *The ego network*, which is a specific look at the direct relationship a specific entity has, such as “who are my friends?”, or how Wikipedia articles are interconnected. *The overall network*, which included all entities and relationship within a particular domain (Figure 8). This allows us to ask questions regarding the general topology of the domain - how dense are the connections between entities? Is it centered around a few central entities or more decentralized in nature? Are there isolated niches? The third analytical level is looking at the *network position* of a particular actor within the general topology, and so we can identify influencers, niche players, dependencies, exposure to risk of a particular agent and so on.

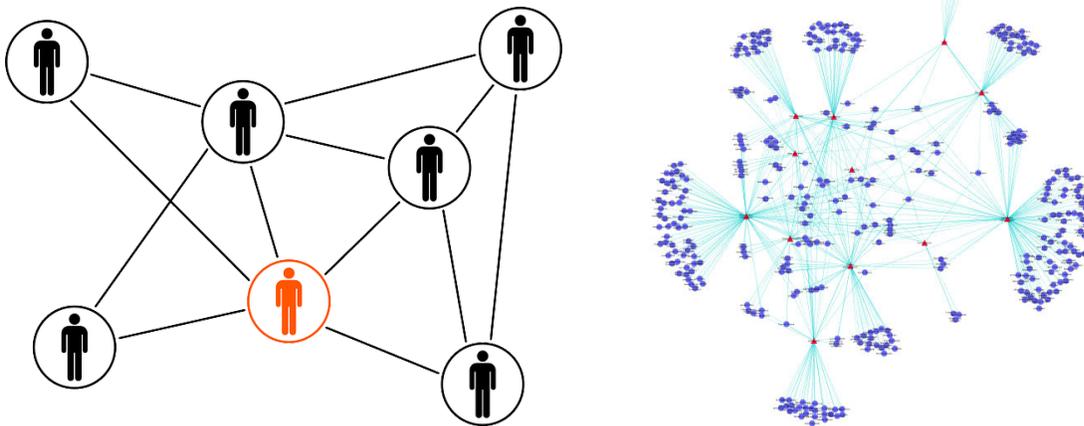


Figure 8: Ego network and overall network

Relationships are often multiplex, meaning that two agents can have more than one relationship type, and within the macro organisational picture, relationships such as project collaboration, political alliances, love stories and so forth, can coexist and aggregate into dynamics that could be unexpected to the naive analyst.

The basic measures used for network analysis are:

Distance - How many connections connect between one entity and the other. E.g. In the context of measuring distance between you and your social network, your immediate friends have a distance of one, while a friend of a friend has the distance of two. Algorithms such as “shortest path” calculate the minimal distance between two points on a network. Average distance between any two entities in a network is an important property that can help predict the spreading pace of disease, information and fashions (Watts, 1999).

Centrality - The most common measure in network position, and can be an indicator of importance of an entity within the network. The basic way to measure centrality is by *degree* - the number of immediate connections an entity has. Incoming connections (such as links to a webpage) are called *In-degrees* and outgoing connection are called *out-degrees*. *Betweenness* measures how often an entity is in the shortest path between other entities. *Eigenvector*, is a concept in linear algebra that is used, among other things, as a measure for understanding how well your connections are connected. For example, Eigenvector can answer the question “Do I have few connected friends?”. The PageRank algorithm, which had significant effect on how Google ranks its web results is a type of Eigenvector as it gives a score to web page, according to the connectedness of the web pages that link to it.

Clustering - When examining a specific group within a network, also called a cluster or sub-graph, we can look at the connectivity within it. Looking if the entities in the sub-graph are tightly-knit

together, or if there are specific entity or entities which connects them together, gives us an insight about the sub-graph. For example we can analyse the relative influence of a person within an organisation who might be a broker, a connector or an innovator within the organisation. This function of connecting together sub-graphs was named “Network closure” by Ron Burt (2001). Burt further argues that people who patch between these structural holes are more likely to receive pay raises, faster promotions and are better positioned to innovate.

Equivalence - Entities that share a similar set of relationships to others. This might indicate such things as competition or redundancy.

Density - Given a set of nodes, the number of possible connections between the nodes is the connectivity potential of this set. Density is the percent of actual observable connections from the potential.

Knowledge Management

Since the dawn of the human race, knowledge management (KM) has been evolving. The need to transfer skills, share memories and retell stories has been a pillar of the evolution of humans as social species. It has began to formulate as an academic field when companies and organisations began to spend significant time formulating theories and strategies to manage knowledge it in a conscious way from the early 80's and peaking at the second half of the 90's, when the foundation of industrial economies shifted from natural resources into intellectual resources, and companies realized that proper usage of their intellectual capital is a competitive advantage. At the same time, networked computers lowered the energy cost of storing, indexing and sharing knowledge, opening completely new venues for innovation in the field (Davenport and Prusak, 2000).

Inside the academic discourse, knowledge management is a broad term that is being used in different contexts for various functions (Scardamalia et al. 2010), where the underlying principle is that knowledge is a community property that can be directly interacted with, rather than siloed mental content inside individuals heads (Scardamalia et al. 1994). A prime innovation in knowledge management theory, is that it considers knowledge as a concept that can be dealt with in practice. Something that can be created, captured, organised, searched, retrieved, communicated and visualized (Ginsburg and Kambil, 1999; Coleman, 1999).

In the context of this research we refer to it in two ways:

First, as a computer-centered strategy that encourages the creation of new knowledge through making knowledge explicit and codified. The codification improves accessibility to the experience, knowledge and expertise of others in order to garner performance, innovation and gain competitive advantage(Beckman, 1999).

A second way to refer to it, is as a people strategy. Connecting people, encouraging knowledge transfer, direct communication as opposed to storage. Improving a sense of organisational culture and practices around sharing, appreciating and accessing knowledge.

It their review of consulting firms, healthcare and computer companies, the authors of a Harvard review paper claim that one should put his primary focus on one of these two strategies according to the the following questions: Do you offer standardized or customized products? Do you have a mature or innovative product? Do your people rely on explicit or tacit knowledge to solve problems? Standard, mature and explicit knowledge requires a computer-centered approach. Customised, innovate and tacit knowledge requires a people-centered approach (Hansen at al, 1999).

For in-depth review on knowledge management see Alavi, Maryam, & Leidner, 2001.

Application of Knowledge Management Within Organisations

While connections have been made from the ability of an organisation to improve the knowledge base of their human capital, to a greater potential to innovate (Gloet and Terziowski, 2004; Carneiro, 2000), the two intangible assets of knowledge and human capital do not easily lend themselves to tracking and quantifying. While survey research is beginning to appear in KM journals, the bulk of it is merely descriptive rather than insightful. The few survey studies that examine relationships between KM and other factors only a few articles empirically examine the relationship between KM and organizational performance (Zack et al., 2009; Rasula et al., 2012). A call has been made to advance the field by applying more quantitative practices to strengthen the conceptual and theoretical work with more rigid methods such as balanced scorecards (Valmohammadi & Ahmadi 2015).

According to the review of Changiz Valmohammadi, there are several key factors related to the successful implementation of a KM project:

Leadership role - The commitment of high level executives must be insured for the better chance of sufficient resource allocations (STRATEGIC DIRECTION, 2007), to establish condition for self and organisation wide learning (Anantatmula, 2008), and ensure proper incentives for employees to engage in KM by including KM criteria in employee assessment practices (APQC, 1995).

Organizational culture - Nurturing a positive tendency to knowledge sharing. Valuing experience, expertise and innovation instead of organizational hierarchy. Recruitment of knowledge-based individuals (Gottsachalck, 2005). Trust between employees, acceptance of change, adaptability and a collaborative culture (Valmohammadi, 2010).

Knowledge management strategy - Knowledge management requires a consistent strategy that includes an explicit description of the strategy, comparison with competitive strategies, and an action plan (Zack 1999). A KM Strategy should also create an understanding of the organization's knowledge resources and where they reside, articulate the role of knowledge in value creation, and comprise a number of integrated projects or activities phased over time including quick wins as well as long term benefits (du Plessis, 2007).

Knowledge as a processes - Knowledge must abide to a cycle treating it throughout its lifetime: Creation, storage, retrieval, transferral and exploitation (Dalkir, 2005). Some of the activities that can be used to apply knowledge management are: Transferral between employees of best practices including documentation and lesson learned, systematic identification of knowledge gaps, and using well-defined processes to bring them together, as well as resource allocation to acquire external

knowledge and transfer into the organization, participation of employees in project groups with other organizations' employees, knowledge obtained from competitors, customers, suppliers and research centers.

Training and education - KM involves usage of IT and learning new practices and so intentional training is necessary.

IT and process balance - The key to achieving harmony between KM and IT is to understand the very basic principles: there are things that computer and technology do well, and there are things that humans do well. Many of the failures of IT and KM, and much of the tension between the two, are the result of repeated attempts to force one paradigm to operate within the realm of the other (Mohamed et al., 2006), Organizations should recognize that IT is only a tool and no an ultimate solution (Wong and Aspinwall, 2005).

Reward and motivation - Literature describes two views to accomplish that: First is combining social rewards, financial rewards, further security, and further opportunity/risk as motivator tools. Second is combining performance appraisals with providing the organization with visibility of people and their ideas (du Plessis, 2007). In my opinion, one of the gaps in KM is lack of proper reward and motivation strategies incorporated into the software and communication protocol itself.

Tooling - Much of current knowledge management relies on a multiplicity of IT tools to formalize existing knowledge in organisation. Such tools include wikis, organigrams, newsletters, HR directories, project management tools, and document repositories such as Google drive or Dropbox with varying access within the organization. These technologies often offer great convenience to individuals and can help organisations aggregate, analyse and share information more effectively than even before. On the other hand, they introduce their own set of constraints such as difficulty to acquire critical mass of adoption and time cost of contributing and accessing information (Garcia-Perez and Ayres 2010)(King, Marks, & Scott, 2002).

Knowledge Identification

Trying to “Boil the ocean” of knowledge is a very common mistake for practitioners of knowledge management. Trying to capture all knowledge is quite simply impossible in real-life scenarios. Knowledge Identification is the process of (i) figuring out what knowledge within the organisation is worth putting efforts in capturing. (ii) figuring out where that knowledge is located, which could, for example, be in a person's head, a database or a document.

Probst et al. (2000) created a well-known framework that position Knowledge Identification in the context of the entire knowledge management process(see figure 9). Within this framework it is the

first process to be executed in a cycle of six activities. It is followed by knowledge acquisition, development, distribution, utilisation, and retention. Being the first process it affects the other steps following it.

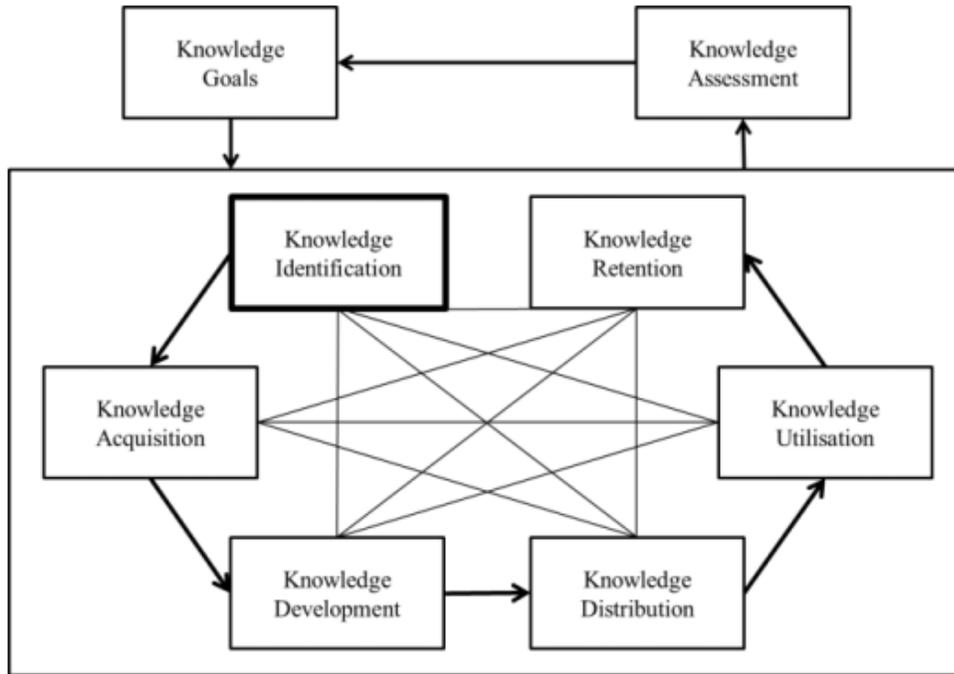


Figure 9: A knowledge management framework and the place of knowledge identification within it. (Probst et al. 2000)

Challenges faced by Knowledge identification

Following interviews with Knowledge management practitioners Tow et al(2015) identified the following factors that limit the Knowledge identification process as reported by the practitioners:

Lack of clear objectives - The organisation is unclear about its' priorities and objectives regarding the knowledge it needs to identify to enhance it's performance.

Knowledge isn't recorded - Staff is not writing down their knowledge in shared repositories, and/or the repositories are not up to date.

KI Methods Effectiveness - The methods are not appropriate for the specific use-case.

KI operational implementation - This complex work falls between the seats of HR and IT, without a clear individual and team responsible to implement it.

Other issues with KI methods - Knowledge stored in knowledge repositories is outdated, Searcher requires subject matter and knowledge to be able to craft queries and interpret results. Poor, inconsistent, not-repeatable search results. Patchiness of information leads to a lack of critical mass. Too much reliability on self-identification or voluntary submission. Employees do not know what to share and finally, lack of incentives for employees to share their knowledge. Not one single consistent KI method. Not knowing which KI method to use. Difficulties in adapting to changes as people learn and move. Overly complex process in how knowledge is captured and collected. Lack of formal processes, no accountability. No trust in currency of knowledge. Lack of KM policies.

As evident, Knowledge Management as a practice is extremely hard, and required a lot of resources, customisation and efforts to succeed. Similarly to Knowledge mapping, described in chapter [1.3.2 Expert identification software and methods](#), it seems the underlying reason for this is an attempt to regulate people's work activities in ways that are not intuitive or enticing to the people involved. Providing value to people in different levels within the organisation is often left unanswered in such projects, and is one of the major challenges I was facing myself in the pilot projects we initiated.

Expertise mapping

One of the key ways of supporting the people strategy of KM is using expert mapping, thus making it easier for members of the organization to identify other members that have certain skills which may be of assistance to them.

While it is relatively simple for individuals to identify experts within their immediate working environment or team, the process of identifying people from more distant circles of interaction is significantly more complex. Previous work to identify experts within organisations were based on methods such as keyword text mining (Dijk, 2009), co-word and proximity analysis (Criscuolo, Paola, Ammon, & Tony, 2007), social network analysis, citation networks and Linked Open data representation (Börner, Conlon, & Corson-Rikert, 2012) and extracting data from LinkedIn, emails and project management software (<https://corp.whoknows.com/>). These tools often serve an external researcher (auditor), or they abstract away so much of the actual representation, that the people who are being documented lose awareness of what is happening. In our previous work on mapping knowledge, [Knownodes](#) (Garbash et al, 2015), we noticed the acuteness of the problem. In the context of a university course, students who were a part of a team mapping research subjects to publications, had in fact, no good awareness that they were building a network together. This lack of awareness hindered both exploration and active participation in the construction of knowledge.

The field of expertise mapping, is one of the core objectives of the Rhizi software. An extended review of the field is available on chapter [1.3.2 Expert identification software and methods](#).

Review of existing solutions

What can the academic literature tell us about ways to overcome the challenge of organisational awareness. In this review I focused on three approaches:

In segment [1.3.1 Review of academic literature on knowledge sharing](#), I provide a review of what people think works for motivating people to share knowledge and collaborate better on their expertise. Considering the complexity of human interaction involved, it can serve as a good reference when applying systems of our own and see if it addresses issues raised in the review.

In segment [1.3.2 Expert identification software and methods](#), I zoom-in to discuss expert identification software. The rationale for this is that it became clear through the pilots in Knownodes that the potential value for mapping experts, being dynamic sources of insight that understand the state of the art and context of the help seeker is much more fitting than a static knowledge base that requires frequent updates. This was continued in the pilot done using Rhizi for expert identification in Descartes university, [Descartes IT](#), and the final proposal in segment [4.1 Proposal for a future project](#).

Review of Academic Literature on Knowledge Sharing

Organisational research teaches us that thinking about knowledge sharing as solely a technological problem is doomed to fail. One must adopt a holistic view that takes into account the cultural, social and structural perspectives. In the chapter, I review several solutions and approaches in the academic literature regarding these aspect of knowledge sharing. Following the literature review by Wang (2010), I separate the review as it applied in different scales of an organisation: whole-organisational-level, team-level and individual-level.

Organisation-Level

Management support

Several research projects have documented the effectiveness of support from management in cultivating within employees a positive perception of knowledge sharing culture, willingness of experts to help each other, trust and be willing to share knowledge (Connelly & Kelloway, 2003; Lin, 2007d, Lee et al. 2006). Examples such as managers spending time to share their own knowledge, encouraging employees, investing and adopting knowledge sharing technologies not just for the tool in itself but also as a symbol are examples of such behavior. This kind of support isn't merely exercising power: "when lower level workers are ordered to 'share' information with those higher up the corporate ladder, a cutthroat information culture of meddling micromanagement can result" (Davenport, 1994).

Supervisor and coworker support increase rates of knowledge exchange and its perceived benefit (Cabrera et al., 2006; Kulkarni, Ravindran & Freeze, 2006).

But how can management proactively support collaboration and knowledge sharing within their organisation? In his influential Book "Collaboration", Morten T. Hansen specifies a methodology for organisation leaders to assess collaboration problems and implement solutions to resolve them. To this end, he introduces the concept of "disciplined collaboration" which frames collaboration as a tool for achieving tangible goals and constantly assessing it accordingly. The three steps he described towards reaching disciplined collaboration are:

Evaluate opportunities for company wide collaboration across organizational units

These opportunities can be centered around innovation, cutting costs and customer acquisition for corporations. Better projects, decisions and cutting costs for non-profits and governments, and solving problems people care about for legislators.

According to Hansen, it is critical to be able to assess if a collaborative project is worth pursuing. In order to do that, he suggests the following formula:

$$\text{Collaboration Premium} = \text{Return on Project} - \text{Opportunity Costs} - \text{Collaboration Costs}$$

Collaboration Premium – Specifies whether you should start a collaborative project or not.

Return on project - The net value gained from the collaboration.

Opportunity cost - The answer to the question: “What else could we be profitably doing with the time, energy and resources we are committing to this collaboration?”

Collaboration Costs - What are the costs of making the collaboration? These include the hassle of working across units to make the collaboration happen, the time and energy spent in defining and negotiating the collaboration, and risks such as delays and budget overruns.

Look at the four potential barriers which might derail your collaborative efforts and assess which are the one’s you concentrate on fixing

(i) The “not-invented-here” barrier – we don’t reach out to others (ii) The “hoarding” barrier – we keep things to ourselves (iii) The “search” barrier – we can’t find what we need anywhere (iv) The “transfer” barrier – we only work with people we know well.

Tailor solutions to these barriers using a mix of three collaborative levers a leader can use

(i) Unification lever – get everyone aiming at a lofty goal: Use leadership status and leverage to signal that collaboration is highly valued and desirable. This is achieved by setting shared, understandable goals that people can get behind on. (ii) T-shape lever – encourage subordinates to work towards goals within and across units. Educating organisation members to deliver results within their units individual performance (vertical) as well as derive results from collaborating across the organisation (horizontal) simultaneously, see figure 10 (iii) Networks lever – get people to use their personal networks. Encouraging formation and strengthening of cross-unit and informal collaboration networks. Networks help achieve two principal goals: (1) Help people identify collaboration opportunities and find the resources they need. (2) Enable the collaborators to capture part or all the added value created.

Results achieved by collaborating across different units in your organization

Results generated by your own unit alone

Figure 10: the T shaped model

In general, the Unification and T-shaped levers are most commonly used to address the “not invented here” and the hoarding barrier. This is because they deal with the lack of motivation to collaborate, which is often the root cause of these issues.

The network lever is most adept in helping with the search and knowledge transfer barriers. It helps to optimise knowledge propagation across the organisation, riding on a pre-existing motivation to collaborate.

Each lever has several solutions associated with it. Below are the solutions related to each lever:

Unification:

- Craft a central unifying goal that sets a common cause greater than individual goals. The goal should be easy to describe and compelling. A well crafted goal follows the following criteria:
 - Articulates an end-goal that all members of the organisation agree to be a potentially tremendous achievement.
 - Simple and concrete. Leaving little room for personal interpretation.
 - Evokes passion and motivating. For this, referring to competition can be very constructive.
 - Puts collaboration on the inside and competition on the outside. Makes people unite against a bigger foe.
- Create and demonstrate the core value of teamwork. This requires specifying collaboration as cross-unit, not just intra-unit. Senior management must provide personal example and not be seen as just cloaking their intentions to better position their own units. Collaboration is presented as a tool to achieve concrete goals, not as end-result in itself.
- Create and use language which encourages collaboration. Ensure you spend a good proportion of time talking about collaboration and emphasizing its importance.

T-Shaped management:

- Managers should dedicate 15-20% of their time for cross unit collaboration.
- Reward system: See more details below in the appropriate section. Hansen claims that a formula of providing 50% bonus for individual performance and 50% of total bonus for collaborative efforts has worked well for many organisations.
- Promote people who adopt T-shaped behaviour (e.g. Dedicate time and energy towards collaboration efforts)
- Specify the criteria for cross-unit collaboration. This means setting explicit metrics for measuring collaboration.
- Consistently gather data about collaboration and evaluate it transparently and candidly with the rest of the organisation.
- Hire proven T-shaped managers (see Figure 10: T-shaped manager).
- Coach for T-shaped behavior: Training, mentoring and usage of vocabulary for promoting collaborative behavior patterns.
- Dismiss or encourage non-collaborators to change their behavior.

Networks:

- Build outwards, not inwards - Try to enrich your networks with people outside your organisation, or outside your own unit.
- Emphasize network diversity rather than network size - Adding people who have diverse perspectives, fields of expertise and know-how is more effective than a large network of people who share your core competencies.
- Weak ties are more effective - Being connected to people you are in touch with infrequently serves as a more effective bridge to rarely accessed resources. Weak ties are also cheaper to maintain.
- Connect with “super-node” people - Some people cultivate and tend to be more helpful connecting between people within their network.
- When contacting someone, use swarming techniques - Try and present more than one connection to the person you are contacting. Mentioning all the influencers you know. Invoke common threads.
- Know when to switch into deep work - As a project gets more complicated, find the few people you can “deep-dive” with, rather than maintaining a lot of shallow connections.
- Hansen recommends a methodology for evaluating the business network your organisation is developing according to the principles listed above:
 - Map the network as it currently exists
 - Evaluate your network - Examine cross-unit ties, compare to the principles listed above and determine what are the deficiencies.

- Tailor interventions - To address the deficiencies, plan ways to create ties between people.



Figure 11: The paint-drip model.

Kent Beck (2016) uses the metaphor of the paint-drip people to contrast against the prevalent T-shaped model. Imagine a drawing a brush horizontally across a canvas, and drips are forming and start to glide down at different heights. Each created drip column represents work done in a different project or initiative, some of them are merely self-teaching, or explorative projects, some go deeper. The basic idea is that people possess a diverse skill set that evolves over time, they are curious, always exploring and can move between focus-areas. They sometimes resurrect their interests.

Reward and Incentive Systems

In their seminal paper, Davenport and Prusak (2000) argue that the most successful knowledge transfer initiatives within companies possess formal and informal systems to reward collaboration and expertise-sharing (see chapter [1.1.3 Human factors limiting organisational self-awareness](#)).

One way to implement such a system is to add collaboration as a factor when conducting employee evaluations. In the Harvard Review article “How to Build Collaborative Advantage“ (Hansen & Nohria 2004), the authors review an annual performance system implemented by Intuit, the financial-software company. The company’s employees are evaluated on two questions: “What was

accomplished?” and “How were the goals accomplished? The “how” part evaluates the employee’s collaborative efforts across functions and business units in reaching these goals. An employee who did not collaborate enough will receive a lower mark for the year, even if he achieved his individual goals.

British Petroleum also integrated measures for collaboration in three ways: First, managers undergo a feedback and review process from peers. This peer-review process is an important consideration when a manager is evaluated for promotion. Secondly, 30% - 50% of bonuses for senior managers are contingent on the performance of the firm as a whole. Lastly, executives keep a close eye on the extent to which business-unit managers ask for peer assistance and will intervene if someone is not seeking enough help (Ghoshal and Gratton 2002).

Harvard’s Business Review article on enterprise collaboration(2011) critiques this approach: “Pay systems are not primary drivers of motivation in collaborative organizations. People will become dissatisfied over time if they feel their pay does not reflect their contributions, but their daily decision making is not guided by the goal of maximizing their compensation.” Several research projects even found negative relationship between rewards and knowledge sharing (Brock & Kim, 2002; Brock et al., 2005). Other studies found no relationship between reward systems and knowledge sharing intentions or attitudes toward knowledge sharing (Kwok & Gao, 2005; Lin, 2007c,d). Chang, Yeh & Yeh (2007) showed that outcome-based rewards and sufficient rewards for effort did not foster knowledge sharing among product development team members.

So how do reward systems do work? They must work together with other factors such as reputation, culture and defining a shared purpose. Some of my findings from interviewing IT staff reinforce this criticism, as some state that their own motivation for collaboration stems more from ethics of contribution and rejected the idea of financial reward as a motivating factor for helping others (See chapter [3.5 Descartes IT](#)).

A secondary yet critical challenge of implementing this type of solution is that it requires the organisation’s leadership to have a complete buy-in and a multi-year commitment to this kind of reward system.

Organisational Structure

Kim & Lee (2006) conducted research within 10 public sector organisations and discovered that employees’ perceptions of the level of centralization in their organisation were negatively associated with their perception of their knowledge-sharing capabilities. In their conclusion, the researchers recommend organisations to engage in a critical assessments of their structure and its impact on vertical and horizontal communications among employees.

Other researchers have suggested concrete ways to “hack” the organisational structure: Providing open spaces to encourage interaction among employees (Jones, 2005). Use fluid job descriptions and job rotation (Kubo, Saka, & Pam, 2001). Encourage communication across departments and informal meetings (Liebowitz, 2003; Liebowitz & Megbolugbe, 2003; Yang & Chen, 2007).

Culture and Climate

Research conducted by De Long and Fahey (2000) on fifty companies found that the benefits of knowledge sharing infrastructure were limited if company culture and practices did not support knowledge sharing across units. According to the literature review made by Wang and Noe (2010), a lot of research attention has been given to trust as a leading factor within organisational culture. Kankanhalli et al (2005) report from a survey of public sector organisations that a culture of trust alleviates the perceived cost of knowledge sharing. Trust is also linked with implementing intranet-based knowledge management systems, individual knowledge sharing, and the organisation’s capability of combining and exchanging knowledge (Chiu, Hsu, & Wang, 2006; Collins & Smith, 2006; Liao, 2006; Ruppel & Harrington, 2001; Willem & Scarbrough, 2006). To establish trust, avoiding emphasis on individual competition and establishing cooperative team perception can help (Schepers & Van den Berg, 2007; Wang, 2004; Willem & Scarbrough, 2006). Schepers and Van den Berg(2007) conclude that the organisational culture type of “cooperative team” stimulates social exchange between team members. “ooperative team” is characterized as possessing a set of values that stress openness, commitment, trust, and morale.

Strengthening certain subjective norms centered around sharing are related to increased knowledge sharing (Bock, Zmud, Kim, & Lee, 2005; McKinnon, Harrison, Chow, & Wu, 2003). These include: perceptions of fairness, innovative culture and a sense of affiliation (togetherness) with colleagues.

Shared praxis

Nielsen(2011) identifies that many challenges to collective intelligence stem from a lack of *shared praxis* - the body of knowledge and techniques used to collaborate. This generates a way for the the collective to identify that progress has been made. An example of good shared praxis is when working on a math problem, and a group of experts can identify when sound mathematical approach has been suggested by a member. Lack of praxis can be identified in politics, when a group with different value-systems engage in negotiation and if the cognitive overhead of identifying how proposals relate to each faction’s self-interests.

Communities of Practice

The norm of reciprocity is defined as the expectation that people will respond favorably to each other by returning benefits for benefits, and responding with either indifference or hostility to harms. Its relationship with knowledge sharing was examined in literature in context of communities

of practice within organisations. A community of practice is a work-related group of individuals who share common interests or problems, and learn from each other through on-going interactions (Lave & Wenger, 1991). A community of practice may be an internal group within an organisation or a cross-organisational interest-oriented group (e.g. people who practice a certain programming language). Research results concerning the relationship are mixed. Chiu et al. (2006) show significant relation between the norm of reciprocity and knowledge sharing while Wasko and Faraj (2005) show a negative relationship. There seems to be a hidden, more nuanced process in work, where other criteria such as individuals personality traits, norms of pro-sharing environment and perceived usefulness of the community are in play.

As an example of such approach, the multinational Chevron's strategy have been to nurture and support these networks when they appear. Chevron assigned a team dedicated to provide on-line technology networks, resource listings, and a Best Practice Resource Map. Dozens of communities of practice have sprung up over the last few years, including networks on customer satisfaction measurement, training, safety, quality, and a variety of technical issues. The resource map purpose is to identify key people in these networks and lists phone numbers and e-mail addresses so that people can find each other (O'Dell & Grayson 1998).

Communication Medium (Online, Face-to-Face etc')

Oshri et al.(2008) suggest that standardization of templates and methodologies across remote sites, frequent teleconferencing sessions and occasional short visits support knowledge transfer within software teams.

Lateral Cross-Unit Mechanisms

Hansen and Nohria (2011) recommend the development of lateral cross-unit mechanisms to deal with difficulties in knowledge transfer. Cultivating strong professional relationships between different units and development of formal cross-unit groups and committees. For this they recommend the cultivation of "connectors", making available an internal electronic yellow pages of expertise within the organisation and implementation of a best-practices benchmark system.

Originally attributed to Jon Udell, but adapted by Michael Nielsen (2011) the concept of Designed Serendipity is the process whereby the myriad problems faced by researchers can be easily unlocked by finding the right expert at the right time to help. Nielsen uses the example of Einstein benefitting from discovering an expert who helped him in the right time:

“He[Einstein] talked his problem over with a long-time mathematician friend, Marcel Grossmann, telling him, “Grossman, you must help me or else I’ll go crazy!” Fortunately, for Einstein, Grossmann was just the person to be talking to. He told Einstein that the geometric ideas Einstein needed had already been worked on in full, decades earlier, by the mathematician Bernhard Riemann.” These kind of systems will be discussed in chapter [1.3.2 Expert identification software and methods](#).

Interpersonal and Team Level

A different perspective to look at knowledge sharing is to examine characteristics and processes within teams and try to isolate factors that are strongly tied with knowledge sharing.

Team Characteristics/Processes

Several factors such as time passed since the formation of a team and team cohesiveness were identified as leading to higher likelihood of knowledge sharing (Bakker et al., 2006; Sawng, Kim, & Han, 2006).

<i>Name of the factors</i>	<i>Items</i>	<i>Factor 1</i>	<i>Factor 2</i>	<i>Factor 3</i>
Group cohesiveness	Personal trust	0.858	0.028	0.025
	Sense of group belonging	0.842	0.055	-0.090
	Personal consideration	0.826	0.021	0.107
	Level of cooperation between members	0.826	0.035	0.119
	Level of intimacy between members	0.783	0.015	0.050
	Group preferences	0.770	0.099	0.003
Task characteristics	Importance	0.139	0.876	0.053
	Identity	0.101	0.841	-0.106
	Functional diversity	-0.058	0.550	0.122
Interdependence	Mutual technical cooperation	-0.071	-0.089	0.763
	Connection between groups	0.141	0.010	0.661
	Interdependency between groups	0.043	0.306	0.512
Eigen value		4.075	1.894	1.347
% of variance explained		34.0	15.8	11.2
Cumulative % of variance explained		34.0	49.7	61.0

Figure 12: The results of factor analysis on group characteristics variables (Sawng, Kim, & Han, 2006).

Another study related to communication style by De Vries, van den Hooff, and de Ridder (2006) indicates that agreeableness as a team communication style (friendliness, empathy, and affiliativeness) and extraverted communication style (talkativeness and enthusiasm) is positively associated to knowledge sharing willingness and behavior.

Lastly, an empowering leadership style within teams, e.g. that which shares power and/or provides autonomy to lower ranked members, had been strongly associated to knowledge sharing (Srivastava, Bartol, and Locke 2006).

Expertise Diversity

Acknowledgment of team members diverse expertise has been shown to increase knowledge sharing (Thomas-Hunt et al. 2003). Moreover, explicit groups in which member expertise is made public share more unique information than groups in which member expertise is unknown (Kim 1997, Stasser et al. 2000, Stasser et al.1995, Stewart and Stasser 1995).

Additional research by Larson et al. (1998) and Wittenbaum (1998, 2000) showed that designated leaders and members with task experience are more likely than others to repeat unique information. Consequently, public acknowledgment of group member expertise should increase members' overall participation as well as their emphasis on their own unique knowledge.

Cummings (2001), in a field study of 182 work groups in a fortune 500 telecommunication firm, had shown that external knowledge sharing (customers, experts and people outside the workgroup) is increased when structural diversity as expressed in geo-location, functional assignments, business units and reporting managers of the team-members is more prominent. Cummings hypothesizes that this is due to a richer ecosystem of external sources.

Social Networks

Research shows that ties among individuals in social network has a positive effect of knowledge transfer and quality of the information transferred (Cross & Cummings, 2004; Hansen, 1999; Reagans & McEvily, 2003). Within the premise of virtual communities, number of direct ties and personal relationship in the ego-network of an individual has a positive relationship to quantity and perceived helpfulness of shared information (Chiu et al., 2006; Wasko & Faraj, 2005). Knowledge transfer between weak-ties is reported by the knowledge seekers to be more useful (Levin and Cross 2004).

Motivational Factors

Gagne (2009) proposed a model of knowledge sharing that is based on theory of planned behavior (Ajzen, 1991), and self-determination theory (Deci & Ryan, 1985, 2000). According to this model, she suggests several actions to improve motivation for knowledge sharing:

- When hiring people: Take into account fit between the candidate values to organisational values.
- Creation of motivating job design: Usage of variety of tasks and skills. Seeing an entire piece of work from beginning to end. Direct contact with those affected by the work. Provide some decision making power. Receive performance feedback.
- Integrate knowledge sharing as a performance criteria.
- Recognition and intrinsic rewards is more likely to encourage knowledge sharing than monetary or other tangible rewards.
- Use group based rewards to foster cooperation.
- Promotion of knowledge sharing by managers - Understanding and acknowledging subordinates perspectives, encouraging self-initiation, minimizing pressure and controls, providing relevant information. Avoiding the use of evaluation, deadlines and tangible rewards. Providing rationale for requests. Choice. Decision making power. Provide opportunities for initiative. Training. Access to information and resources. Optimal

challenges and goals. Constructive feedback. Increased interactions. Support for cooperation. Sharing information and experiences. Acknowledging feelings.

- Leaders must act as role-models by sharing their own knowledge and help subordinates synthesize incoming information and articulate a common goal that will facilitate knowledge creation.
- Use training as opportunity to create norms about sharing behaviour - Teaching communication skills. Teaching what knowledge to share and how to share it.

Beliefs of Knowledge Ownership

Research shows that when employees believe information is owned by them and not by the organisation they are more motivated to participate in knowledge-sharing activities (Constant et al., 1994; Jarvenpaa & Staples, 2000). This is attributed to people's internal satisfaction of sharing with others. Aspect such as solidarity and need for achievement were related with ownership belief (Jarvenpaa and Staples 2001).

Perceived Benefits and Costs

According to social exchange theory, individuals evaluate the ratio of cost and benefit of potential action. Expectation is that their actions will lead to rewards such as respect, reputation and tangible incentives (Blau, 1964; Emerson, 1981). Research shows consistency of that principle in relation to knowledge sharing. Several research papers show that knowledge sharing is more related to individuals' beliefs that sharing knowledge is useful to others, than to the benefits they personally gain, and especially in a professional context (Chiu et al., 2006; Siemsen et al., 2007; Wasko & Faraj, 2000). The more time and effort employees perceived as necessary to codify knowledge in order to share it, the less likely they are to use electronic sharing repositories. This was aggravated when there was low trust in the likelihood other employees will contribute and reuse the knowledge (Kankanhalli et al. 2005).

Interpersonal Trust and Justice

Expectation for reciprocity, establishing a community of practice and collective actions as a team or community are key elements in knowledge sharing that depend on perception of interpersonal trust and justice (Wu, Hsu, & Yeh, 2007).

Abrams et. al (2003) conducted interviews of employees in 20 organisations and identified ten "trust building" behaviours that promote these perceptions within a knowledge sharing context:

Trustworthy behaviours

- Act with discretion - Be clear about what information you are expected to keep confidential. Don't reveal information you have said you would not ... and hold others accountable for this.

- Be consistent between word and deed - Be clear about what you have committed to do, so there is no misunderstanding. Set realistic expectations when committing to do something, and then deliver.
- Ensure frequent and rich communication - Make interactions meaningful and memorable. Consider having some face-to-face (or at least telephone) contact. Develop close relationships.
- Engage in collaborative communication - Avoid being overly critical or judgmental of ideas still in their infancy. Don't always demand complete solutions from people trying to solve a problem. Be willing to work with people to improve jointly on their partially formed ideas.
- Ensure that decisions are fair and transparent - Make sure that people know how and why personnel rules are applied and that the rules are applied equally. Make promotion and rewards criteria clear-cut, so people don't waste time developing a hidden agenda (or trying to decode everyone else's).

Organisational factors

- Establish and ensure shared vision and language - Set common goals early on. Look for opportunities to create common terminology and ways of thinking. Be on the lookout for misunderstandings due to differences in jargon or thought processes.
- Hold people accountable for trust - Explicitly include measures of trustworthiness in performance evaluations. Resist the urge to reward high performers who are not trustworthy. Keep publicizing key values such as trust—highlighting both rewarded good examples and punished violations—in multiple forums

Relational factors

- Create personal connections - Create a “human connection” with someone based on non-work things you have in common. Maintain a quality connection when you do occasionally run into acquaintances, including discussing non-work topics. Don't divulge personal information shared in confidence
- Give away something of value - When appropriate, take risks in sharing your expertise with people. Be willing to offer others your personal network of contacts when appropriate.

Individual factors

- Disclose your expertise and limitations - Make clear both what you do and don't know. Admit it when you don't know something rather than posture to avoid embarrassment. Defer to people who know more than you do about a topic.

Baker et al. (2006) examined trustworthiness according to three criteria: capability, integrity and benevolence. Their two principal findings were that (i) knowledge sharing lessened between team members when the team member perceived his colleague as being very capable.(ii) They shared

more knowledge when they believed other team members to be honest, fair and followed principles (integrity). Interesting to note that perception of good will (benevolence) of the knowledge recipient towards the knowledge giver was not significantly related.

The role of justice in affecting the quality of social exchange relationship has been well established (Rupp & Cropanzano, 2002). Schepers and van den Berg (2007) found a positive relation between procedural justice to perception of knowledge sharing among employees.

Individual Attitudes

Individual's' expectations of usefulness of sharing their knowledge and that by sharing they can improve relationships with others is related to positive knowledge sharing attitudes which are in turn related to knowledge sharing behaviors (Brock & Kim, 2002). As mentioned before, individual attitudes of leadership regarding knowledge sharing is an important factors in adoption of knowledge sharing activities in organisations. Research by De Vries, van den Hooff, & de Lin and Lee (2004) found that managers intention of encouragement was related to employee sharing behaviors.

1.3.2 Expert Identification Software and Methods

The goal of this chapter is to review the state of the art within expert identification methods and software. This field has seen a surge of activity in the last few years with new software solutions, techniques and papers coming out during the course of my thesis. I therefore adjusted the review to reflect the latest changes to be later compared to Knownodes and Rhizi within chapter [4 -](#)

[Conclusion: Met and unmet challenges analysis](#). To review the methods and software, I will use two criteria: (i) It improves a part of expert identification system (ii) The method or software helps to understand current trends and available solutions for expertise identification.

Background

A group's awareness to each other's expertise is also called transactive memory (Wegner 1986). It has been shown to affect group performance (Moreland and Myaskovsky 2000, Moreland 1999, Hollingshead 1998, Littlepage et al. 1997, Liang et al. 1995) as well as individual's perceptions within the group. Team members who are identified as experts are often afforded power and status (French and Raven 1959) that alters their performance within the team (Stasser et al. 2000, Stasser et al.1995, Wittenbaum 2000).

Yiman-Seid and Kobsa (2003) outline some of the reasons for seeking an expert:

- Access undocumented information
- The need for expert dialogue to identify or specify a problem
- The need to use the expert to filter large amounts of information
- The need to interpret or contextualize information, including assessing or recommending information sources
- Social need and preference for human interaction
- The need for the expert to perform a task such as becoming an employee, contractor, team or committee member. Performing activities such as public speaking, teaching, serving as an expert witness or interviewee.

Sources of knowledge within an organisation can either reside formally and explicitly within a knowledge management system or informally and intrinsically inside people's heads. The challenges of creating and maintaining an explicit knowledge management system has been reviewed on previous chapters (See chapter [1.1.5 Technical factors limiting organisational self-awareness](#), and chapter [1.3.1 Review of academic literature on knowledge sharing](#))

The need for a system to identify experts stems from the fact that the default solution of asking a peer to identify an expert does not work in the following scenarios:

- You or your peers are unable to identify an expert.
- You or your peers are able to identify an expert, but it is not the best expert within the expert pool to solve the problem .
- You might have reservations to ask help from your peers or boss, for example, you do not wish to distract someone or reveal your ignorance about a particular subject.
- The perceived cost of engaging with the expert identification process might not be worth the perceived value.

The discrepancy between what is desired from knowledge identification and what is currently available to organisations has been demonstrated by previous research. For example, a survey of Australian KM practitioners that specifically addressed knowledge identification (Newk-Fon Hey Tow et al., 2012) found a severe mismatch between the respondents' perceived importance of knowledge identification (4.36 out of max score of 5) and the extent to which their organisations have identified their internal knowledge sources (2.74 out of 5).

According to Community Roundtable survey, people's skills and experience with projects is often undocumented. Only 5–20% of organisation members publish their expertise (SOCM 2013, communityroundtable.com)

Solutions specifically directed at expert identification are Expert Finding Systems, with which organisations mine existing data to identify who knows what (Maybury 2006), Organisational Network Analysis (ONA), with which organisations identify who knows whom (Parise et al. 2005), Knowledge Mapping, which resembles ONA, but provides a graphical representation (Wexler 2001), and the ExTra approach, used by knowledge managers to identify key holders of knowledge (Weber et al. 2007).

Expert Finding Systems

Expert finding systems are defined as systems which allow end users to discover subject-matter-experts in order to hire them(if external) or acquire their knowledge. Their main benefits are: (i) Accelerating R&D activities by rapidly locating experts (ii) Enhance rapid formation of operation, proposal or interdisciplinary teams that can respond to emerging opportunities or threats (iii) Assess organisational skill sets and predict how staffing changes would affect the organisation's capabilities (Marbury 2006).

These systems provide the following functionalities:

- *Identify* experts via self-nomination and/or automated analysis of expert communications, publications, and activities
- *Classify* the type and level of expertise of individuals and communities.

- *Validate* the breadth and depth of expertise of an individual
- *Recommend* experts including the ability to rank order experts on multiple dimensions including skills, experience, certification and reputation.

Marbury has further broken down the differentiation between these systems through the criteria of data sources, data processing, search, result organisation and system properties. Following this framework, I will review the state of the art within each of these components.

Data Sources

What data is used to determine expertise?

- Curated user profiles (Reichling & Volker 2009) - these can include self reporting in a form of open text, tags, keywords or selecting from a list of predefined categories and reporting by others e.g. Reviews by peers, supervisors or recipients of services from the expert.
- Data from ticketing and version control systems - Many project management system use a ticketing system that include issues and details about who addressed them. This can be an excellent source of data as it demonstrates past behavior that is directly related to the person's expertise. Version control system include who modified particular pieces of code, which might hint the person have knowledge of that code.
- Social Bookmarking (E.g. [Evernote business expertise discovery](#) 2016)
- Email and chat messaging (Streeter & Lochbaum 1988; Foner 1997)
- Extracting names from documents, web pages and other published work (Kautz et al. 2007)
- HR and project management reports
- CV's
- Social media activity such in information stored in general and enterprise social networks Facebook and LinkedIn (Lin et al. 2009)
- Patents, scientific publications and grants

Data Processing

What processes does the software make to treat the data?

Balog et al. (2006) propose two approaches for retrieving expertise from existing information:

(i) A candidate-based approach, which builds a profile for a candidate expert based on the relevance of the information generated by them to the user's query and (ii) a document-based approach, which retrieves all documents relevant to a user's query and then finds the link between documents and the authors of the documents.

In her analysis of the different approaches to treating expert data, Paul (2016) reports that

recent text mining techniques, such as topic modeling of employee-authored documents have replaced the candidate and document based approaches. Topic modeling approaches address the drawbacks of the profile-based and document-based approaches proposed by Balog et. al. The candidate approach underperforms the document based approach due to two reasons. First, noisy documents that don't represent the candidate well, may get added to the candidate profile which, though associated with the candidate, do not represent the candidate well. Second, there is a risk that even if the supporting documents do match the candidate well, they won't match the topic of the original user query. While the document-based approach addresses the first drawback, it only relies on query terms matching documents and does not consider any semantic concepts underlying queries and documents. Momtazi and Naumann (2013) show that topic modeling based on Latent Dirichlet Allocation (LDA) outperformed several profile and document-based approaches for expert finding by "(1) modeling each document as a probability distribution, which indicates the likelihood that it expresses each topic, and (2) presenting each topic as a set of words." Having query terms and the expert names as items in the vocabulary. Therefore, the name of the experts are represented in the same topics as the query terms that ask about that field of expertise. "If a candidate name and a query term appear in the same topic(s) with high probability, it is very likely that the candidate is an expert for the given query."

Search

The kind of retrieval supported by the tool. This could be query-based or discovery based. Discovery systems can be based on expert recommendation such as Twitter's who to follow recommendation or WhoKnows recommendation based on your google searches. Other forms of discovery include expert ranking systems and expert indexes.

Expert Result and Selection

McDonald & Ackerman, who did one of the first empirical investigation about expertise seeking within organisations (1998) identified the expert selection process as a critical component within the expert identification process. Following the search process initiated by the user, how are the results organised? What kind of criteria works best for users trying to select an appropriate expert?

Morten Hertzum conducted a literary review of 72 relevant research papers (2014) concerned with these questions and looked for patterns that might provide insight to the question. Following is a summary of findings relevant to this thesis and a breakdown of the component in the expert selection process (figure 13):

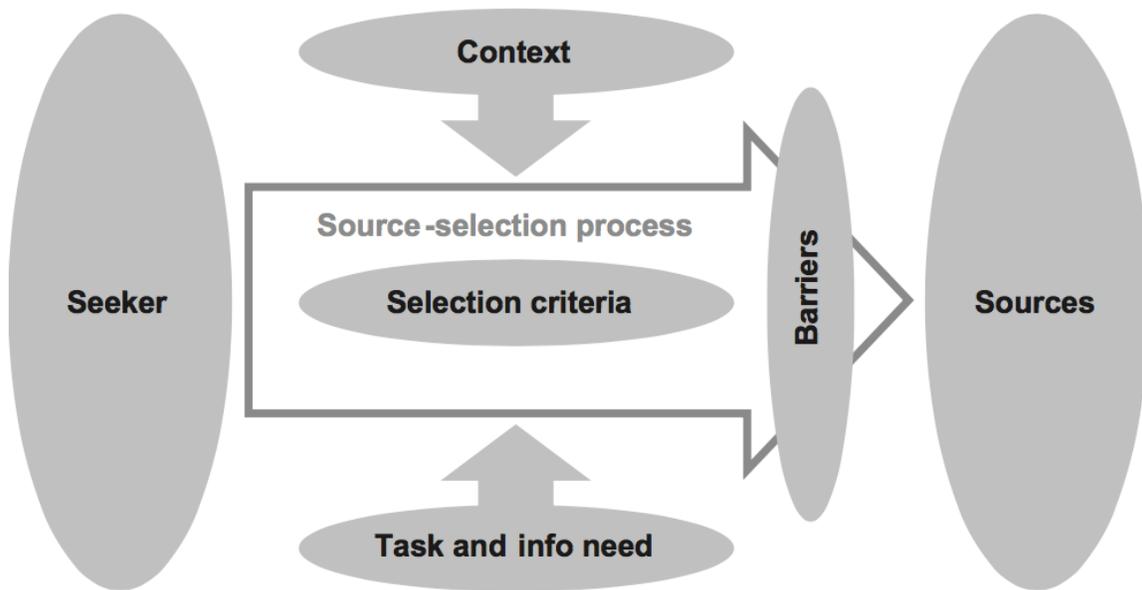


Figure 13: Components of expertise seeking: Each one of the components might affect the preferred selection criteria (Hertzum 2014).

Expert source vs document-based sources: Which is preferred and what is the balance between them?

Statistical analysis of the papers reveals a 31–81% range for usage of people as sources and 9–65% to documents. Robinson (2010) found that after 20 days of measurement, engineers were spending 7.8% of their working time seeking information from other people and spending roughly an equal amount of time seeking document sources. Curiously, other research, made by Choo (1994) and Lu and Yuan (2011), found that people sources were significantly more looked at than documents. Hertzum suggests that the difference lies in the fact that latter studies focused on important tasks while Robinson’s study involved all tasks, big and small.

Other identified patterns were that people sources were used more commonly on complex, non-routine and short-term tasks. Yuan, Rickard, Xia, and Scherer (2011) found that document sources were more appropriately used for general information and background knowledge while people sources were more appropriate for helping determine if specific information was applicable for a specific situation. While women stated their ideal information source is a people-source (51%) and men stated preference for document resources (65%), there was no significant difference when examining actual use (Julien and Michels 2000).

Internal vs external sources: What is the balance between seeking information within and outside the organisation?

This question is important when we try to understand the value of inter-unit and interorganisational collaboration. Regarding the balance, research results are inconclusive. One mentionable finding is that five studies indicate that low-frequency use of external resources is correlated with their low-accessibility (Auster & Choo, 1994; Chakrabarti et al., 1983; Cool & Xie, 2000; Culnan, 1983; O'Reilly, 1982). This suggests that a good expert identification system might help to facilitate more frequent use of external people resources.

Quality vs accessibility: Is the resource selection process influenced by expert quality and/or expert accessibility? How do you define each?

Regarding the preference, studies are inconclusive. Studies do point out several ways which improved on perceived quality and accessibility: Information about participation in an internal social-network was a strong indicator of accessibility as it implied openness to be contacted (Shami Ehrlich, Gay, and Hancock 2009). Humans functioning as gate-keepers between seekers have been identified as a source for obtaining info about accessibility and quality of experts (Civan, McDonald, Unruh, and Pratt 2009). Documentation of work artifacts, such as change history of software components, served as a useful indicator of expert relevance (McDonald and Ackerman 1998).

In Paul's (2016) research examining this question among technicians using formal requests for collaborative troubleshooting, the conclusions were that expert quality was the top criteria. Other criteria ordered by importance were: overall expertise, evidence of related-to-task experience and crowdsourced judgements of expert quality (rating). Furthermore, it was the combination of overall expertise and task-specific expertise that worked particularly well. See figure 14, for how the final interface looked following the initial criteria selection. While the accessibility feature of online presence was rated highly by the technicians, in practice, it did not have a statistically-significant effect on the experts selected. The paper concludes that user defined availability measures are more effective than automatic inferred ones. It's important to understand how much context sensitive the quality/accessibility criteria is. While within non-formal settings accessibility measures are important as seekers usually don't wish to contact experts uninterested to help, in formal settings, where collaboration is expected, it is the quality criteria which is important. This can serve as an important reminder that as long as an organisation does not give some formal recognition or permission for its employees to collaborate, the measurement of accessibility will trump quality, Scarcity changes priorities, which might hurt the organisational outcomes.

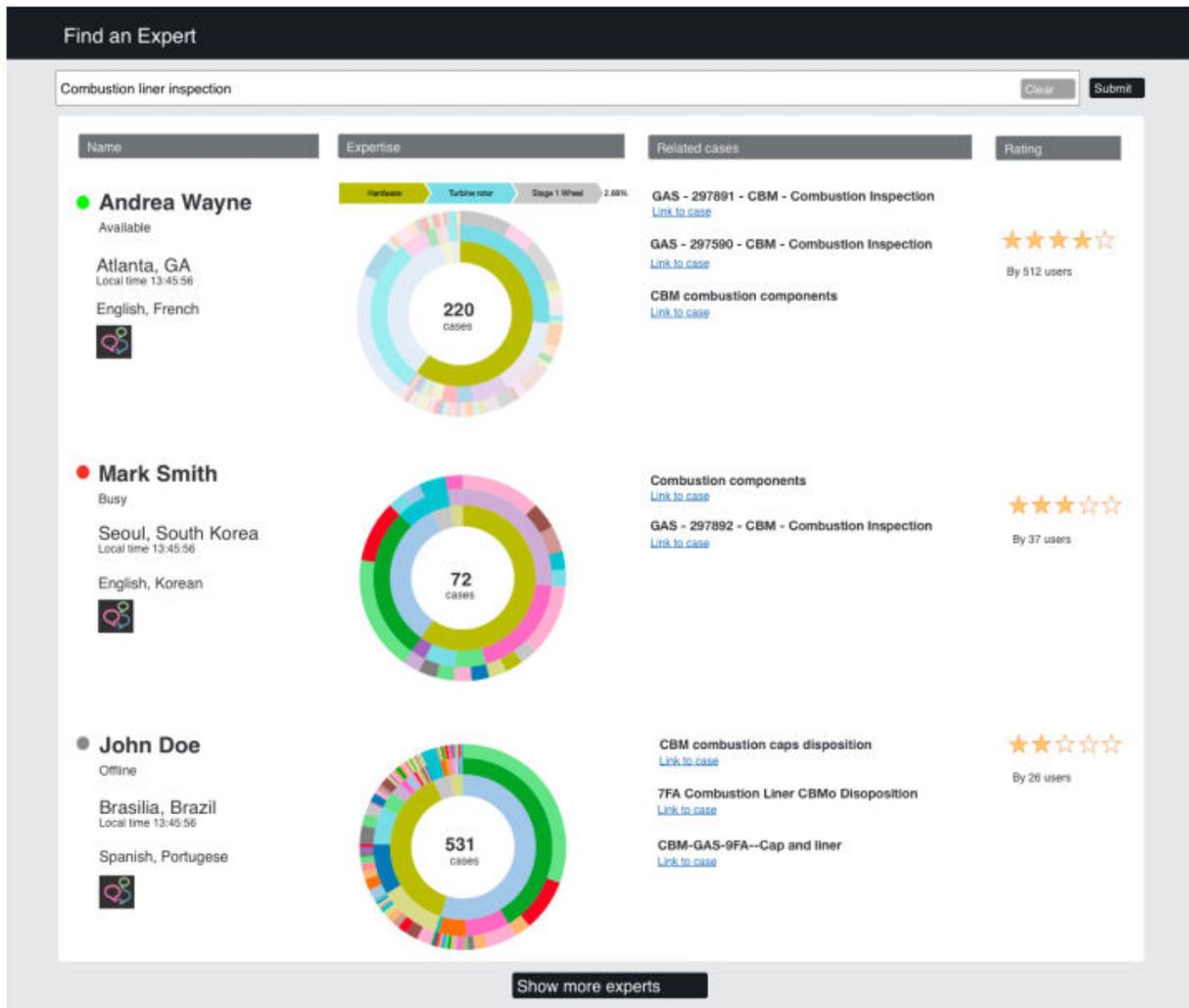


Figure 14: Expert selection interface taking into account the top criteria for selection: Overall expertise, case based expertise and crowdsourced rating (Paul 2016).

Task-related factors: What properties of the task (e.g. complexity, importance, ambiguousness) effect the expert selection criteria?

No direct influence was detected between task-related factors and other influences such as quality and accessibility. However, task importance and task urgency moderate the influence of quality and accessibility, but not consistently between studies. E.g. Within the study made by Agarwal et al. (2011) task importance increased the importance of expert quality, while Xu et al. (2006) found importance to lower the importance of expert quality. Other related findings were that task ambiguousness increased tendency to seek out experts as opposed to other sources of knowledge such as documents (Christensen and Bailey 1997).

A study in a hospital uncovered that physicians and nurses asked for help less when the help-requiring task was related to the core competency of the hospital, e.g. if the core competency of a hospital is dental surgery, physicians would ask for less help in that subject. This is hypothesized to be related to the higher social cost of asking for help (Lee 2002).

Seeker-related factors: What properties of the information seeker (experience, self-esteem, gender) effect the expert selection criteria?

Expertise seeking tends to decrease with increased job experience (O'Reilly 1982; Baldwin and Rice 1997). In addition, Vancouver and Morrison (1995) found that the likelihood of requesting feedback from expert sources was higher for participants with a high need for achievement. Tan and Zhao (2003) found that with increasing self-efficacy seekers became more willing to inquire about information. DePaulo and Fisher (1980) found that study participants who asked for help expected the helper to view them as less competent and felt more nervous and uncomfortable asking for help. These psychological aspects are key in understanding how computer-mediated system can help people feel more comfortable seeking help. This will be further discussed in chapter [4. Conclusion: Met and unmet challenges analysis](#).

Contextual factors: How do social ties and formal organisational ties factor in the selection criteria?

Social ties are important as unless absolutely necessary, seekers avoid seeking help from unpleasant colleagues (Casciaro & Lobo, 2005; Yuan et al., 2011). Positive social ties compensate for lack of prior working experience when selecting a source (Cross & Sproull, 2004). Woudstra and Hooff (2008) found that 11% of selectors in their study mention prior working experience as a selection criteria. In Hirsh and Dinkelacker (2004) study, 60% of engineers found experience with the source significant or very useful when choosing a source. The importance of experience was explained by Hertzum (2002) as a strategy to assess the credibility of an expert, and by Gerstberger and Allen (1968) as a way to reduce the perceived cost of reaching out and of collaboration with the source. Borgatti and Cross (2003) hypothesize that having a prior experience working together means that awareness to the skills of a colleague have been calibrated and updated, which affect the likelihood of them reusing that source in the future.

Nevo, Benbasat, and Wand (2012) add that if the source of information is given the option, he/she will select by its expectation of reciprocity from the receiver.

Barriers to expertise seeking: What are the barriers blocking or degrading expertise seeking?

Within the 72 papers reviewed, the barriers for selection were concerned with perceived costs (time and money) of accessing the expert (Quigley, Peck, Rutter, & Williams, 2002), that the quality of the

source knowledge is lacking (incomplete or unreliable) (Helms et al., 2011), and perceived cost to communicate the need to the source (McDonald & Ackerman, 1998).

Another relevant barrier is lack of practical experience of the expert to solve the given problem (Helms et al., 2011). This is strongly tied with the practice of assigning formal experts which can sometimes become a reverse self-fulfilling prophecy, where, once an expert status has been achieved, the expert is removed from the practical, hands-on problems within his field and serves more of an advisory or researcher role. Studies has shown that professionals such as lawyers, engineers and managers often prefer to tap into informal networks of expertise rather than the formal ones (Cross and Sproull 2004; Hertzum 2002; Wilkinson 2001).

System Properties

What kind of interoperability is supported by the tool? E.g. Integration with other knowledge repositories, project management software or external databases. Privacy, real-time communication, representation, export features, data structure are all relevant. However, these sort of features are general to software within IT environments in general so will not be expanded on here.

Examples of Expert Systems

- [GuruScan \(http://www.guruscan.nl/\)](http://www.guruscan.nl/)
- AskMe
- [Expertise Finder \(http://expertisefinder.com/\)](http://expertisefinder.com/)
- Tacit Knowledge Systems' ActiveNet
- MIT's [ExpertFinder](http://web.media.mit.edu/~lieber/Lieberary/Expert-Finder/Expert-Finder-Intro.html)
(<http://web.media.mit.edu/~lieber/Lieberary/Expert-Finder/Expert-Finder-Intro.html>)
- MITRE's XpertNet
- Thomson's tool
- Microsoft's SPUD project
- [WhoKnows \(http://corp.whoknows.com/\)](http://corp.whoknows.com/)
- [Skillhive \(http://intunex.fi/skillhive/\)](http://intunex.fi/skillhive/)
- [The Firm Directory \(https://www.neudesic.com/products/the-firm-directory/about/\)](https://www.neudesic.com/products/the-firm-directory/about/)

Challenges Faced with Expert Systems

Marbury identifies that the main challenge of expert systems is that “experts and their skills and knowledge are rare, expensive, unevenly distributed, difficult to qualify, continuously changing, varying in level, and often culturally isolated and oversubscribed. To complicate this, expert seekers typically have poorly articulated requirements, are ignorant of expert’s past performance, and are not fully enabled to judge a good expert from a bad one. Finally, their complex problems often require the combined wisdom of multiple experts.

Other challenges identified are dealing with situations where there isn’t an expert stored in the database. Handling lack of knowledge by novice users regarding the keywords related to their query. Avoiding expert fatigue due to forcing experts to deal with too many requests. Avoiding system gaming to gather unjustified reputation. How to rank experts? Can an expert be objectively assessed? How can the information remain up-to-date? These challenges will be further discussed in chapter [4 - Conclusion: Met and unmet challenges analysis](#).

Organisational Network Analysis (ONA)

Organizational network analysis is a method for studying communication and socio-technical networks within a formal organization. It is a quantitative descriptive technique for creating statistical and graphical models of the people, tasks, groups, knowledge and resources of organizational systems.

Parise et al (2005) define organisational network analysis systems main goal as to (i) identify key knowledge vulnerabilities in a network by virtue of both what a person knows and how their departure will affect a network and 2) address specific knowledge loss issues based on three roles in networks – central connectors, peripheral players, and brokers. Central connectors are defined as those who has many direct information relationships with their peers, usually because they are considered experts in one or several domains. Peripheral players are those who have fewer ties and typically placed on the boundaries of the network. They have a tendency to be disengaged and dissatisfied and more at risk of leaving their work. Despite this, there are several key values they provide to the organisation, namely niche expertise, early adoption of novel technologies and practices and knowledge resulting from external relationships (See chapter [1.2.3 The Network Paradigm Within Organisations](#)). Brokers are defined as those who have ties across subgroups in a network and so have a disproportionate ability to help an organization in order to capitalize on opportunities requiring integration of disparate expertise

Furthermore, Parise et al. identify that when an expert departs from an organisation there are several key things that leave with him that are not traditionally addressed: Subject matter expertise, organizational memory of why certain key decisions were made in the past, awareness of past

company projects (the results of which may never have been documented), and relationships with both internal employees and external business partners and customers.

ONA providers list several typical organisational scenarios that could be treated through ONA methods:

Change management - When a new process or initiative is chosen (e.g. Implementing a new safety measure, shifting to a new customer service process), ONA claims to help identify the key employees that are required to onboard in order to facilitate the change. In the webinar of the company Networkmapper (2016) the following scenario is specified:

“The customer was implementing a digital transformation with the purpose of improving customer service. This transformation meant going digital and leaving paper completely out of processes. Although the change itself was positive and made the daily jobs of employees much easier, our client organization wanted to make sure that the shift was as smooth and seamless as possible. So, they used [OrgMapper | INFLUENCE](#) to find their most influential employees and invite them to join their group of change managers leading the change. The role of the influencers was to minimize employee resistance, and to make sure employees had sufficient knowledge about what was required of them.” (Elect, empower, enhance – TOP 3 influencer involvement scenarios 2016)

Internal Communications - Map communication pattern of actors within an organisation. While organisations often have a formal communication/reporting structure in place, how does the communication actually flow? This might detect, such in the example given in figure 15, how “Cole” has a surprisingly central role within the informal communication network vs low-ranking role within the hierarchy.

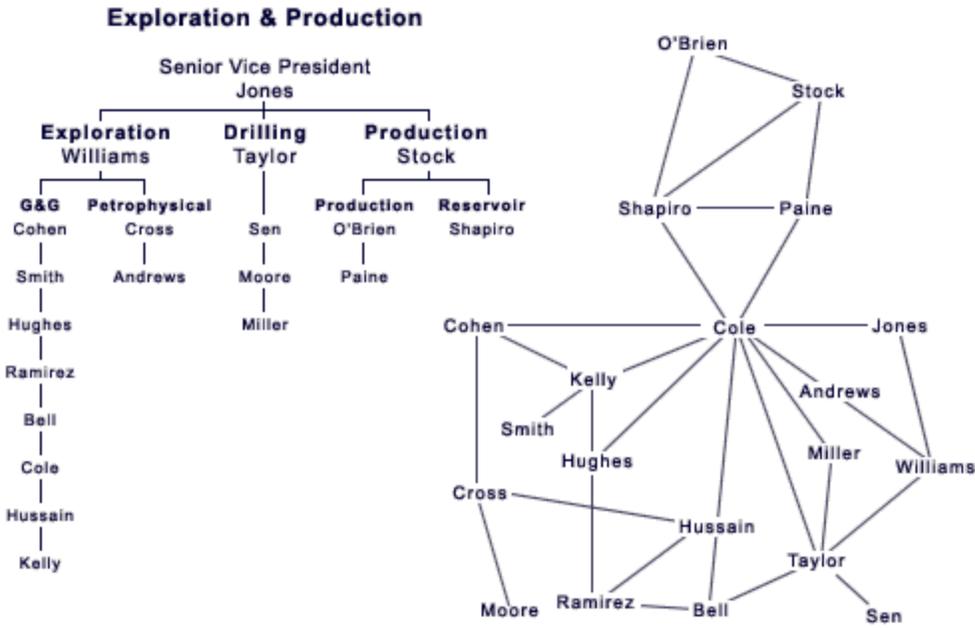


Figure 15 - formal and informal communication networks (What is ONA? 2016)

One criticism of ONA is that these one-off snapshots might change dramatically when sampled across time. Also, because the edges are not weighted, or weighted in according to subjective self-reporting it's hard to differentiate meaningful informal communication with less consequential ones.

The following four value statements have been retrieved from [HOW4 website](#) (2016).

Employee Engagement - The peer review process of understanding who collaborates with you, provides employees a unique feedback about what they mean for the organisation and to their peers, this is claimed to improve employee engagement.

Leadership diagnostics - Detect and characterize key leaders, compare leaders and team performance, assess leadership influence over team members and colleagues, pinpoint recruitment needs.

Talent development - Support talent training and development, succession planning, increase engagement through collaboration, support strategic human resource planning.

Organizational processes - Detail the information flow within projects and processes, act on interactions to strengthen the organization, understand the relational patterns within each process, identify critical links and strengthen key connections, make decisions to allocate projects or compose Teams

For more details about specific analysis methods used by these service providers, please refer to chapter [1.2.3 The Network Paradigm Within Organisations](#).

During the last two years ONA has reached industry maturity and many production-ready software solutions are available for organisations in order to convert surveys into networks and provide automatic or expert-assisted analysis of the results. Below (figure 15) is a list of state-of-the-art ONA's:

Frank	Web-based diagnostic tool that provides actionable insights based on the organizational network analysis and current team situation.	Any organization
Graph Commons	Web-based collaborative network mapping , analysis, and publishing platform.	All
HOW-4	Web-based self diagnostic tool (previously named DNA-7) that allows managers and consultants to have full organization diagnostic including insights and action items for improvement. The platform is solid so it can easily manage up to 300.000 employees. Moreover, it is fully integrated with Business Intelligence tools like tableau.com or http://qlik.com	Any organization
KeyNetiQ	All-in-one Organizational Network Analysis platform. Modern and easy to use web-based application designed to capture, visualize and analyze complex networks of formal and	Any organization

	informal relations within an organization. Highly customizable to fit into various types of ONA projects in companies of up to 50.000 employees. Includes training on the use of the tool.	
Kumu	Web-based Relationship Visualization	Social Impact, Business, Government & Policy
OrgMapper	Cloud-based tool designed to capture, analyze and visualize complex networks of formal and informal organizational relations with an automated reporting feature. OrgMapper currently has five modules to fit into various types of organizational development projects, and offers a certification program for users to be able to conduct their analyses independently.	Any organization
orgnet.com	Social Network Analysis & Organizational Network Analysis	All
Polinode	Powerful and easy-to-use web-based tool for conducting organizational network analysis. Includes integrated visualization and calculation of advanced network statistics.	Any organization
Socilyzer	A complete, web-based organizational network analysis tool for managers and management consultants.	Any organization
SYNAPP	Web-based Organizational Network Analysis, Visualization and Recommendations	Any organization
Syndio	Web-based platform that provides executives and senior managers with tools to embed social network analysis intelligence into their daily workflows.	Any organization
TEAMsCOM	Cloud-based solution. Automatic organisational analysis tool, which offers instant overview of relationship in a team or company.	Any organization which is using Google Apps for Works
TrustSphere	TrustSphere’s Relationship Analytics platform processes metadata (not the message content) from corporate email and communication systems in order to provide analytics about relationships between people, teams and organizations. It is capable of doing so at enterprise scale – including large organizations.	Any organization

VoloMetrix	People analytics solution providing global enterprises with time allocation reports and social network analysis to inform corporate strategy.	Enterprise
----------------------------	---	------------

Figure 16: List of leading ONA providers (Organizational network analysis, 2016)

Knowledge Mapping

Wexler (2001) reviews knowledge mapping as a feasible method for institutions to deal with complex information. And more specifically, to coordinate, simplify, highlight and navigate through information silos. Its purpose is to mobilize, diffuse and evaluate intellectual capital.

For explaining the method, Wexler describes the who, what and why of knowledge mapping.

Who

Wexler described knowledge mapping as a multi-actor interaction between people who possess defined roles, the different roles are described in figure 16 below.

Character	Process	Type
Map maker (Creates the details and sets the usage pattern of a knowledge map)	Whether out-sourced or in-house, the process is one of mapping the shared explicit knowledge and graphically representing it in a form which facilitates navigating problems and issues, finding new relations and developing alternative routes. Map makers can work in committees, project teams or as individuals. The map making function is often not the everyday role of the map maker at work	The type of map maker depends on the problem field being mapped, the end-users of the map and the perceived longevity of the information on the knowledge map. Examples are competency maps, strategy maps, operations maps
Map users (use maps in order to accomplish their tasks and to develop learning potential)	Whether members of the organization or stakeholders (investors, customers, contractors, etc.), map users must be motivated to use the map. Ease of access, clarity of icons, interrelatedness of parts, ability of users to locate their	The type of map user depends on the information imbedded in the map, the accessibility of the map to particular users and the value of the map in solving problems or providing shortcuts

	position and reliability of information in the map or key. Map users return to maps which work	
Map innovators alter existing maps through use, reuse and diffusion of innovation)	Whether frequent users or specialized users, some map users modify the knowledge maps as they use them. They modify the map by finding shortcuts, clarifying relationships and/or suturing the map to new or different aims than those held by the mapmakers. To become shared map innovators must find a means of diffusing their innovations for potential adoption	The type of map innovator depends on how radically they alter the initial knowledge map. Tinkerers nuance the map. They develop minor adaptations. Map innovators as map makers overhaul the map. They begin a new generation of knowledge maps. Both must diffuse these innovations.
Map champions uphold the need for knowledge maps as providing a competitive advantage for the organization)	Whether top-down, bottom-up or peer-based, map champions provide the push for map development and the creation of a new generation of knowledge maps. Map champions compete with others claiming other factors which can generate competitive advantage. The process can be seen as political	The type of map champion depends on where in the structure of the organization the push for the knowledge map originates. The top-down champion creates legitimacy. The bottom-up increases buy-in. The peer process assumes a good knowledge base

Figure 17: *The who of knowledge mapping*(Wexler 2001)

What

Knowledge maps are problem centred, meaning, that maps must address the solution of at least one concrete problem. Wexler details types of problems association with Knowledge maps:

Intellectual capital guide - Display current and potential of company intellectual assets including skills, document, policies and organisational structures, and their location within the organisation.

Implement change in the firm's technology/policy - Coordinate technological/policy transitions within the organisation, directing to mentors and experts that can help those who are having difficulties with the new technology

Acquisition plan - Locating a strategy, including contingencies(if... then..) in case one course of action isn't successful.

Accountability tree - Maps responsibilities to people within the organisation. Useful for joint or inter-unit projects.

Stakeholder orientation guide - Map how to best get information and have needs fulfilled. Identify resource portals for outsiders.

Outline of new product launch - Provide those involved in product launch with a timed sequence of interrelated activities related to the launch, who's in charge of what and the process for sharing knowledge between the different entities involved.

Why

Knowledge maps which bring returns to the organisation are financed and maintained, while those that not are abandoned or left to deteriorate. Returns can be divided into the following categories:
Economic returns - Direct energy to opportunities and threats, enhancing the organisation's competitive advantage. Help outside actors that organisation possess the capabilities to deliver on its goals. Lower transaction costs of socializing, coordinating and renewing human, intellectual and technical capital.

Structural returns - Help deal with complexity of inter-unit and inter-organisational collaborations. Communicating results of restructuring efforts such as mergers or downsizing. Capture best practices and propagate them. Locate blocks and breakdowns in organisational routines.

Organisational cultural returns - Disseminate shared views, values and coordinate values to enhance predictability and trust among participants. Help onboard new members. Increases awareness to others within the organisation. Identify mentors.

Knowledge returns - Help locate effective route, plan, scenario or procedure in order to accelerate learning. Keeping older knowledge maps, help retain organisational memory. Assist to discover and communicate new relations and ideas in a shared vocabulary.

Difficulties in application knowledge maps

Throughout his review Wexler stresses several important difficulties encountered by knowledge map initiatives:

- Lack of champions of these methods to compete against other heuristics of analysis, leading to the project becoming “Stillborn”
- Lack well identified problem to solve.
- Lack of motivation from map users stemming from not being able to solve a problem in a way that competes with alternative solutions
- Lack of feedback mechanism to improve the map
- Lack of clear legend communicating the meaning of the representations in the map
- Not encouraging self-correction and learning therefore not generating actionable insights.
- Lack of frequent updates to the maps, making them obsolete.
- Overpromising or oversimplifying expected value from the knowledge map.

While important to review, knowledge mapping has been superseded by ONA, but should be recognized as pioneers in application the methods within the field.

the ExTra approach

While the ExTra approach is very “analogue” and different from the other approaches reviewed before it provides several key insights among them, how to deal with cases where there is a critical amount of knowledge which must be preserved within the organisation. Demonstrating the type of processes that are being used as alternative to more software based solution and ,lastly, show some interesting human aspects related to dealing with unmotivated participants and capturing knowledge in a rigorous way.

Weber et al(2007) describe a the Expertise transfer (ExTra) approach. The report covers their two year experience of applying the approach for more than a 100 cases. The goal of the system was to reduce loss in expert knowledge and client relations, associated with departure of an engineer. This approach has been applied not only between people, but also between groups.

The approach has been applied to the case of retirements, job changes, or management position change.

The ExTra process involves four steps(see figure 17):

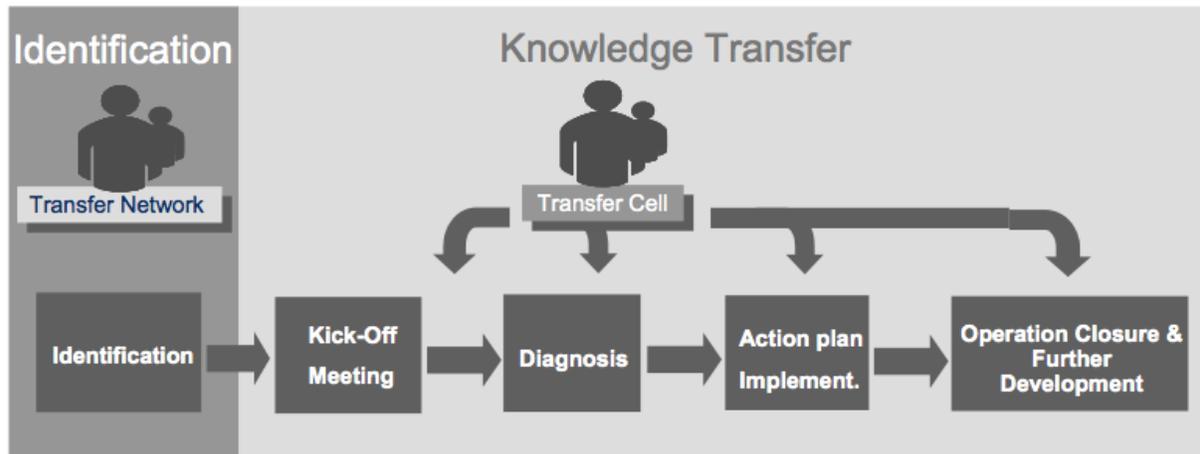


Figure 18: The main steps of the ExTra process(Weber et al., 2007)

Identification - Representative of management, HR and Knowledge management meet twice a year and compile a list of candidates.

Kick-off meeting - A transfer cell composed of the expert and the replacement, direct superior and facilitator/coordinator. These individuals meet to a kick-off meeting to establish trust and achieving joint commitment. During this meeting the process facilitator/coordinator prepares the action plan, mediates between parties, and facilitates specific transfer actions.

Diagnosis - Following the kick-off meeting the facilitator/coordinator performs semi-structured expert interviews with the knowledge provider, the knowledge receiver and relevant colleagues. Requirement and objectives are detailed, priorities are defined, and nature of the knowledge transferred is categorized. An action plan is prepared and defines next steps to be taken by the transfer cell.

Action plan implementation - A set of transfer methods is implemented, these could be: Facilitated talks, Transfer of personal contacts and network, Facilitated workshops, Documentation of lessons learned(The RISE method), Organise a “document, archives and devices clear out”, Contribute to a book of knowledge, Manage a forum/faq, create/animate training module/lecture etc’.

Operation closure - Feedback is given through a dedicated questionnaire, and if necessary, further actions are evaluated.

Challenges summary

Below is a summary of the principal challenges relevant towards the methods and results parts of this thesis. The challenges presented here are addressed in chapter [4 - Met and unmet challenges analysis](#).

Why human capital awareness?

In this chapter I discuss both work satisfaction and efficiency. It is also important to mention fairness. We are so used to think of knowledge about human capital being a scarce resource that we often do not see how this scarcity is unfair to those who are new, shy, peripheral or are introverted or have self-confidence. Moving into a state where knowledge of human capital is abundant can help people dare to take on more innovative, interdisciplinary projects because they know there are others who will have their back. This process of course needs additional resources beyond mere awareness, but this is a first step in the right direction. See chapters [3.2 CRI - Opportunity mapping](#), [3.3 Descartes IT](#) and [3.4 IGEM](#) to review several pilot projects made to explore this question further.

What stands in the way of organisational awareness?

Self programmable workers can turn into self programmable teams if we give them the means to do so. There are understandably technical challenges in recommending, storing and modelling data about human capital, but the major challenge is that of shifting (i) the culture within organisations and (ii) the perception of organisation members. After all, most members want to do what they are good at, interested in, have a sense of autonomy and a sense of directly helping other people(<https://80000hours.org/articles/job-satisfaction-research/>).

Human factors limiting organisational self-awareness

Main human related challenges are related to an imbalance between the perceived cost and benefit.

These include the following challenges:

- How can we help experts share their expertise despite their cognitive biases? (related to asking students and practitioners in the field to engage in context rich interactions on a when need to basis)
- How can we provide motivation for organisation members to take part in knowledge sharing activities? What incentives can we provide? What barriers can we remove?
- How to overcome the “Knowledge is power” perception?
- How can we prevent the “norm of reciprocity” from inhibiting people to seek help? How can computer-mediated systems help people feel more comfortable seeking help?

- How can we avoid the formal form filling and following rigid sets of rules that make knowledge sharing a lot less satisfying than informal ways?

Structural factors limiting organisational self-awareness

- How can we decouple collaboration activities that can be done in a networked way from hierarchical organisational structures?
- How to address Reluctance and sanctions against bottom-up information and norm sharing?
- How can we overcome the downside of Preplanned, static division of labor?
- How can we provide collaboration systems with measurements outside scope of defined role or task?
- What alternatives can be provided for Cultural resistance: Resistance of hierarchically minded, bureaucratic, and centralized minded managers to give up control?
- How to reduce the high barrier to entry for collaboration? These include Tasks that are too broad, conversational critical mass and production of a knowledge commons.
- How to bridge the time transition cost of moving from a self-serving to collaborative behavioural model?
- How to rank experts? Can an expert be objectively assessed? How can the information remain up-to-date?

Technical factors limiting organisational self-awareness

- How do we reduce knowledge retrieval within organisations? How do we deal with dynamic change in data, objectives, units and high price for centralization and standardization of knowledge storage and retrieval?
- How do we deal with lack of knowledge structure? How do we define ontologies for organisational awareness?
- How do we deal with knowledge fragmentation?
- How can we guarantee a shared organisational awareness state? With information overflow?

Design and Development Methodology

The overall goal of the design and technological development process was to provide an agile framework to (i) explore in-vivo which use-cases provide real-value to a core group of users and (ii) address the challenges presented in the introduction chapter. This concerns mainly the question of how design, technology and dissemination of them can enable people to engage in organisational awareness activities such as mapping and exploring human capital and how their perspective can change as a result of using these tools.

This is extremely similar to what startups do, but with several important differences:

(i) The focus was non-commercial - we focused on creating an open technology and open data that can be freely used. (ii) Unlike a startup, we didn't target a lucrative market such as enterprise or businesses, but underserved groups that were bringing a lot of value to society, just not in the strict monetary sense: Universities, research institutions, non-profits and governmental departments. (iii) Our teams mostly consisted of interns, part-time volunteers. (iv) Due to budgeting constraints, we only worked in 6 month development cycles.

A report (Marmer et al. 2011) analysing 3,200 high growth web/mobile startups, showed that within 3 years, 92% of them failed. Being well aware of such risk and the fact that we were imposing on ourselves additional constraints beyond that of startups - we had to be creative about ways to produce valuable outputs.

Projects Background

The exploration process of this thesis was structured around four development cycles. Each cycle was headed by a different team, towards a different pilot project:

Project name	Core-development team	Date	Associated pilots
Knownodes V1	Liad Magen - full-time developer intern Alexandre Lejeune - part-time dev intern	1.10.2012	CRI - M2 Open Your Phd hackathon
Knownodes V2	Dmitry Paranyushkin - part-time intern Alexandre Lejeune - part-time intern Mikael Couzic - part time developer	2.6.2013	CRI - M1, M2 Mozilla workshop
Rhizi V1	Eyal Rotbart - part-time volunteer Owen Cornec - full time developer intern	13.1.2014	Team formation Conference demos

Rhizi V2	Alon Levy - part-time developer Amir Sagie - full-time developer	15.7.2014	CRI - M1,M2 Ohalo IGEM Descartes IT Leadership workshop
----------	---	-----------	---

figure 19: Different software projects during thesis.

Besides the core-team, we had a wide range of volunteers who helped with community building, hackathons and promotion of the software.

Stakeholders background

CRI - Centre for research and interdisciplinarity in Paris. They were the incubator and facilitator of the project, supporting the project in every way possible: Funds, talents, facilities and pilot projects audiences.

IGEM - International Genetically Engineered Machine (iGEM) Competition. This competition is between student teams competing around various categories in synthetic biology. The CRI has a formidable tradition and students have used Rhizi as part of their competition in 2015.

Sage Bionetworks - A non-profit dedicated to open medical data and software. They have partly financed Rhizi V2 and intended to use Rhizi as coordination tool between researchers.

Ohalo College - A teacher training college. They have partly financed and used Rhizi within some several courses.

Descartes IT - Paris Descartes University is a leading medical school. They have helped to finance part of Rhizi V2 and a pilot within their IT department was made in 2015-2016.

Leadership workshop - The CRI has organised a leadership workshop with participants from around the world. During the workshop the Rhizi V2 was introduced as a tool for capturing resources, concepts and ideas discussed.

Open your Phd - As part of a larger conference within the CRI, during a dedicated workshop , participants were invited to structure their projects(some of them Phd projects) inside Knownodes V2.

Methodologies background

Before diving into the specific methodologies that were used for the development of the different projects, it is important to explain the rationale behind choosing them. Therefore, it is important to understand the nature of the problem the team developing Knownodes and Rhizi, including myself, were trying to address and the constraints we were facing.

My claim is that the problem I was setting out to solve is within the domain of wicked problems (Rittel, Horst 1973). A wicked problem can be defined as a problem that is incomplete, contradictory, and subject to changing requirements that are often difficult to recognize.

Environmental degradation, terrorism, and poverty and software development are classic examples of wicked problems, and as Judd Ruggill and Ken McAllister claim, interdisciplinary collaboration is also one of them: “Collaborators negotiate different schedules, work demands, and ways of seeing, as well as resource pinches that hinder the process by which innovative digital media collaborations come to fruition. And yet, it is precisely because collaboration can be so wicked that it is so valuable. In constantly requiring collaborators to assess and reassess their rationales, artistic visions, and project objectives, collaboration makes for reflexive, complex, and innovative projects”(Ruggill & McAllister 2006).

It is how interconnected the problem is with other systems and issues that gives it these characteristics. Knownodes and Rhizi had to pave their own way within the french education and research systems and their associated incentive systems which are primarily directed towards individual or unit accomplishments rather than a inter-unit or whole-organisation view. Where hierarchical power structures that are deeply internalized despite everyone’s noteworthy efforts. Furthermore, our users possessed a lifetime of habits of digital-literacy that favored a linear and list/table based interfaces. A software ecosystem already rich with optimized and feature rich solutions that suit these habits and institutional culture was already established. As a result, as much energy that went into planning, designing and developing the system also went into generating conditions for adoption, including partnership building, user onboarding and coordination between the different groups who wanted to use the system but did not necessarily have the means to support its’ development by themselves.

In an article “Strategy as a wicked problem”(Camillus 2008) the author, a strategy consultant, points to several indicators of wicked strategy problems. Listing the points he made, I describe similarities with issues encountered during this thesis:

The problem involves many stakeholders with different values and priorities: Directors of institutions, middle-management and file and rank employees (or students) are motivated and incentivised by different things. Many students simply want a good grade and to finish writing a compelling thesis proposal at the end of their masters. A teacher wants the class to collaborate and exchange ideas. Different organisation leaders have different visions about the outcomes they hope to generate with the software, and these visions can change over time. For us, at one point we had to forfeit €15K of budget because one of the stakeholders expectations were completely different from the specifications of other clients.

Rather than the usual stable setting of a tight research group, the project span three continents with the core development team changing five times during that period, dealing with a web of interests from high education, research, non-profits, enterprise, startups as well as an open-source value system.

The issue's roots are complex and tangled: Subject of the thesis is extremely interdisciplinary: Social science, computer science, complex systems as well as user-interface and design were all important parts.

The problem is difficult to come to grips with and changes with every attempt to address it: Specific implementation of a software design and features may fit one group, but change as you move between groups. It was quickly noticeable that while we gathered some “super-users” that were extremely excited about the solution, there were others that were absolutely non-participative.

The challenge has no precedent: Each organisation and each pilot within the organisation has a similar yet discretely different set of challenges. For all of us this was the first time attempting these solutions. Review of previous literature was useful to some point, but since the technology and social landscape has changed in such a rapid fashion in the last 20 years, there was no reliable source to truly benchmark against.

There's nothing to indicate the right answer to the problem: As in any innovative social computing project, the only way to truly find out if a system works is to test it against users in context of their tasks and work.

So what kind of design and developmental methods can be adopted to tackle such problems?

The article writers suggest the following four strategies:

- **Involve stakeholders, document opinions, and communicate**
- **Define the corporate[your organisation's] identity** - Values, competencies, aspirations
- **Focus on action** - experiment with a number of strategies that are feasible even if they are unsure of the implications. Use pareto analysis to choose the most effective ones.
- **Adopt a “feed-forward” orientation** - Collect vision of the future, and anticipate future developments and needs.

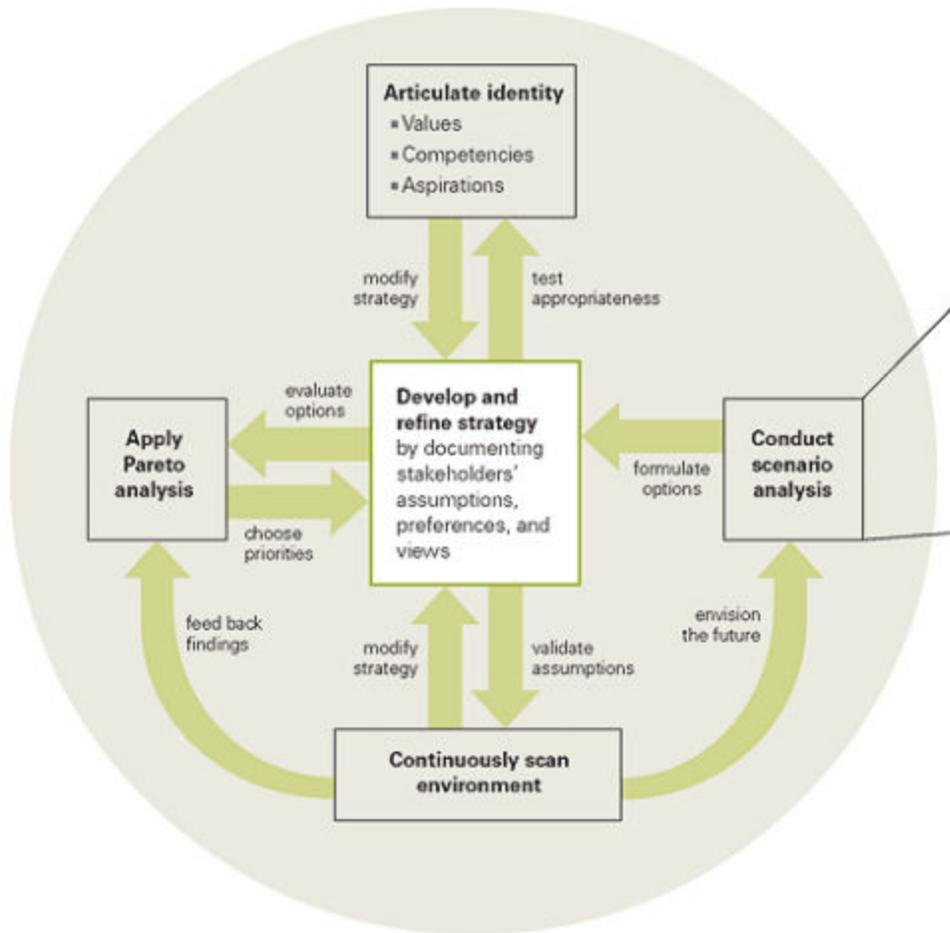


Figure 20: PPG's Framework for Responding to Wicked Issues(Camillus 2008)

Each of these methods will be detailed in the following chapter as well as specification of design and development practices that are part of the product development strategy. Our approach can also be attributed to certain schools within the social-computing field that think that within a rapidly-expanding technological frontier, building systems, rather than merely studying existing systems is a source of innovation(Bernstein et al, 2011)

My hope is that this will serve to inform others regarding their own strategies when working on projects that are both exploratory and software oriented.

Our strategy for approaching our wicked problem

Involve stakeholders, document opinions, and communicate

The thesis and it's process and documentation was completely transparent and open:

- The code, tasks and open issues were publicly available on Github. (See: <https://github.com/Rhizi/> and <https://github.com/CyberCRI/KnowNodes>)
- Mockups and designs were shared with stakeholders using Google Drive. Knownodes: <https://drive.google.com/drive/folders/0B7JAEailsrgMSXMwUFpLcmd4V1k> Rhizi: <https://drive.google.com/drive/folders/0B7JAEailsrgMfmZEUE5sbFNsNVA1QVJxVkdHSEJ4WHYyTXZOX3JfYU5xdXdFeENBdWpFZHc>
- Project description, list of objectives, deliverables and schedule were documented and hosted on Google drive and made available for stakeholders to edit and/or comment on.
- Major updates, milestones and prototypes were shared with stakeholders and discussed over meetings in person, email and videoconferencing tools.
- Special feedback sessions throughout the development process were initiated with end-users. From reviewing mockups, navigating wireframes, to in-class observations and feedback sessions, surveys and conversations, the end-users were always taken into considerations. This was following a user-centric design paradigm that will be later expanded on.
- Within the internal development team, we followed a daily standup methodology, where we took 10m in the start of the day, through IRC or Slack to discuss the challenges ahead of us and inform others what we are up to.

Relationship with all stakeholders were defined and maintained as partnerships. All involved knew we were working together to create a better technology for sharing knowledge.

Define your organisational identity

In order to navigate the uncertain waters of exploratory development, a strong organisational identity was always maintained and often updated to give a sense of identity and purpose to all involved. From the various documents created, I have curated some of the texts and images we used to convey our identity:

Values - What is fundamentally important to the organisation. In Knownodes, the emphasis was on open-source, open-data and enhancing software accessibility to non-experts. Our value declaration was: “We want to empower people to learn and innovate by making knowledge and ideas easier to access.” and “Let’s empower our society to better collaborate and innovate by providing insightful connections between knowledge and ideas.” as an organisation we described our raison d’être as: “Knownodes is an organization dedicated to building and maintaining an open sourced knowledge-building platform. Not only that this technology will enable learners from all over to world to collaborate in exploring how knowledge from all domains is connected and by doing so enhance interdisciplinary thinking and collaboration, but it will also help to identify open questions in science and research and help devise ways to answer them.”

Other description of our identity emphasized the potential of the open technical and legal structure as a tool to bring innovation(also see figure 21): “This open source, open data infrastructure paves new paths for collective intelligence, interdisciplinary and creative collaborations.”

Every part of Knownodes is open to reuse and extend

Open source and open data(cc0)

Full stack javascript

Interoperable and flexible



Figure 21 - Presentation slide of Knownodes project 24.10.2013

With Rhizi, while still emphasizing the value of open-source and open-data, our organisational values were also directed toward how we treated ourselves, this an extract from the mission statement made in 2014:

- “To connect between 7B+ people on this planet and empower them with a revolutionary new way of sharing, creating and exploring knowledge and insights.
- To create an open ecosystem around the technology. Open source and open data. Enabling the emergence of a global collective intelligence.
- Help communities become autonomous.
- We want to create a fun-loving organization where the Brightest and kindest people in the world talk between themselves from the heart, and not from the ego.
- To reach gender equality as well as eastern-western representation within us.”

In 2015 I've updated our mission statement to be shorter, incorporating the specific value we were offering to organisations:

“Mission statement: Empower people to do things together within institutions without needing middlemen”

- (1) To enhance organisational awareness by building and distributing an open tool to map and find resources and opportunities.
- (2) To build a sustainable business where bright and kind people can joyfully work together.“

Define product identity

A large part of the work within social computation is to convey the value and vision of the software to users. The reason for that is to encourage adoption, to attract funding and interest from collaborators and also for inner team morale who is dealing with wicked problems and could use having a sense of the potential and value of the software. Below are a few examples of how we articulated the product's identity. Some are short sentences, other are longer explanations that convey the philosophical meaning behind it.

Competencies

What is it we do better than others do?

The biggest differentiator we have emphasized was our rich technical background, and our diverse network of innovative collaboration institutions: “Our core team resides in Tel-Aviv and includes a development team with background in Israeli startups and software(Waze, Red-Hat). We have contracted customisation services for the software(CRI, Sage-bionetworks), and have several potential clients in the Pharma and higher education industry interested in our solution.”

Aspirations

How does the team envision and measure success?

Knownodes

The name Knownodes was chosen because it it both meant “nodes of knowledge” as well as when spoken out loud could be heard as “No nodes”. As the main value we tried to convey was the use of information rich edges, and putting meaning in the connection, it was a concise way to name the software.

“An open, customisable middleware for groups of people to organise their knowledge and actions.”

“Knownodes is a web platform that enables people to create emergent knowledge networks by creating contextual connections between knowledge resources, problems and ideas.”

“Knownodes is a web platform that makes it easy for people to work together on open projects while encouraging interdisciplinary, community participation and accessibility.

Within Knownodes, we had stated the following as our long-term aspirations:

- Raising awareness and collaboration between different scientific disciplines.
- Making the problems of science and research accessible to scores of non-experts.
- Creating an online community for problems solving and idea making that enables sharing of knowledge in a powerful way.
- Approaching knowledge management with a radically new perspective that will induce creative solutions to old problems.

Another aspirational aspect was added to Knownodes thanks to the help of Alexandre Lejeune who helped design and articulate a model for a parallel layer above the web, where connections could be made between ideas, people and web-pages:

“We have the rather ambitious goal of applying the same logic to the rest of the web, allowing people to create a participatory parallel universe of the web that gives the community a set of tools to share connection, peer-review and discuss them and get updated about knowledge resources they follow.”

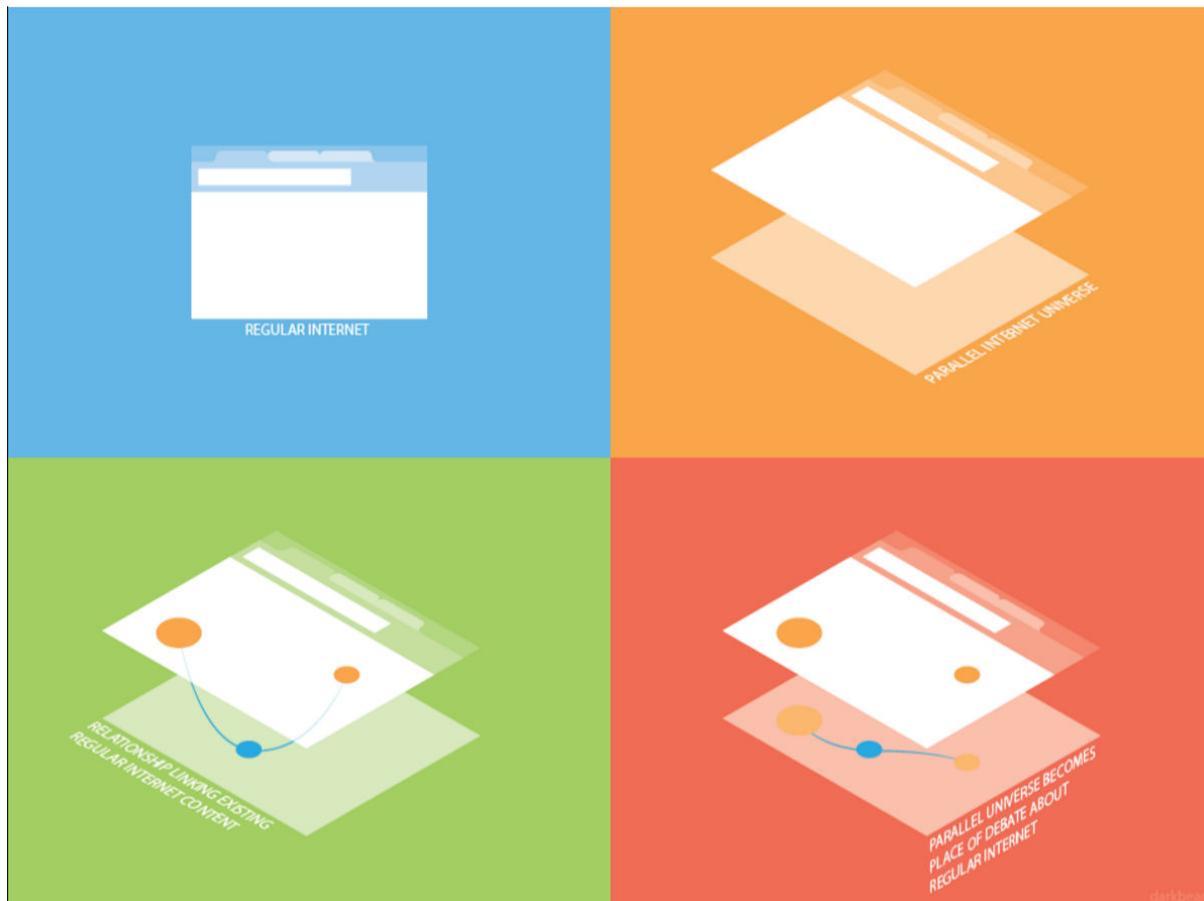


Figure 22 - Knownodes vision: Create an open parallel universe on top of the world wide, where people can connect and debate about content.

Rhizi

We had a prolonged discussion in our team about how to name our project. Below is a short description I wrote to our team suggesting the name Rhizi:

Rhizi is a funny name.

Fun to pronounce. Short. Can act as a verb. Serves a similar functionality to the prefix “Wiki”. You can “Rhizify” your community, create your own Rhizipedia, RhiziCourse, RhiziClass, RhiziData.

Pronounced out loud it sounds like “easy”, and with an american pronunciation can sound like Rise(y), a small rise :)

It’s a bit quirky as well with this funny “h” stuck there, why not rizi, anyway? Or why not wizi, pisi, fizi or whatever?

“In botany and dendrology, a rhizome (from Ancient Greek: rhízōma "mass of roots", from rhizóō "cause to strike root") is a modified subterranean stem of a plant that is usually found underground, often sending out roots and shoots from its nodes.

Rhiza = root.

Gilles Deleuze and Félix Guattari use the term "rhizome" and "rhizomatic" to describe theory and research that allows for multiple, non-hierarchical entry and exit points in data representation and interpretation. In *A Thousand Plateaus*, they oppose it to an arborescent conception of knowledge, which works with dualist categories and binary choices. A rhizome works with planar and trans-species connections, while an arborescent model works with vertical and linear connections. Their use of the "orchid and the wasp" is taken from the biological concept of mutualism, in which two different species interact together to form a multiplicity (i.e. a unity that is multiple in itself). Horizontal gene transfer would also be a good illustration.”

This funny sounding name is shrewd because it packs within it the most profound philosophical aspects of Knownodes. The things that bond all of us together to this journey in the first place.

We are a bottom-up movement. People driven structure. Freedom to connect, contextualize, and revolve communities around content for communities by communities. It’s about people inter-sprouting roots of understanding from their individual minds.

"As a model for culture, the rhizome resists the organizational structure of the root-tree system which charts causality along chronological lines and looks for the original source of 'things' and looks towards the pinnacle or conclusion of those 'things.' A rhizome, on the other hand, is characterized by 'ceaselessly established connections between semiotic chains, organizations of power, and circumstances relative to the arts, sciences, and social struggles.' Rather than narrativize history and culture, the rhizome presents history and culture as a map or wide array of attractions and influences with no specific origin or genesis, for a 'rhizome has no beginning or end; it is always in the middle, between things, interbeing, intermezzo.' The planar movement of the rhizome resists chronology and organization, instead favoring a nomadic system of growth and propagation.

"In this model, culture spreads like the surface of a body of water, spreading towards available spaces or trickling downwards towards new spaces through fissures and gaps, eroding what is in its way. The surface can be interrupted and moved, but these disturbances leave no trace, as the water is charged with pressure and potential to always seek its equilibrium, and thereby establish smooth space."

Everything is connected, but the connections are obscured by outdated perceptions and interfaces. Rhizi empowers us to make these connections explicit. It is the cause for us to strike root.

In another document, and as our vision evolved, a shorter explanation of the product's vision was put down in writing:

“To build the spreadsheet of tomorrow: Make it accessible to non-experts, empowering communities to better manage and communicate information.“

For Rhizi in 2015, our aspiration was attached to our core value proposal:

“Make organisations self-aware.”

Focus on action

In order to optimize the exploration process, we have adopted the following three strategies:

- 1) Adopt rapid-prototyping and agile process so to test our hypothesis as fast as possible - can also be referred to as: The lean startup approach and expanded on chapter [2.1 On usage of Lean startup methodologies within a Phd](#). This was done by adopting rapid deployment practices within the code. Following deployment, following up with users and stakeholders and gather feedback. We also organised did many ideation sessions, meetings, workshops (See chapter [3.7 Other pilot projects](#)) and interviews with collaborators from all levels: end-users, administrators and feedback sessions within the internal development team.
- 2) Initiate several pilot projects, targeting different value propositions within different contexts

to optimize chances of discovering a useful test-case that can be then scaled up. From a code perspective we developed a customisable data schema configuration file, that allowed us to modify node types, change URL's, deploy multiple instances and provide each instance with their own database instance. Description of the different pilot's and process is specified in chapter [3 - Pilot projects and results](#).

- 3) You learn by doing: nobody taught us how to run a team, manage product cycles or create a community. These were things that were in our responsibility to experience and expand through learning and doing. We did so by contacting and gathering feedback from mentors and experts that had different perspectives and more experience on the different domains. Reading and applying project management techniques, adopting new collaboration technologies such as Slack and Github. We took on-board people who were extremely adaptive and competent with regards to their development abilities and were not afraid to learn and experiment with novel technologies.

Adopt a “feed-forward” orientation.

The meaning is rather than just learn from feedback processes, we should continuously scout for diverse opinions from different stakeholders to form an ever-evolving landscape of the opportunity space. We facilitated this by in-person interviews, regular meetings, regular reviews and thinking sessions about new technology initiatives with potential collaborators.

Lean Startup methodologies and their application

In his seminal book, “The Lean Startup”(2011), Eric Reis captures a lot of insights and methodologies for helping startups build innovative systems. The basic premise of the book is that with technology and contemporary management practices we can increase productivity in an unprecedented rate. Thinking lean - small teams, small scale products and just in time solutions shorten product cycle times. Core measurements are not of the physical products created, but by validating value hypothesis - in other words, what we assumed people would find useful in a product, and what within that assumption is validated and which parts need to be changed or rethought. The objective is system wide - to quickly build the right thing to solve a problem, rather than assessing work productivity in number of hours or direct output such as lines of code. And lastly using the build-measure-learn to quickly test assumption and learn if a sharp turn in the product, a “pivot”, is needed.

Validated learning is a method in which the startup learn from end-users what works. It is backed up by empirical data from users “in the wild”, rather than within a lab. The objective is to synthesize, through the experimentation process, between the founder’s vision and what end-users would accept. It’s a tricky dance between understanding the underlying needs of your users that might not be exactly what they explicitly ask for and being a deep listener that avoids telling users what they are supposed to want.

The experimentation process begins with formulating two types of hypothesis:

The value hypothesis - Verify that the product delivers the expected value to the end-user.

The growth hypothesis - Test adoption and spread of a product across end-users.

As growth per-se was not our focus, we concentrated on the value hypothesis. However there’s a clear connection between value provided and the way people contribute and help spread the software. This was something we have encountered and surveyed for.

To test the value hypothesis is to measure whether people are willing to invest energy to use the product and/or spend money to support it.

As part of our value-hypothesis testing - we conducted discussions with many different potential partners where we presented our value proposal and discussed potential collaborations. Among those participating in the discussion: Israeli ministry of education, Weizmann institute, PNAS, Sage Bionetworks, Servier, An insurance company, The leadership of Sorbonne Paris Cite and others. The end-result from these conversations was a collaborative spreadsheet reviewing potential

use-cases, which initiated a collaborative process involving our team members voicing their opinion on the different options. The document containing the synthesis of our exploration process can be found in [annex 6.6 Product scenarios](#).

Following this methodology creates a feedback cycle illustrated in figure 23:

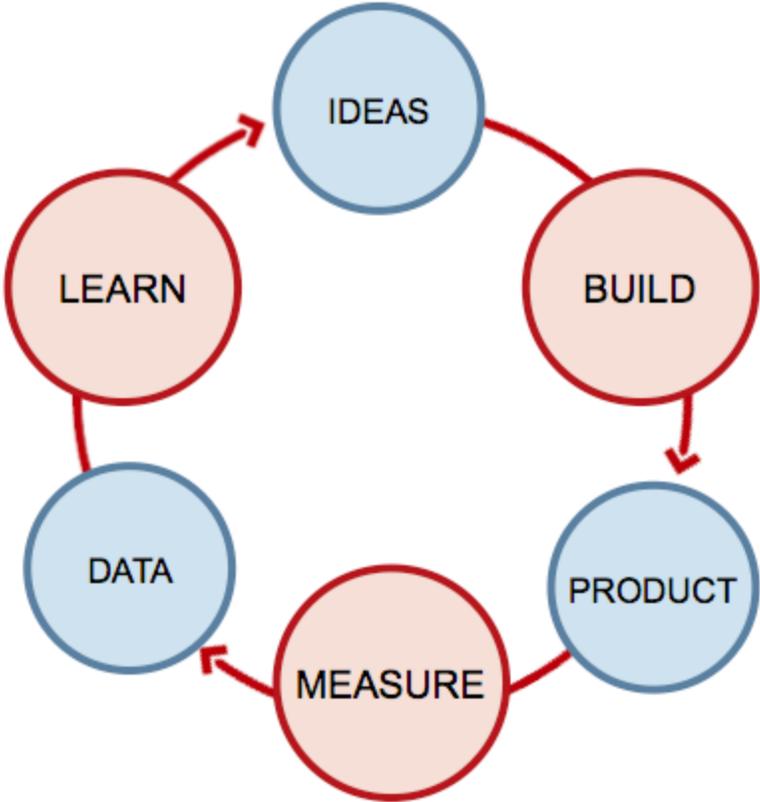


Figure 23: The lean startup feedback loop.

In order to support the learning process, it's important to minimize the total amount of time each cycle through the loop takes. The first step in achieving this is through following a Minimal Viable Product methodology(MVP). It is defined as the product that has the least amount of features and requires the least amount of work that enable a full turn of the feedback loop.

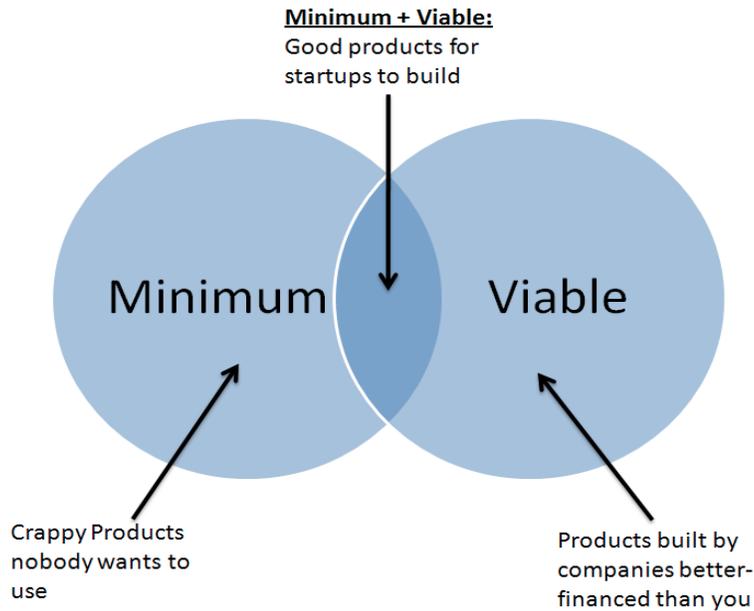


Figure 24: Venn diagram of an MVP

Product	Value Hypothesis tested
Knownodes prototype	Capturing knowledge with Information rich edges will revolutionize the ability of people to share knowledge and related ideas to non-experts
Knownodes V2	In order to collaboratively build knowledge additional social features such as notification, reputation and voting must be built
Rhizi prototype	A full visualization interface will provide users the tools to be fully engaged in co-constructing knowledge as well as explore it
Rhizi V2	Mapping human capital allows organisation members to better collaborate between units.

Figure 25: Main development hypotheses for each product.

Inevitably, an innovative hypothesis must be a leap-of-faith one, after validating it, one can go and change variables in an effort to optimize value and growth. In the initial leap-of-faith phase a customer arch-type is constructed to humanize the target, and inspire creative thinking and understanding.

Below is an example document we have created in early Rhizi V2 for the CRI:

Veronique and Laurence - Administrative Team

Position

Coordinator of administrative staff. Employee at the CRI.

Goals

- **Handle student registration and other administrative management**
 - collect applications for CRI
 - maintain a database of the student body in the CRI - demographic statistics and information of students' specific educational activities
 - handle the Internship submission and approval process
- **Interact with external entities**
 - maintain a database of contacts for the CRI, such as contacts in labs or in universities
 - collect data about internships from students and updating the labs database.
 - contact lab supervisors and check regularly that the contact information is correct and up to date
 - target the right persons for communication. flag contacts that are known to be dead ends.
- **interact with the students**
 - send notifications and updates
 - filter notifications by courses, and other organizational rules
 - two major categories of announcements - general information for all CRI students such as events and specific information required as part of the curriculum.
 - follow students careers after graduation

Current tools in use

- custom made database
- moodle - open source software for managing moocs
- direct mail tool for messaging students
- webform for internship submissions

What's working for them

- **moodle**
 - The moodle handles the students interaction as part of their course requirements pretty well
 - easy to announce administrative information to a specific group of students
- direct mail tool
 - customizable mail templates - 'Dear Jeremy'...

- easy to send reminders only to students that hadn't completed some requirements

Pain points

- **Handle student registration and other administrative management**
 - webforms for internship submission propose no way to effectively follow the status of forms filled, no overview on status, to see if students need help in finding internships (need to manually check: export pdf to excel)
- **interacting with the students**
 - the moodle is considered as spam because of insufficient filtering and too frequent emails
 - **“This is why (the moodle) is not really a place for discussion, its more a place for announcements”**

Anirudh - Student

Position

M2 Student

Goals

- To fulfil their course requirements: upload articles and comment on them
- apply for internships
 - get information about specific labs
 - find out about other students that went to a lab
 - read student reports from internships

Amodsen - Teacher

Position

Teacher, head of program

Goals

- **manage the academic activity (courses)**
 - To monitor students interaction with other students about articles discussed in class
 - organize and share course schedule
 - announce administrative stuff
 - direct and comment on discussions in a course
- **announce events and general information to the CRI community**

- **communicating to the outside world that the CRI has an active community that share information**
- **communicate to the outside world quantitative data about projects and students' success**
 - what students did after attending the CRI
 - following people who briefly worked with the CRI - e.g. interns
- **approve internships**
 - learn about specific labs the student is interested in
 - get students abstract proposal
 - assess whether the lab and the student are a good fit
 - connect students with past students that had experience with specific labs

Current tools in use

- use moodle to message groups of students or all students
- use offline tools such as meeting, clubs, seminars to encourage collaborations between students
- uses word for producing project reports and excell to generate charts

Pain points

- have difficulties messaging students using the Moodle - treated as spam
- when you post an event on Moodle there is no way to anticipate the turnout
- **“After 10 years that the CRI exist... we are repeating the (internship) process every time without information from the past”**
- there is no centralized database that contains valuable data about past internships an labs

Margot - community manager

Position

manage communication between students, clubs, teachers, external stakeholders.

Goals

- get a clear overview of what everyone is doing - students, projects. etc.
- communicate the interesting things happening in the CRI community
- promote externally and internally what students are doing in the CRI
- keep a database of researchers from all over the world to help in organizing workshops

Current tools in use

- SPIN blog platform
- moodle - read posts students are publishing about talks

Pain points

- no easy way to collect up to date data about stuff going on
- talking face-to-face doesn't scale
- 'spamming' on Moodle is inefficient and don't get results
- **“we all have our own databases that we're creating from time to time. Maybe in the next office they have very valuable information for me, but how could I know?”**

Which is processed into a summary towards an MVP:

Rhizi UX

User Experience

What makes for a happy user

CRI Student:

- I can quickly find internships, projects and people related to skills I'd like to learn and fields of research that interest me.
- It's simple to understand and complete the mapping tasks I'm assigned to do.
- My contribution to the map is meaningful as other people see it and it helps them out.
- I can see what my classmates and collaborators are interested at.

CRI Admin:

- I can get a comprehensive information about a student: Internships he did, skills he has, things he's interested in, projects and labs he's involved with.
- I can advise students about projects and people related to their interests by using the Rhizi map and search.
- Students can get notified about relevant things to them without getting “spammed”
- I can identify staff members that can help me with small tasks I need to do.
- As HR manager, I can plan training sessions from sending out a mapping task to staff.

Descartes IT Employee:

- I can quickly find the person related to my skill/project and contact him for help
- I can see what colleagues are working on and what they know

Descartes IT Admin:

- I can show outsiders the system level view of what we are working on.

- Outsiders of the department can find out who can help them with a technical task or technical consultation.
- I can figure out what skills we lack or if we have bottlenecks and plan training and hiring accordingly.

Which turned into a simplified user cards and profiles:



Figure 26 - Example of user archetypes used in Rbizji V2

After exploring what users want, an MVP document is created to capture the important features that are included in the product MVP.

The next step in the lean startup methodology is measurement. This is done through a framework called innovation accounting and consists of measuring what is happening right now inside the system, devise additional experiments and/or optimizations to improve on results and finally, decide whether the current direction is working and requires additional iterations or it's time to pivot. The sign of a successful pivot is seeing higher adoption, activity and participation or whatever it is you set as the overall metrics you have set for yourself.

Within our project we had a total of 4 iterations, our first iteration(Knownodes V2) was around optimization of the tool and adding additional features. This was a result of encountering bugs and problems in adoption, which in retrospect might have been better solved with a pivot.

The second development cycle led to what is called a zoom-in pivot, meaning that one of the features in Knownodes, the map interface, became the product. The rationale was that following a pilot project, surveying and user interviews(see pilot results in chapter [3. Results and Pilot projects](#)) we saw users were lost within the list-based interface. We decided to a network visualization-first approach would be at the basis of our new hypothesis.

The third development cycle, in which we prototyped text-to-graph interface was also a customer need pivot, we noticed a much more focused need for students to discover internship and potential collaborations, then knowledge mapping within a course.

Our fourth development cycle, led us to uncover frustration and serious blocks in innovation and workflow due to lack of human capital awareness. This led to both a Client segment and client need pivot as we moved into large organisation and inter-unit collaboration as well as the need to find problems solvers outside one's team as the focus of the product.

Limitations within academic research

While using this framework was definitely productive we have encountered various problems that are important to mention, in the book the author mentioned three structural aspect for a startup that was badly lacking in our case:

- 1) **Scarce but secure resources:** Since clients insisted on breaking of payments according to milestones, we found that this often caused any unexpected delay to potentially sabotage the entire project. This was a cause of stress and demoralized the team. In addition, clients not paying in a timely way introduced significant loss of time and energy.
- 2) **Independent authority to develop the business:** Working with clients instead of an investment forced us to develop features we didn't deem as necessary, just in order to please clients. In other cases, we managed to explain our rationale to the clients, but only after a lot of efforts, which defocused us from our mission.
- 3) **Personal stake in the outcome:** While me and the team working on the Rhizi V2, had a personal stake and this had dramatic effects on the quality and output of the work done. In previous iterations, when working with interns and temporary hires, it was clear that the lack of personal stake led to considerable losses as interns went to work on other projects and the energy that went into their training as well as their specific expertise went with them.

From Human-centric design to Activity-centric design

The human centric design approach has been described by Giancomin (2012) as a framework that poses questions about a certain design, explores answers to these questions through observation and interaction with those who are experiencing the product or a representation of the product (mockup, wireframe etc.), and lastly, iterates around these answers to produce a better design. The questions the designer asks are structured as a hierarchy (See figure 27) and range in complexity: from basic scientific assumptions about human capacities such as motor-skills, cognition and emotions, to more complex question about human interaction, socialization, and ultimately, about the meaning people get from their interaction with the design.

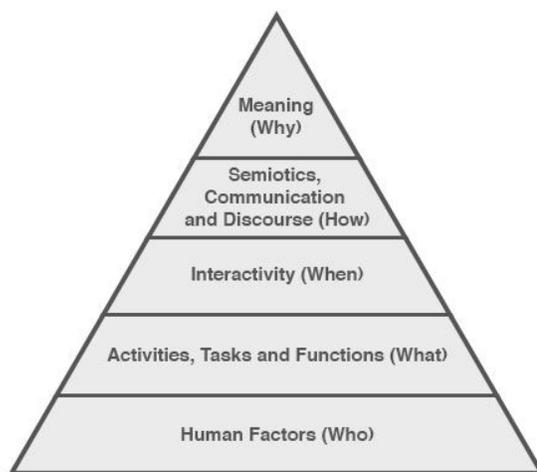


Figure 27: The human centered design pyramid (Giacomin, 2014)

Conversation and feedback between the designer and the end-users is emphasized in human-centric design. The designer functions as a deep listener and facilitator who translated the will of the people involved into a tangible result. A set of tools such as questionnaires, ethnographic studies, development of personas (documentation of user profiles and their backstory) and tools to simulate the design such as prototypes, mockups and role-playing are used to capture those needs, desires and meanings.

In this chapter I will detail the various ways I engaged in this conversation, as well as share some of the insights gained from the process and describe the migration into an activity-centered design methodology.

The question of meaning is key as this is determines if a product will be adopted, loved or used.

A design that introduces new meaning to a person's life is considered within this framework as the highest achievement and something that has high commercial value. This might explain why product landing pages start introducing their product with an answer to the deeper meaning behind it.

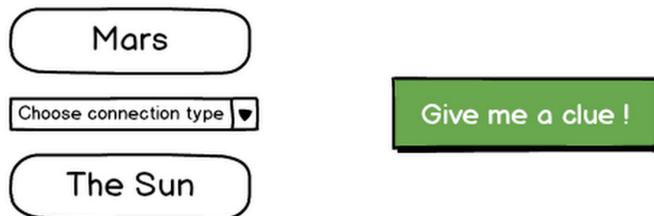
Defining objectives with stakeholders

In addition for meeting and conversation, for each pilot project, we have prepared a set of collaborative documents that were shared with the stakeholders and defined objectives, a project description and project details. These documents were a rich tool for conversation regarding expected outcomes and needs of the stakeholders. For an example document, see appendix 5 [Digital organisation - example project page](#).

Mock-ups, wireframes and illustrations:

In order to quickly receive feedback about our ideas, we had to translate them into a visual representation of the finished product or feature. This helped us filter out many ideas before spending lots of time on implementation, and engage in discussion with stakeholders in the project. Below are examples of some of the mockups that were developed:

Do you know the relationship between...



The mock-up shows a quiz interface. At the top, it asks 'Do you know the relationship between...'. Below this, there are two rounded rectangular input fields: the top one contains 'Mars' and the bottom one contains 'The Sun'. Between these two fields is a dropdown menu with the text 'Choose connection type' and a downward-pointing arrow. To the right of the input fields is a green rectangular button with the text 'Give me a clue !'.

Figure 28: Mock-up of connection centric quizzing platform (Credit: Mikael Conzic)

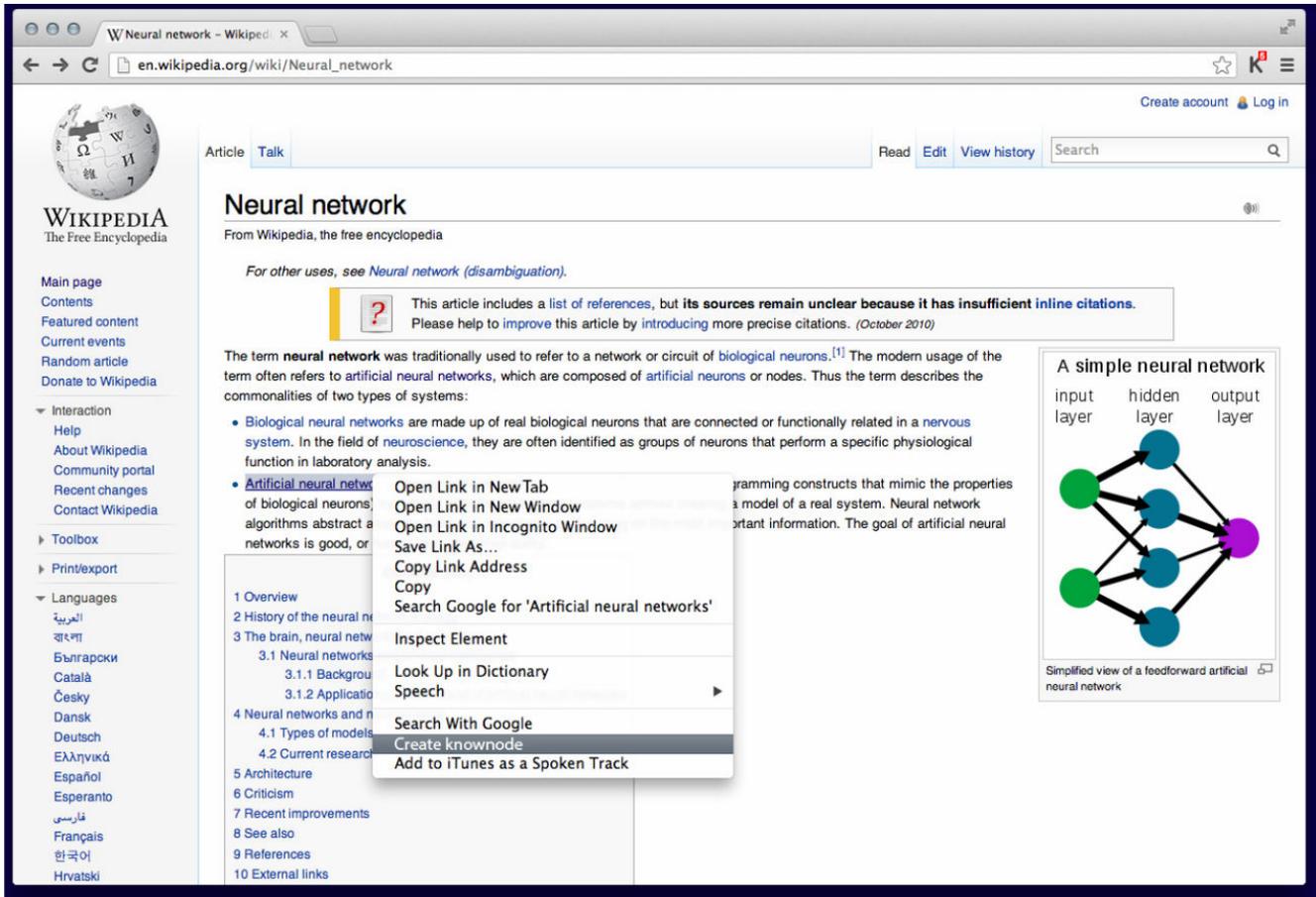


Figure 29: Mock-up of a browser plug-in to connect between webpages.

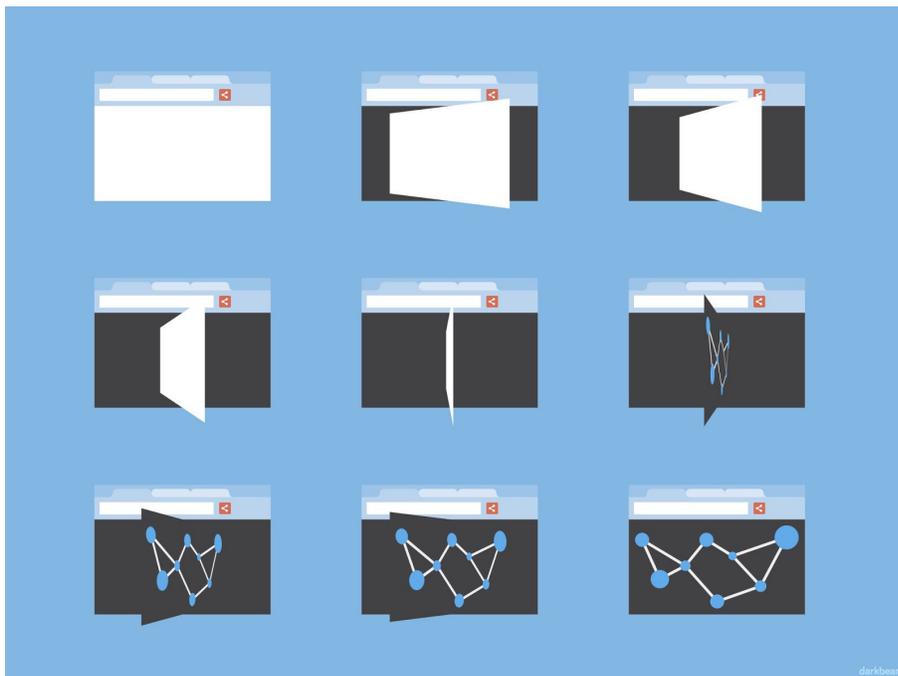


Figure 30: Flip animation of browser plug-in for Knownodes. (Credit: Alexandre Lejeune)

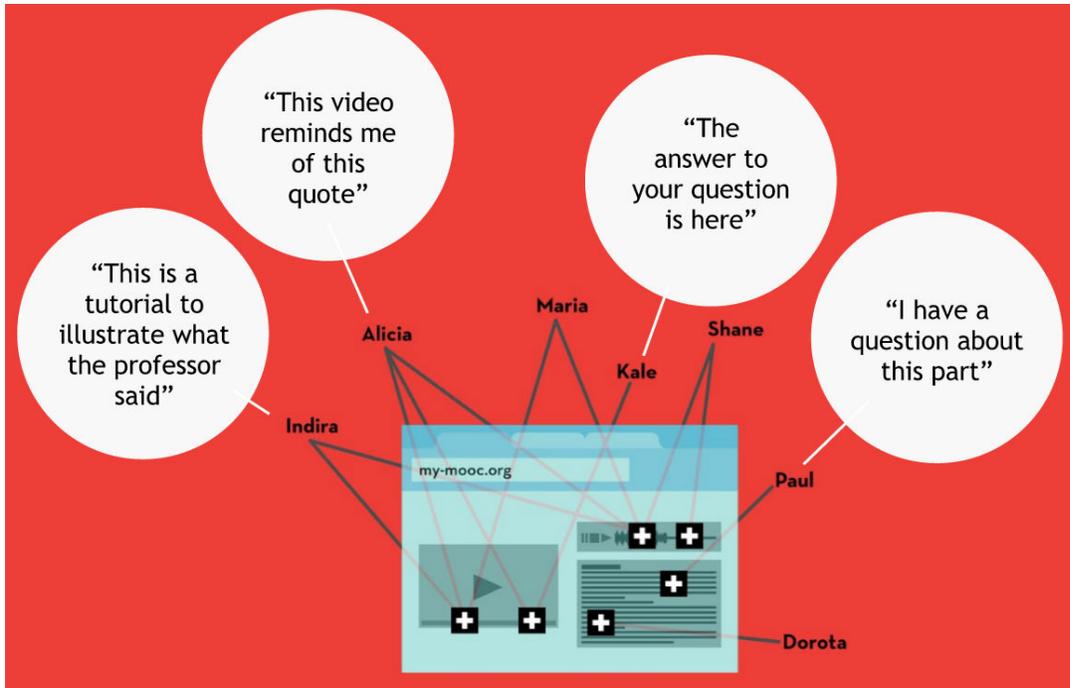


Figure 31: Mockup of an annotation system for MOOCs

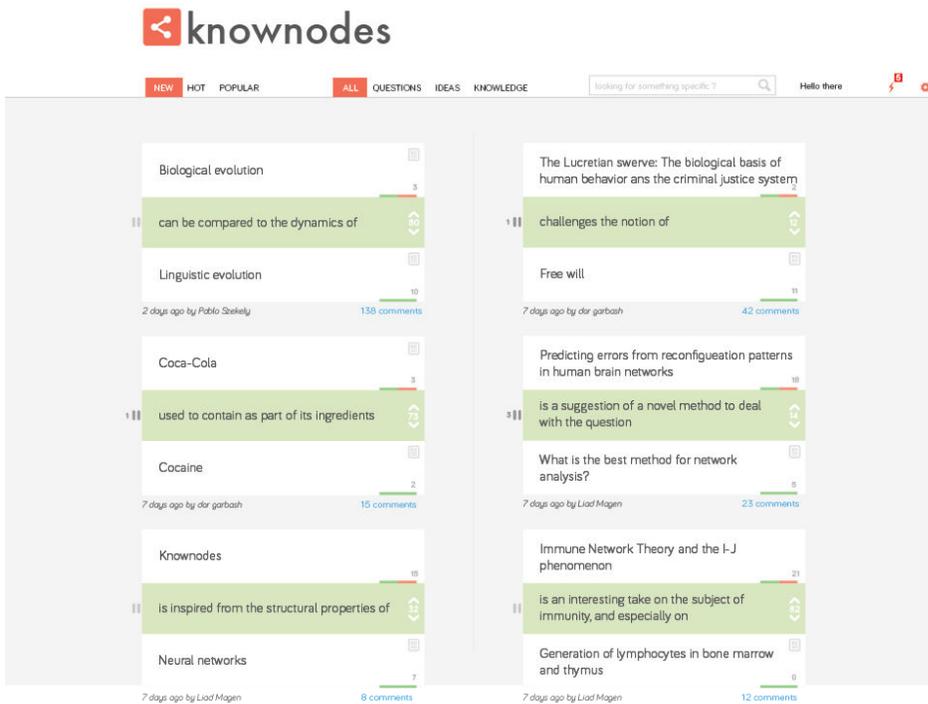


Figure 32: Mockup of Knownodes page interface

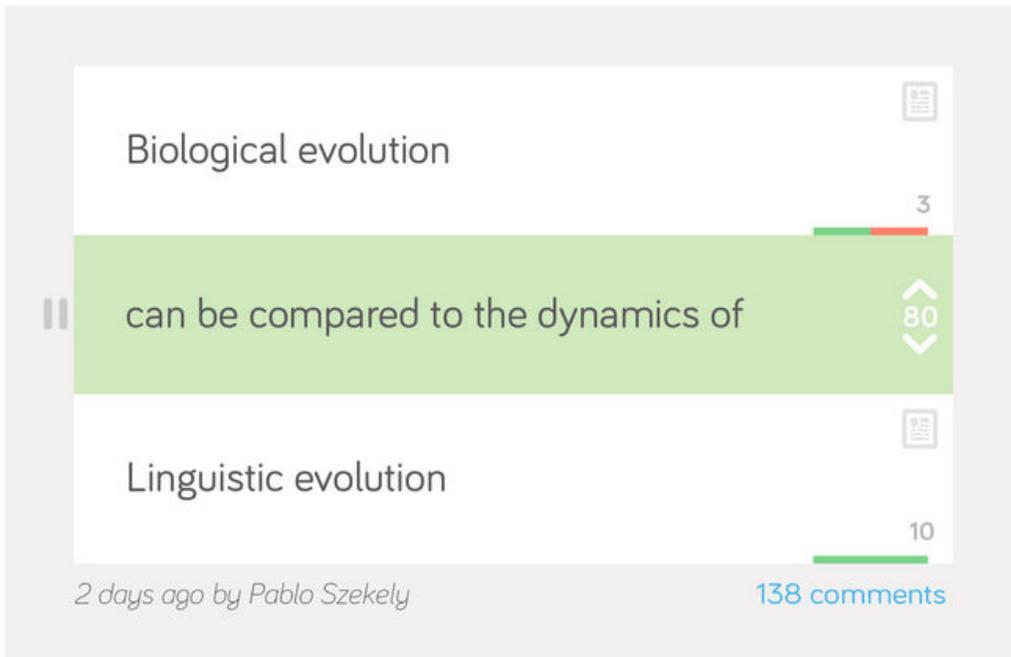


Figure 33: Mock up of a connection between concepts in Knownodes

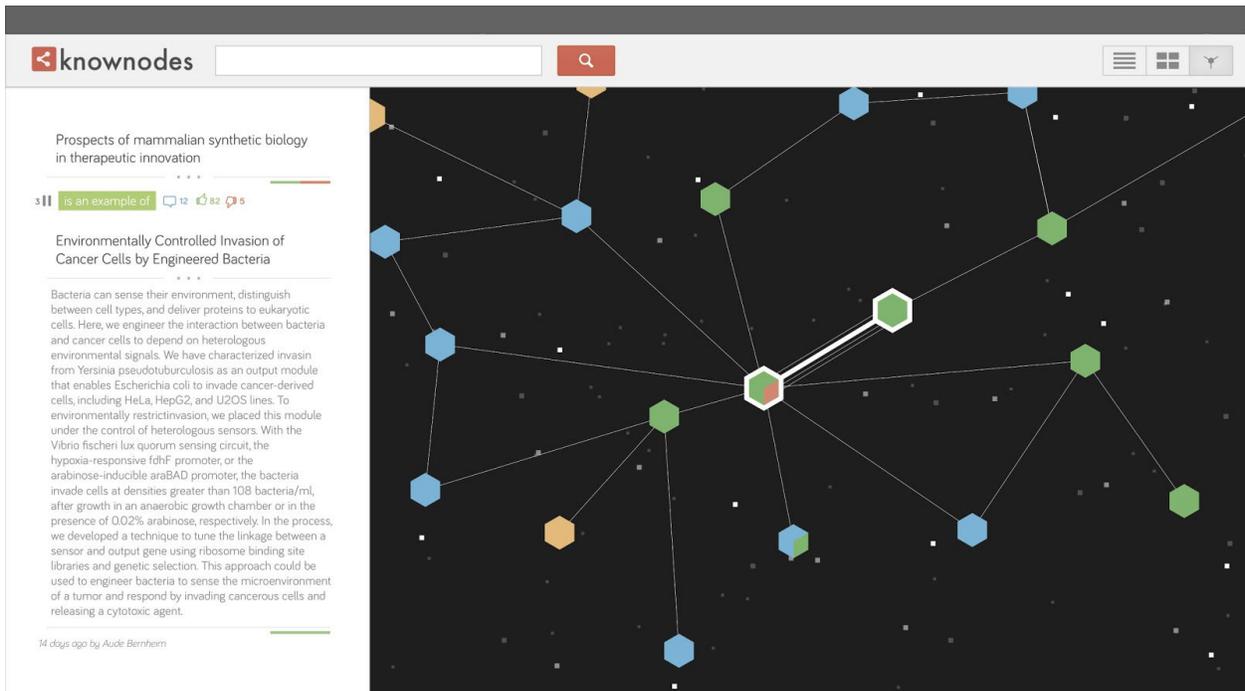


Figure 34: Knownodes map interface. (Credit: Alexandre Lejeune)

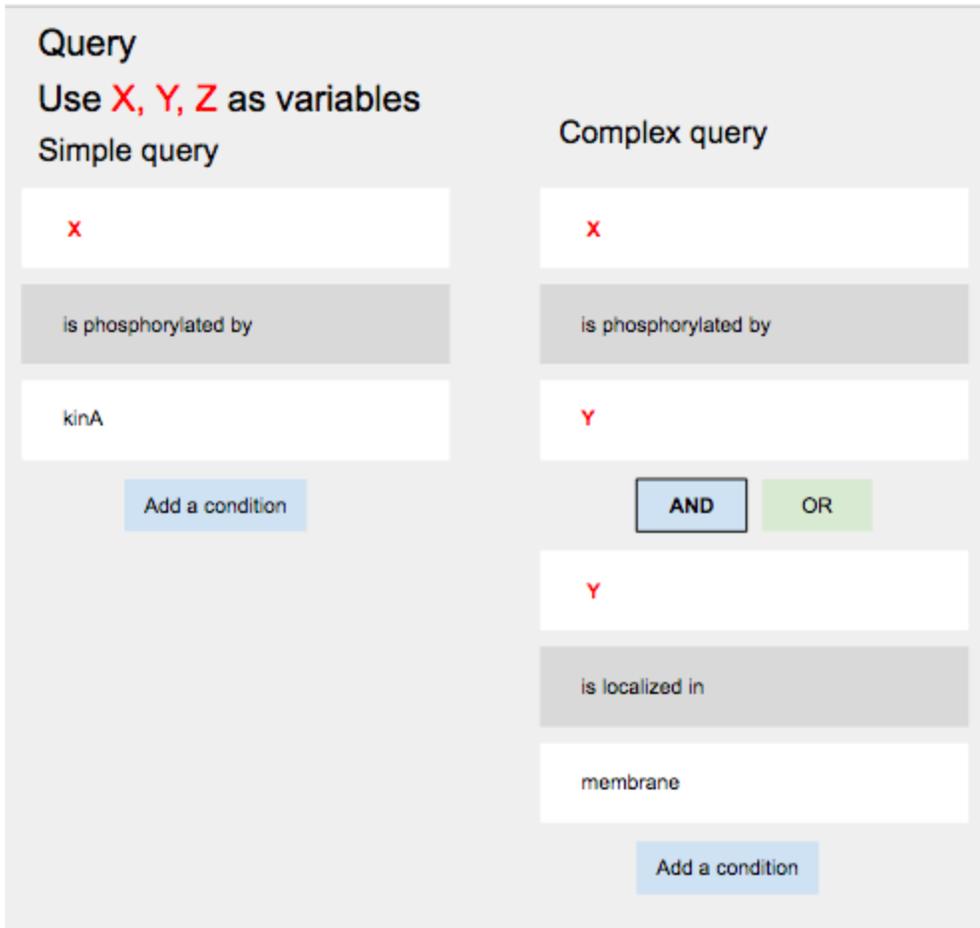


Figure 35: Knownodes query interface mockup (Credit: David Bikard)

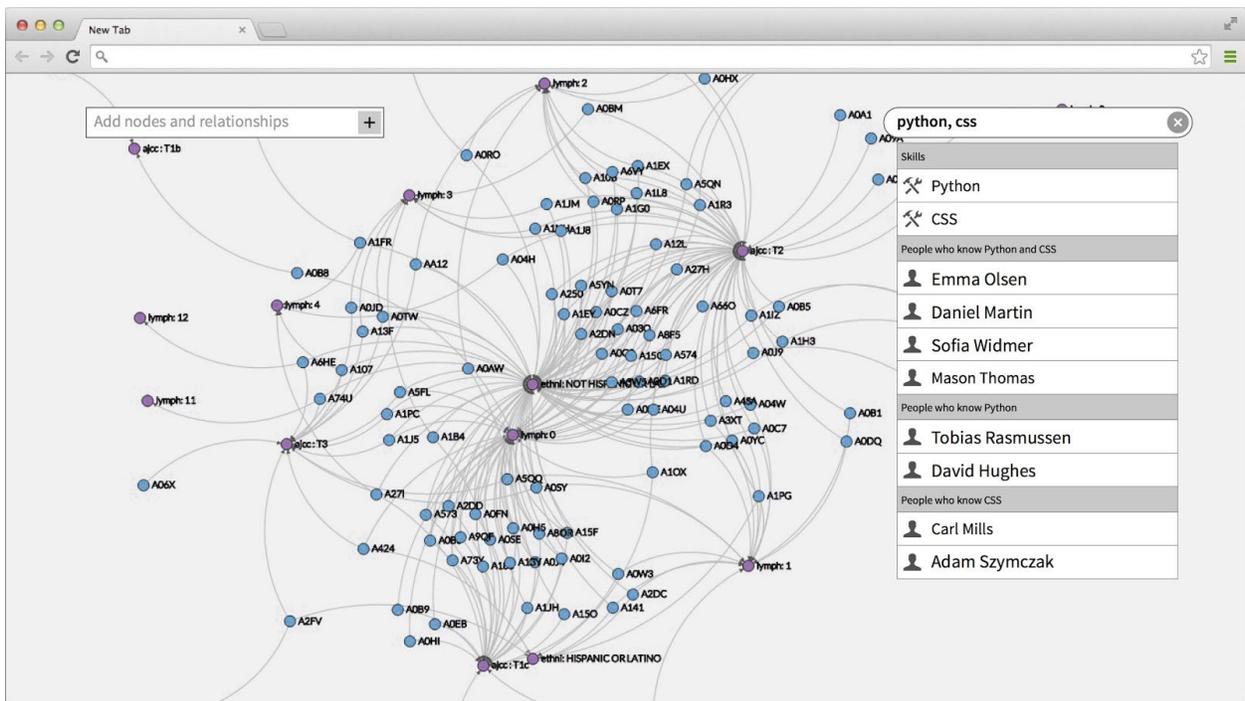


Figure 36: Rbizi search results mockup



Figure 37: Rbizi Gantt chart view mode

Rapid-prototyping

After initial user research, mockup and wireframing to filter out ideas and understand users' needs better, the next step is getting a prototype into the hands of users for further feedback. This process is covered in Wexler & Fishbein(2015), which state that the main goal of this methodology is to mitigate risks by validating assumptions and tackling the risks associated with the product early in the process. The process starts with generative experiments, where problems and pain-points are identified through an exploration together with the target audience. It continues with evaluative experiments where a simulation of the end-result or certain features of the software are presented or used by the target audience in order to gain effective feedback. For example, In several projects we gave participants a prototype to use within a workshop. See appendix 4 [Other pilot projects](#) for a description of the Leadership conference, the Open science conference and map your Phd project for a detailed description of the scenarios. Within these cases we got into a discussion with the communities regarding problems they were having, we tested together with them a solution, where they participated in a mapping task and/or viewed simulated information we added to the system. Following these we gathered feedback from participants, and following all these cases, saved a lot of work by deciding to not put more energy into them.

Interviews and surveys

Interviews were done with end-users, using semi-structured interviews.

From user-centric to action-centric

In his seminal article *Human centric design considered harmful*(2005), Don Norman argues that traditional human-centric design puts too much emphasis on understanding and adapting to the needs of users, rather than focusing on the specific activity that needs to be done using the design. Citing usage of tools ranging from garden, woodworking tools, music instruments and writing tools, vehicles and technologies such as word processors, he determines that it is actually the understanding of the activity, rather than the user, that allows for them to be more-or-less with a similar design globally. It is through understanding activities that good design is generated. People will adjust and learn the tools if the tool is well designed for the activity. One of the dangers of human-centric design is that over fitting a tool to a specific group might result in the design outcome being less appropriate for others and/or adaptive to future uses. Human-centered design can defocus the designer from understanding and working towards best enabling the activity: The methods criticized include emphasis on static views instead of looking at the interrelated tasks that an activity contains, over-listening to user complaints rather than having a cohesive model of the activity that the tool should support.

One of my main take-aways from the design and technical development during this thesis is that a lot of the features developed in Knownodes and Rhizi were over-emphasizing listening to immediate feedback from users, often reaching a local minimum, where hard to produce features results in very small incremental improvements in user experience. Developing a good model of the activity you intend to support should be the focus when building a tool, and it is often better to insist of building something that users are willing to learn how to master it in order to accomplish tasks that are important for them.

Technical approach

While there are many frameworks for software development, within the various development teams we tried to favor the Agile methodology as it works well with the lean startup methodologies to ensure that the software gets as quickly as possible into the hands of users and achieve as many feedback cycles out of the development process.

Within the agile manifesto the principles of adaptive planning, evolutionary development, early delivery, and continuous improvement, and rapid and flexible response to change. The principles of agile development are stated in the “principles behind the agile manifesto”(2010):

1. Customer satisfaction by early and continuous delivery of valuable software
2. Welcome changing requirements, even in late development
3. Working software is delivered frequently (weeks rather than months)
4. Close, daily cooperation between business people and developers
5. Projects are built around motivated individuals, who should be trusted
6. Face-to-face conversation is the best form of communication (co-location)
7. Working software is the principal measure of progress
8. Sustainable development, able to maintain a constant pace
9. Continuous attention to technical excellence and good design
10. Simplicity—the art of maximizing the amount of work not done—is essential
11. Best architectures, requirements, and designs emerge from self-organizing teams
12. Regularly, the team reflects on how to become more effective, and adjusts accordingly”

Other methods related to the agile methodology were adopted including creation of a product roadmap(Example document in chapter [4.1 Proposal for a future project](#)), product vision, user stories(See chapter [2.1 On usage of Lean startup methodologies within a Phd](#)), defining what is “done” within issue and feature tickets, customer feedback, continuous integration, rapid deployment and daily standups.

While we tried our best to use a lean startup approach and be completely agile, the material and systemic conditions forced us to incorporate several waterfall techniques. Mainly to agree in advance to an extensive set of features. In order to negate the issue, we have broken down the set of features into smaller release cycles and after every feedback round, I tried my best to renegotiate with the stakeholders and explain to them our findings and challenges that explain why a shift in requirements is essential.

For our collaboration we used Github to host our code, manage issues, and maintain an internal wiki, Google drive to share documentation and IRC and later Slack for real-time communication.

In chapters [3.1 Knownodes Technology and data structure description](#) and [3.2 Rhizi Technology and data structure description](#) the specifications of Rhizi and Knownodes will be shared in greater detail.

Pilot projects and results

In the next chapters I describe the Knownodes and Rhizi tools that were developed during my thesis. These descriptions might give the reader a general idea about their features for enabling non-experts to explore and collaborate around building knowledge graphs. These tools are ultimately made for people to use to support activities that fulfill specific needs. During the course of this thesis, I tested these tools within several organisation in order to answer the questions we have set for ourselves.

Can these tools help people become more aware of the human capital around them? Can they encourage collaboration? What problem did people encounter using them? In which use-cases are these tools appropriate and which are better dealt with other tools?

In order to assess this, we had launched several pilot projects during the thesis period, as you can see arranged as a timeline in figure 38. Blue indicates project we tested with Knownodes. After determining to pivot away from Knownodes, there was an ideation phase, in which we explored different opportunities during workshops, meetings, hackathons and collaborative grant-writing. These are reported in chapter [3.7 Other pilot projects](#). After the ideation phase, we settled on several stakeholders to run the pilots colored orange.

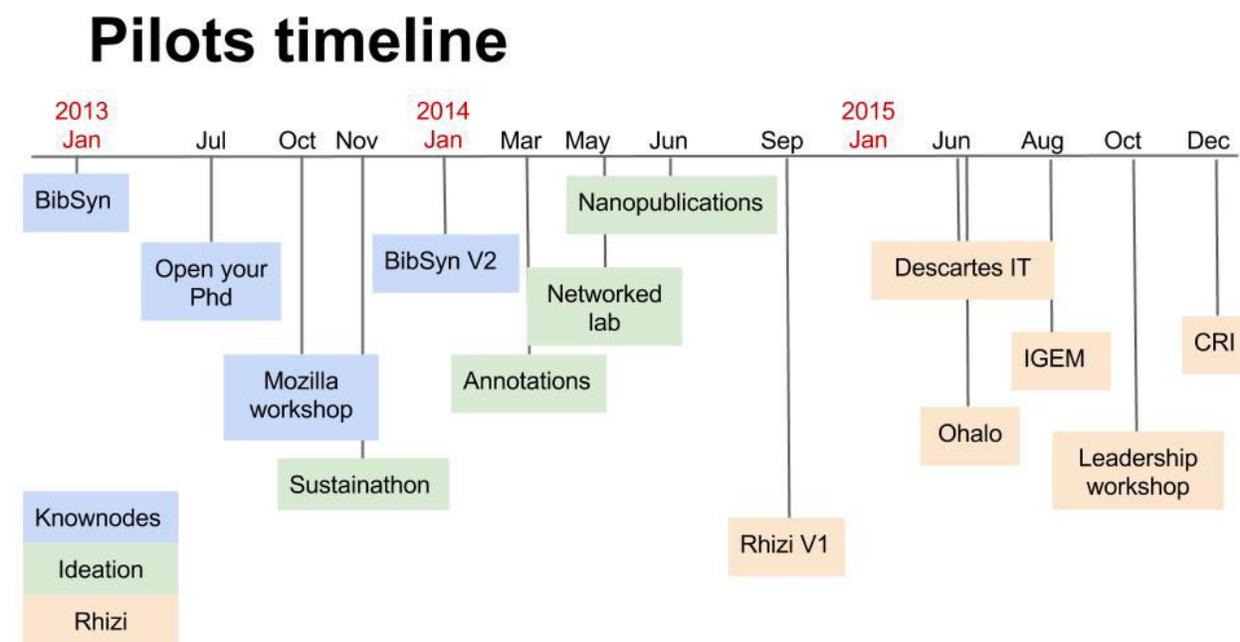


Figure 38: Timeline of the different pilots during the course of the thesis.

Of these pilots, three were mature enough to justify in-depth interviews (Descartes IT, CRI, and BibSyn) and the Descartes IT use-case achieved sufficient levels of adoption to justify extended research and analysis; a manuscript depicting Rhizi and its implementation in this pilot was submitted to Human Computation Journal (see appendix 3 [Manuscript submitted to Human Computation Journal](#)).

Several pilots did not reach sufficient maturity because of software issues, implementation issues or simply lack of time because the thesis was rolling to an end. Specifically, I think that the CRI use-case deserves further work, and I hope someone will be able to pick up where I left off.

Within each part of this chapter I describe the background about each pilot participants. What specific questions we were trying to address. The concrete objectives that were set. Outcomes, and the conclusions from the pilot.

Knownodes technology and data structure description

Innovative thinking, collaboration, participation and inquiry-based learning in the classroom is sought to be the cornerstones of 21st century learning skills. My objective was to design an open online collaborative platform to facilitate and stimulate these learning modes.

To this end I focused on a central theme essential for sustainable learning: Nurturing the ability to curate, filter and successfully communicate information with peers about what you and they learn and the ability to associate abstract concepts and learning material with real-life questions and challenges.

The Knownodes web platform was based at its core on the premise that the connections ('edges') that structure the network between knowledge resources and open questions are centers for social interactions. These *information-rich edges* generate a network, effectively mapping the state of the art and the interconnectedness of knowledge domains. It's primary use is to create an organized and scalable system, designed to serve students and researchers to actively create, share and search for qualitative connections between questions and knowledge resources.

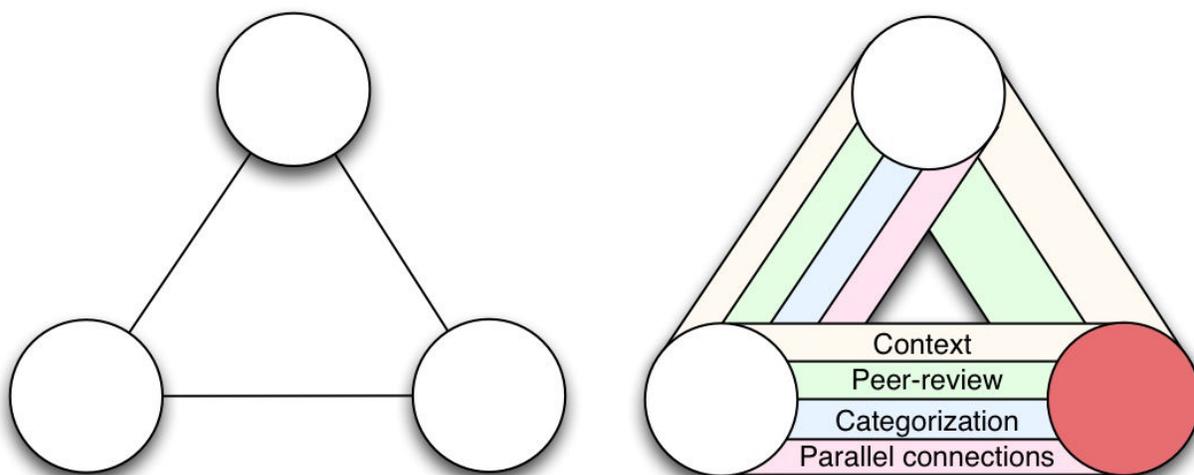


Figure 39: Knownodes features information-rich connections between entities.

The information rich connection is in the heart of the system, it connects between two resources to form a simple and readable sentence. Each time an information rich connection is created, it also functions as a center for interaction: generating discussion, peer-review and categorisation activities around it.

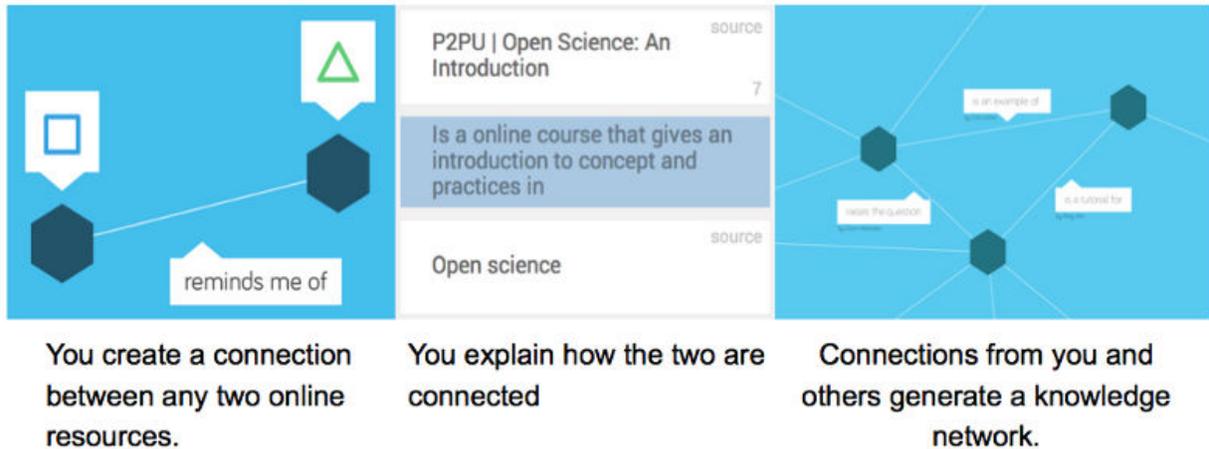
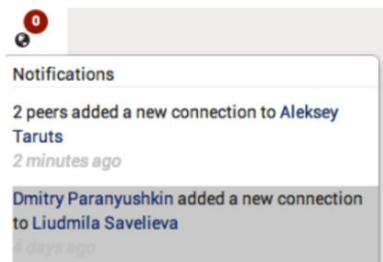


Figure 40: How To use Knownodes

The system offers students a way to explore the initial knowledge base both as highly readable structured text as well as a visual map representation. It also offers them a way to enhance the knowledge graph by creating their own knownodes, connecting multimedia resources such as videos, images and hyperlinks as well as providing a discussion platform and peer-reviewing tools.

Features for self-organisation



Notification system



Voting system



Reputation system



Commenting



Knowledge stream

Figure 41: Advanced social features implemented in Knownodes V2.

These connections are discoverable to the teacher and fellow students, but most importantly they are also available to students from future years classes, enabling them to build upon the knowledge of the

past learners to learn and add new knowledge more efficiently . By providing feedback on the younger students work, they stand to benefit from learning by teaching, save the teacher time and create a virtuous workflow where students constantly teach and learn.

Knownodes wiki: <https://github.com/CyberCRI/KnowNodes/wiki>

Knownodes repo: <https://github.com/CyberCRI/KnowNodes>

Knownodes demo video: https://www.youtube.com/watch?v=LBt5vsHh_h0

Technology stack

Neo4j for the database, node.js+express.js for the server. Angular.js for the front end.

Data model

Resources these are the entities connected by connections, they either point to a url, a file or a user-generated resource. User-generated resource is a type of resource that does not refer to any existing URL but is original content by a user, for example an open question or an idea are user-generated resources.

Concepts a type of resource that is always user generated and unlike other types of resource does not need to be connected to anything to exist.

Connections are the entities that connect two resources together. They contain a description of how the two entities are connected, have a Connection type and can be commented and voted on.

Parallel connection - When there is more than one connection between two resources we call the connection parallel connection.

Start-Resource and End-Resource - When a connection is directional the Start-Resource is displayed at the start and the End-Resource is displayed in the end.

Connection type are sort of tags that are associated to connections and allow the sorting of the connections. Currently there are three basic types: Explaining, Inspiring and Criticizing. Though soon many new subtypes will be prototyped.

Triplet A combination of two resources and the connections between them. This is the usual way information is organised in the system.

Database

In the Knownodes main database, Neo4j, data is organised like a graph composed of two types of units: nodes and edges.

nodes contain information about users, group types, resources, connections, and comments.

edges contain information about how each node is connected to another node, the direction of the connection, and who voted on what.

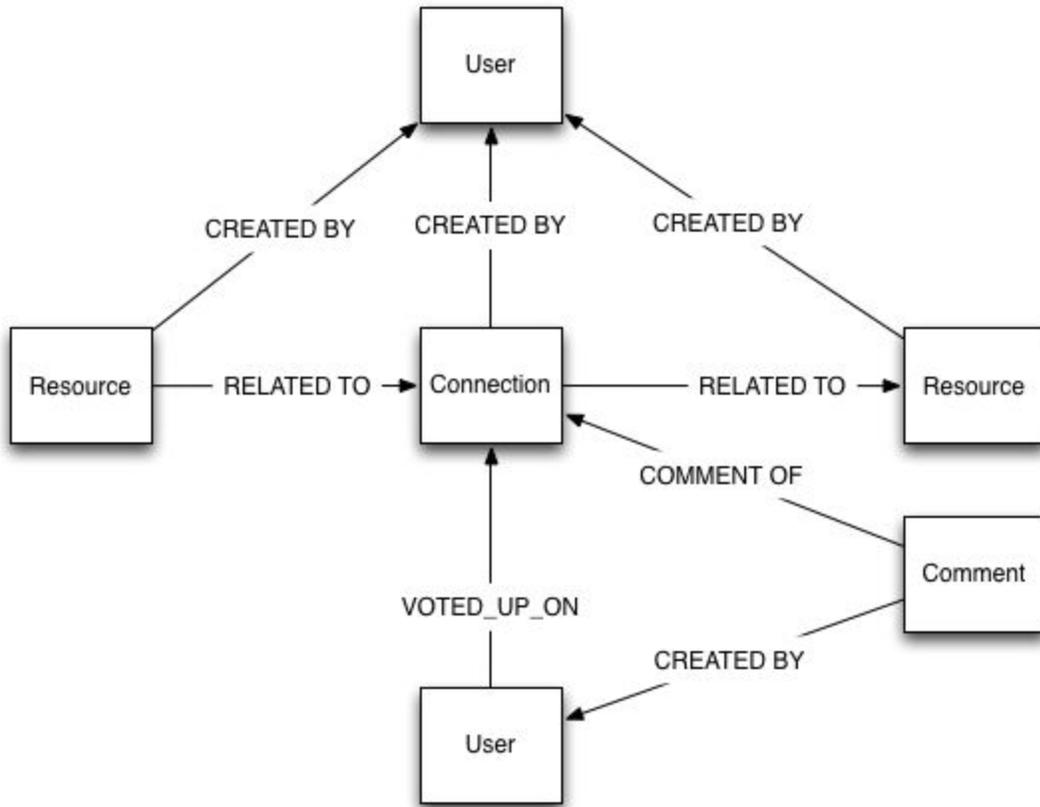


Figure 42: Knownodes data structure

For example, figure x describes a minimal triplet. It is composed of 6 nodes and 8 edges : A node for the user who created the triplet, two nodes for the resources, a node for the connection between them, a node for a comment, and a node for the user who created the comment. Two edges of type [related to] connect between resource and the connection, four edges of type [created by] connect the two resources, connection and comment to user nodes and a [comment of] edge connecting the comment to the connection node, and a [voted up on] edge connecting a user to a connection node he has voted on.

In the database there are several types of nodes depicted by the property name KN_type:

- Kn_userGroup = for defining groups of users in Knownodes.
- Kn_User
- Kn_Tag
- Kn_KnowledgeDomain - to be deprecated.

- `Kn_ConnectionType` - To group together types of connections (To be deprecated?).
- `Kn_Post` - for resources - `KN_Post` gets new properties if it refers to a file, a url, or is a user-generated resource.
- `Kn_Edge` - for connections.
- `Kn_Comment`

API modules

Relation module - responsible for creation of edges.

User module - creating, retrieving and deleting user nodes.

Comment module - creating, retrieving and deleting comment nodes, calls edge module for edge creation.

Concept module - handles the special type of resource called concept.

Edge module - responsible for connection nodes.

Knownodes module - responsible for Resources. For retrieving triplets and for the complicated operation of connecting between users, resources and connections to form a triplet.

Back-end

As a whole, the application is organized following a [Three-tier architecture](#) :

- Data tier Neo4j Graph Database hosted on Heroku
- Logic tier Node.js application hosted on appfog
- Presentation tier AngularJS single-page application executed in the browser

Communication between Logic and Data layers

Neo4j provides a RESTful API and allows our Logic layer to manipulate the data. Many requests are written using the Cypher language, which a graph-oriented SQL-like language specific to Neo4j.

Only our Logic layer is allowed to communicate with the Data layer.

Communication between Presentation and Logic layers

The Logic layer in turn provides a REST API with the services required by the Presentation layer.

The goal of this page is to describe how those web services are implemented and organized.

Service Endpoints

The endpoints are implemented using a stack of third-party modules, including :

- `express` : HTTP requests handling
- `express-resource` : Automatic binding of CRUD-like functions
- `passport` : Authentication

In the project, the service endpoints are declared in [app.js](#).

The implementations can be found in the [controllers folder](#). The most important controllers are :

- resources
 - connections
 - users
-

WARNING !!!

For historical reasons, obsolete controllers are still present in the source code. During a massive redesign, most controllers were re-written, however there are still some chunks of logic that need to be moved. In a near future, our goal is to completely get rid of the following controllers :

- concepts (replaced by resources)
 - edges (replaced by connections)
 - knownnodes (should be cut down in smaller chunks)
-

In a controller's folder, you will find two files :

- index.js
- Defines the functions that will actually be made available to the world. Completely delegates to the other file
- (Name)Controller._coffee
- [Streamlined](#) CoffeeScript implementations. Those components are responsible for getting all the data from the HTTP requests and delegating to DAO or model components.

Controllers and DAOs already implement basic CRUD functions. If you want to make those functions available to the world, all you have to do is define them in the index.js files.

Rhizi technology and data structure description

The features described below are divided into two parts: Front-end focusing on interaction design and back-end focusing on data structure and storage. When necessary, I added additional information about the implementation.

For a more dynamic overview of Rhizi's interface, here is a link of video tutorial that was made for the CRI students <https://www.youtube.com/watch?v=n4pMaG24q78>

The project is released under an AGPL open source licence and can be viewed on: <https://github.com/Rhizi/rhizi>

Notable Front-end features

For a complete list of features see Appendix 2 [Rhizi manual](#). For detailed review of specific usage see segment [3.5 Descartes IT](#).

Text-to-network Input

Input in Rhizi utilizes a novel text-to-network real-time visualization. As the user types, the preview of the output is displayed in real time. If the system recognizes that an entity already exists user is shown it's current location within the map view (see figure 43). To differentiate between nodes and links users use double-space in their assertions.

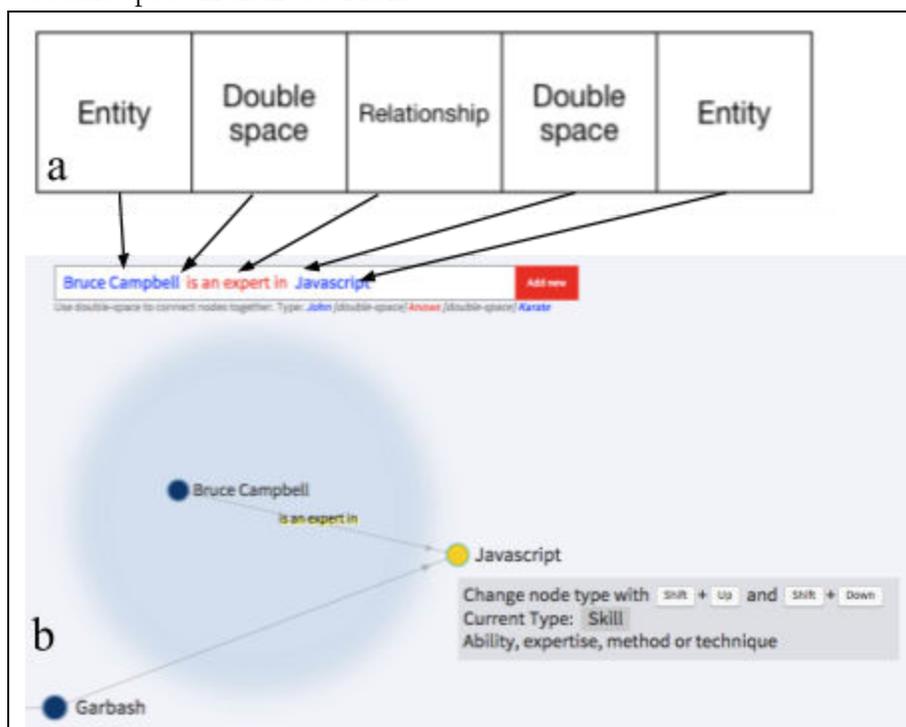


figure 43: (a) Syntax of a simple phrase to input in the graph. (b) Syntax and text-to-network preview in interface.

Edit operations on selected nodes and links

Once a selection has been made, users can select to delete, merge and connect nodes together. Merge allows the user to merge entities that in fact represent the same thing. This tells the system to remove one node, while moving its links to the node it is merged into.

Shortest path

Once two or more nodes have been selected, clicking on the Shortest path button will highlight the minimal distance between the original node and the nodes selected.

An example of the usefulness of this function is when a user wants to know how he can reach a specific expert and the system can highlight specific people that can introduce him to the expert or shared skills and expertise they might have in common.

Zen mode

Zen-mode is a dynamic way to explore the map. It filters out all nodes and edges that are not highlighted or selected. By selecting and deselecting nodes, the user can travel along the map.

Node properties

Each node type contains properties as defined in the data schema. These properties are editable. There are several special properties that have specific effects: URL property adds a hyperlink icon next to node in the map interface that links directly to the Url's target. Image-url property will display a thumbnail of the linked image.

Filtering

Nodes can be filtered via a dropdown menu according to type.

Navigation between maps

Users can navigate or create maps by using the "New Map" or "Open Map" buttons. Each map contains a subset of the nodes and links in the database and represents a separate mapping task.

Layout

Map layout can be modified by changing the layout. The three options are: (1) Force layout (Fruchterman & Reingold, 1991) (2) Ring layout - arranging the graph according to concentric rings according to node type (3) Custom layout - which is a manual and collaborative arbitrary placement of nodes according to users (See figure 2). This allows users to get different perspectives on the data.

In our development process we started with the Force layout because it optimized ease of viewing as it minimized connections crossing each other, making the network less readable. Following user-feedback, we saw that when maps grow beyond ~80 nodes, users started having problem reading the map, and the connections were hard to read as well. The ring layout enabled grouping of the different node types in circles, and an alphabetical order, making information easier to look at

without losing the network view. Finally in order to further explore additional layouts, and curious to see how the “wisdom of the crowd” will reposition nodes, we developed the custom view.

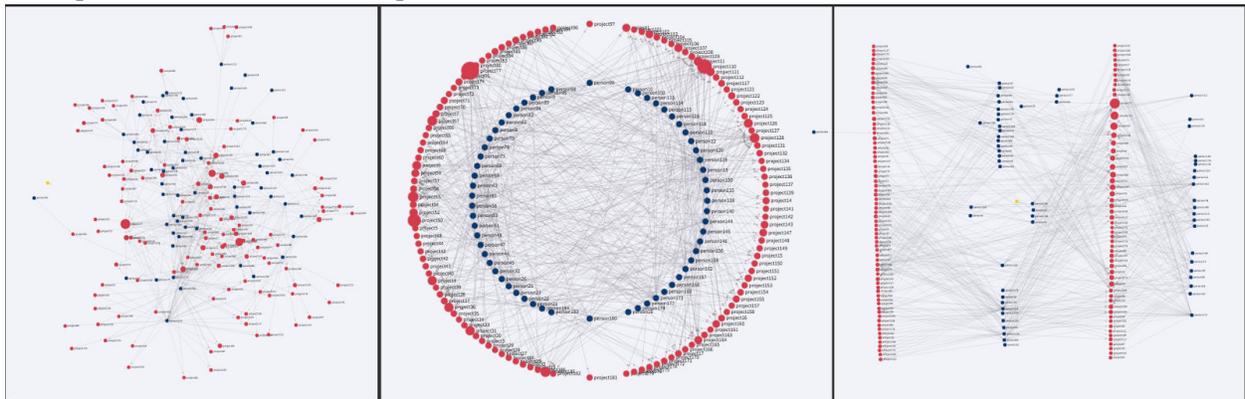


Figure 44: The different visualization layouts in Rhizi: Force layout, ring layout and custom layout.

Free assigning of connection types

From previous experience with earlier prototypes, we didn't want to burden users with selecting from an ontology of existing relationship types. We also didn't want to disrupt the flow of using our text-to-network interface and so we decided to allow users to express relationship types via freeform text.

Back-end

Graph database backend

Data inside the system is stored in the graph database Neo4j. Making it easier to use graph based analysis, algorithm and applications.

Data model

Data is modelled according to the schema in figure c. Inside the Database, we have two separate graphs: A graph depicting the relationships between the entities, which have a 1-to-1 relationship with the representation shown through the UI. A second graph which represents the different maps (referred here as Rz-docs) and which specific nodes and links appear in which map and which are shared between maps.

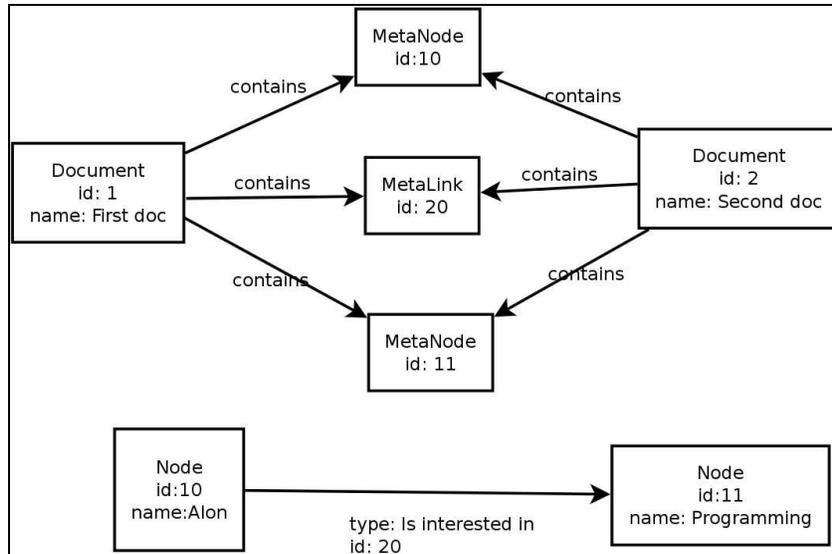


Figure 45: Rhizi data model

Real-time collaboration

Using Websockets and TCP connection, changes in maps are updated in client in real-time.

Import from Spreadsheet

An additional repository(<https://github.com/Rhizi/rhizi-api-client>) contains a configurable tool that imports information from a spreadsheet into Rhizi. A rules file defines which specific column get translated into an edge, node or a node property.

Example of a simple rule file in which the letter denote a specific column in a spreadsheet:

- node:

- type: person
- name: B + A
- work address: X + Y + Z + AA + AB
- subtype-tags: Phd student 2015

- node:

- type: project
- name: C
- description: D
- subtype-tags: Phd project

- edges:

- '[A+B] worked on [C]'

Configurable Data Schema

Within each pilot variation between data models were used. A configuration file can be modified to specify the node types and their properties. However, to illustrate the general structure, I am using a detailed explanation of a specific use-case - Within the Descartes University IT department.

The Rhizi data schema was adjusted for the IT department according to the following table:

Node types	Node properties
Person	name, email, image, URL, description
Project	Name, Description, URL, Status, Start date, End date, country, city, address, facility
Skill	Name, Description
Keyword	Name, Description, URL
Organisation	Name, Description, URL
Media	Name, Description, URL

Figure 66: Node types and their properties as defined for the IT department use-case

These types were selected following extensive discussions with the heads of the IT department, following a hybrid of human-centred and activity-centred approaches to offer a good balance between structure while not hindering the user experience.

Customizing system to different use-cases

As we were working with use-cases that used three different languages(French, Hebrew, English) and different data-models, we had to build customisation tool so each use-case can adapt itself. We also had to adjust URL's, whitelists of users, logo's, tutorials and documentation.

Logging of user actions

Using the Piwik web analytics software, we logged users visits to the different sites including which maps they have accessed. Within our database we logged all content editing actions. In case you are interested to log a single user action within a map, you can use the advanced settings, which is accessible by adding “?debug=1” at the end of Rhizi’s URL and you can record and replay sessions.

Knownodes Bibsyn Pilot

Background

To test the knownodes platform, we deployed a Knownodes instance for the usage of a second year Master's course within the CRI called Bibsyn (Short for bibliographical synthesis). About 20 students participate in this course each year. From the course page: "The course consists of a series of seminars, each initiated by two students from different backgrounds, presenting their review of scientific literature concerning a subject of their choice.



Figure 48: Students using Knownodes during a Bibsyn session

All students of the course are expected to contribute for each session their point of view on the subject of choice by: 1. reading (at least one) peer-reviewed paper or well-documented website. 2. posting it in the database together with writing a 200 word paragraph describing the resource and its relevance to the review subject."

In 2013 we have deployed a first version of Knownodes to be used within the course. Due to numerous bugs and systems crashes, we were unable to truly test the usage within the classroom. In 2014, we did a second iteration, this time with a more robust system.

Challenges

The main innovation of Knownodes was the ability to categorize and add content to connections between knowledge. The main two questions we wanted to answer were:

- 1) How would people use these feature.
- 2) Would these information-rich connection help people better explore the knowledge base that was created.

Methods

During the course, students were divided into groups of two and requested to present a scientific concept of their choice, submit the concept to Knownodes and have other students connect relevant publications and concepts to it.

The screenshot shows the Knownodes web interface. At the top, there is a search bar with the text "Find or create a resource...", a "Map mode" button, and a "Login/Register" link. Below the search bar, there are sorting options: "sort: Popular New". The main content area displays a grid of concept cards. Each card has a title, a "source" link, a count, and a "comment" link. The cards are arranged in two columns. The left column shows connections for "Translation (biology)" with a count of 4, and "Cell (biology)" with a count of 1. The right column shows connections for "Gene" with a count of 2, "Translation (biology)" with a count of 3, and "Messenger RNA" with a count of 3. Each card also has a "+" and "-" button next to the count.

Concept	Count	Author
Protein	2	
is a final product of	4	
Translation (biology)	2	10 days ago by Urszula Czerwińska
Gene	2	
holds hereditary information that the cell is able to use to make proteins, a process called	3	
Translation (biology)	3	10 days ago by Yara A
Translation (biology)	1	
occurs in	3	
Cell (biology)	1	10 days ago by henry de belly
Messenger RNA	1	
contains the instructions (code) for	3	
Translation (biology)	3	10 days ago by Antonio VL

Figure 49: Knownodes page displays all connections related to “Translation”, a concept in biology from the course.

We conducted several in-class feedback sessions, a tutorial and conversations with small groups of students. At the end of the pilot we sent an online survey for students to fill out.

Results

We quickly discovered that the first challenge we have set for ourselves presented a very big challenge. The ability to get a group of students to use the system was difficult because of three types of issues that rose up during the pilot:

Input interface issues

Because students were used to conventional interfaces to input information, we needed to put a lot of work into making the input interface usable. This involved the following challenges:

How to allow people to input “triplets” (two nodes and a connection between them), without feeling they are filling a very long form?

The process of inputting a triplet consisted of the following steps: Selecting the name of the first resource. Selecting the type of first resource (Link, post or wikipedia link), Selecting the category of the resource (such as article, concept, question, idea etc’), Selecting the description of the connections, selection the category of the connection. And then going through the same procedure for the second resource as for the first one, and then, selecting the directionality of the connection.

When we tried a naive implementation of these specification as a step-by-step form, it was extremely difficult for users to handle (See figure 50 and 51).

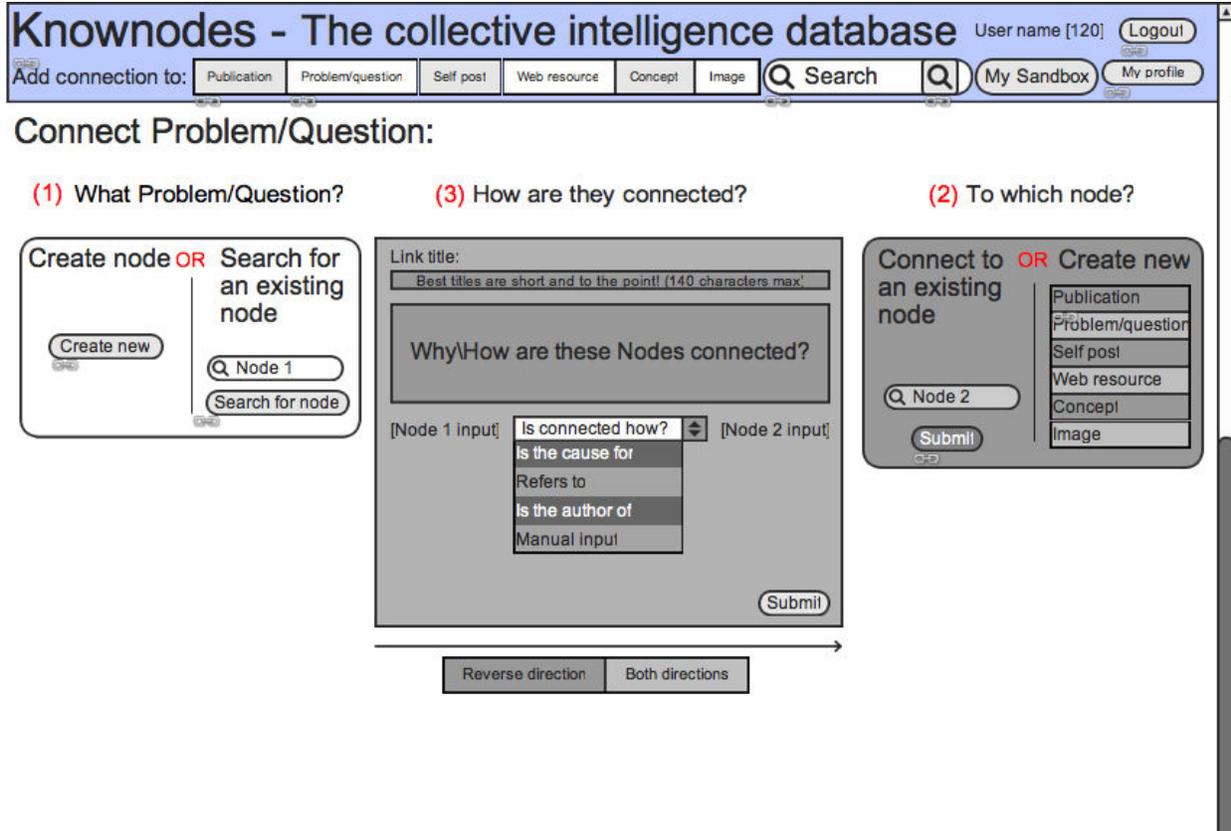


Figure 50: Mockup Knownodes step by step form

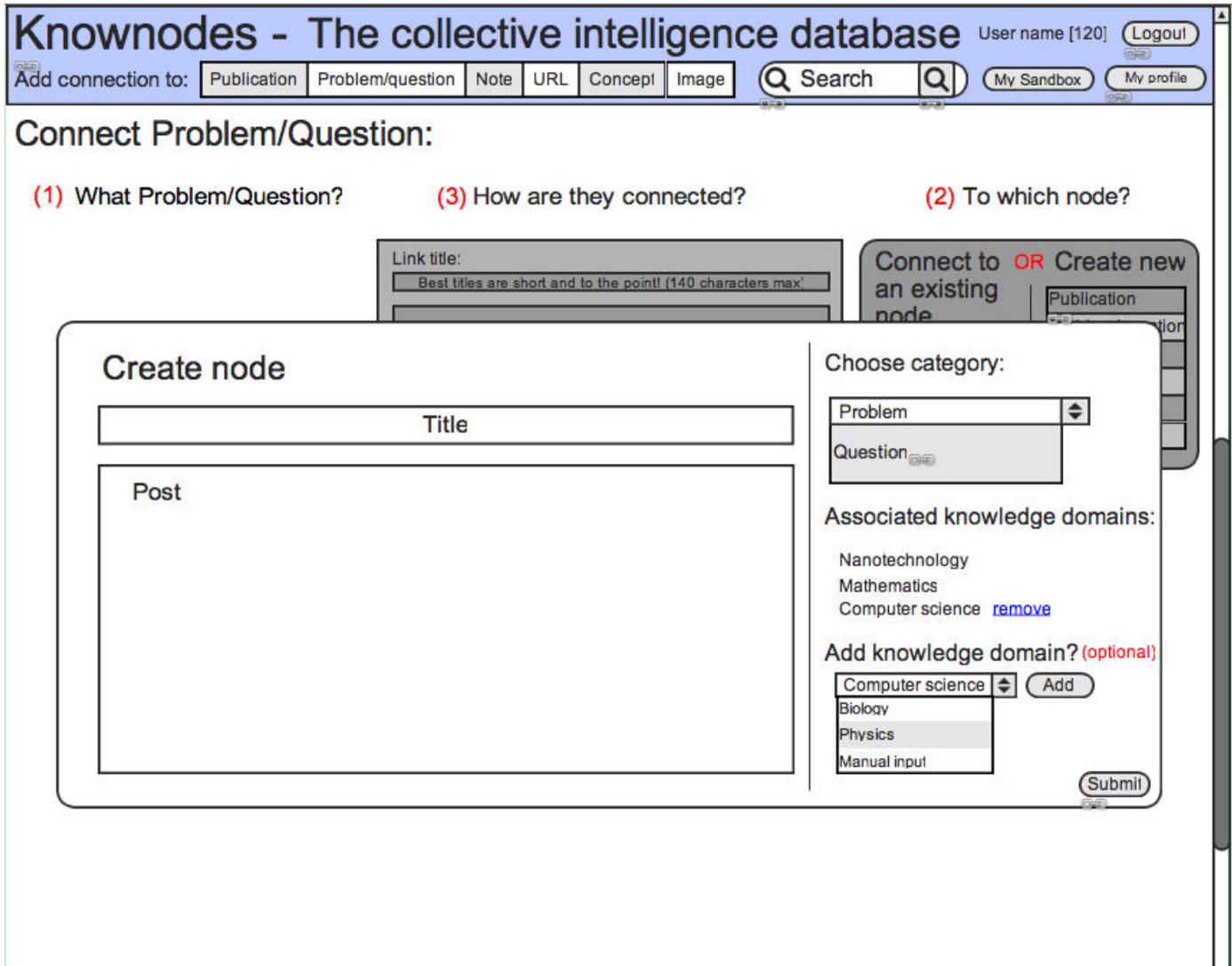


Figure 51: Mockup Create new Problem/Question dialogue box

We needed to add many different features such as auto-import from wikipedia, auto-detecting url's and scraping websites, changing the directionality of the connection between the concepts with a button, unifying the display of both nodes and connection in the same view and many other small adjustments to overcome these problems. You can see the work on simplifying the interface on figures 52 and 53.

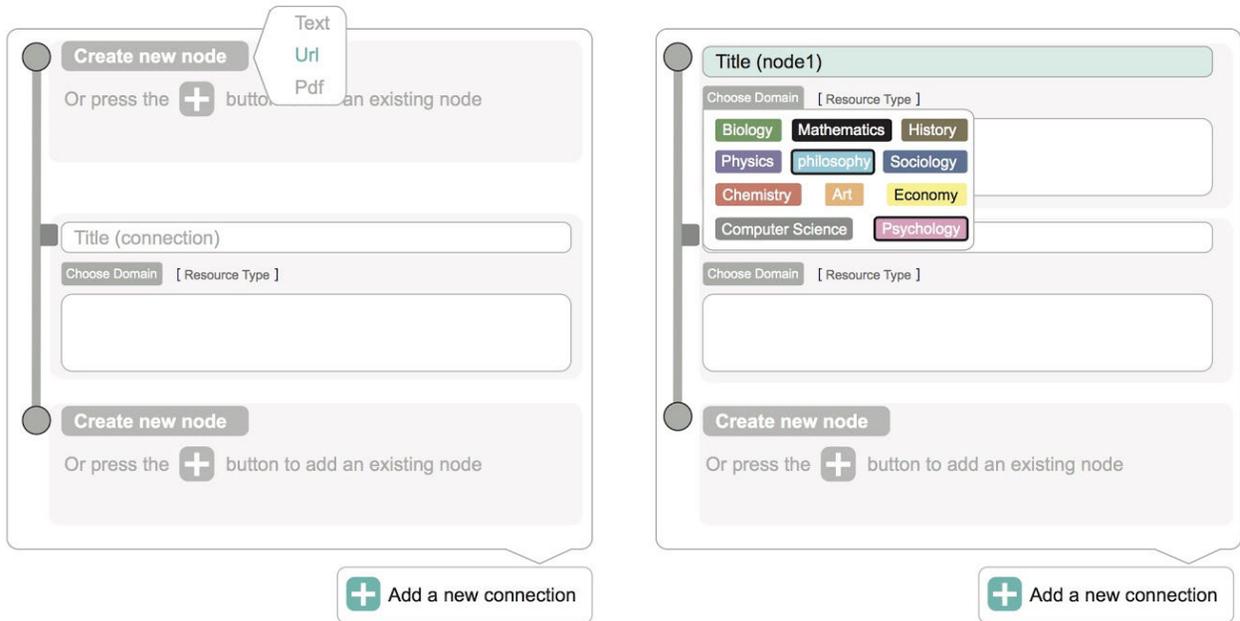


Figure 52: Mockup following first iteration

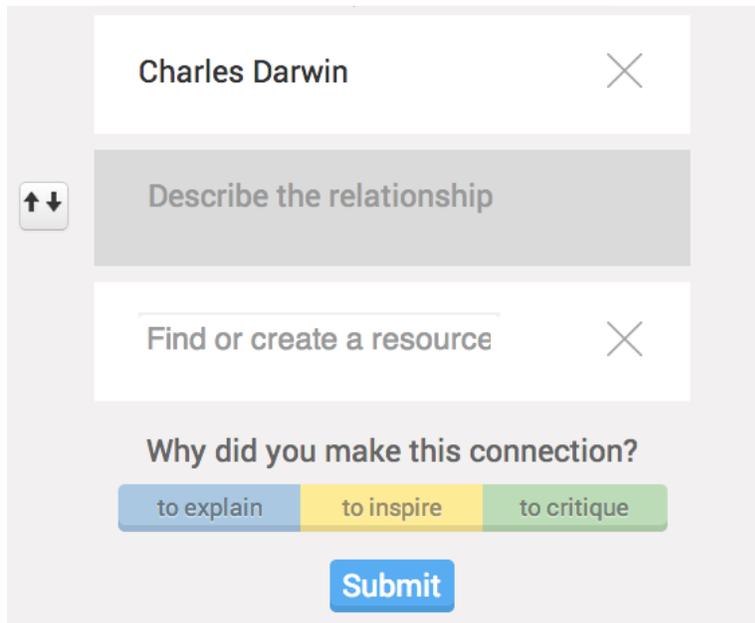


Figure 53: Final input form that was implemented

How to enable people to place the right category on the connection when, if there are too many options in the interface the user just gives up, or puts the wrong type of category?

Categorizing connection was a surprisingly challenging task for the students. We had to quickly evolve from providing a closed list of potential relationship types into allowing people to select between three categories that do not describe the nature of the connection, but the intention of the

creator of the connection. In other words, why did he create the connection in the first place(see figure 53).

Communication of interface components and value of software

How do you explain to students the concept of nodes, connections, knowledge graphs and other unintuitive concepts related to the software? How to articulate the value that this work provides?

While usage of the software was mandatory for certain parts of the exercise, it was quickly apparent that most students did the absolute minimum(or below that) when interacting with it. The main value proposition was that building these kind of knowledge networks can help future students and peers to better understand the concepts the student was working on, but ultimately there was no “buy-in” from the students. The value was not inherent in the action itself, as student performance did not affect their grades.

As specified in the discussion section below, we did learn a lot about the types of challenges associated with getting students to use the interface, which was productive towards future iterations. Unfortunately, the data students put in didn't have sufficient quality to justify further investigation regarding the viability of information rich edges to those who read and explored the graph.

The database

A total of 10 concepts were uploaded to Knownodes, to each concept had an average of 14.4 connected resources and an additional 1.4 resources that have been connected as a second degree connection(connected to a publication that has been connected to a concept).

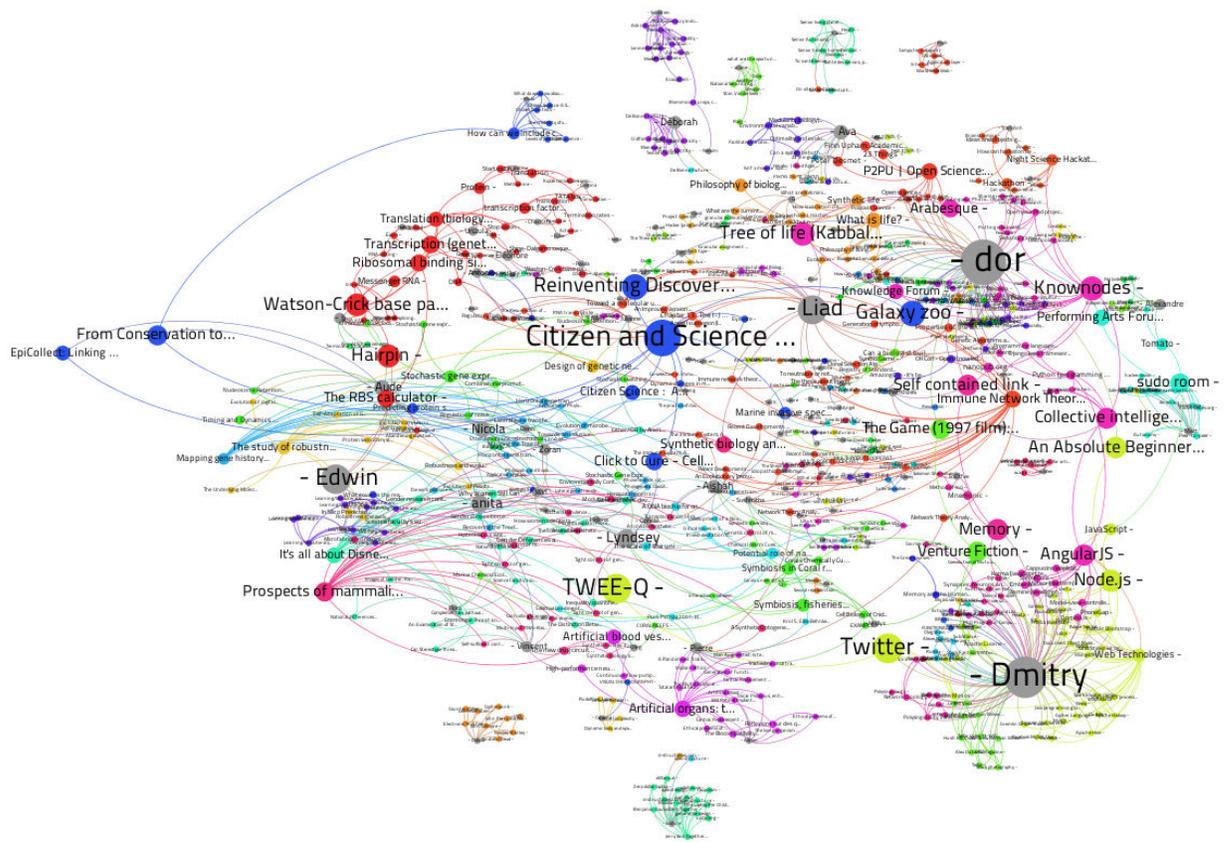


Figure 54: Visualization of Knownodes database including rapid prototypes and workshops.

Survey

Survey results show that out of the seven students who filled the survey only two would recommend it to friends, and the platform itself was rated on average 5.4 out of a scale of 10. We have also received a lot of negative feedback regarding technical problems and bugs in the system, information overload, and a feeling of “easily getting lost”. The positive feedback received was around appreciation of the concept, its potential and value.

Result analysis

Following conversation with students, teachers and analysis of the survey, we attribute the negative feedback to the following reasons:

- Bugs and crashes at early deployment diminished trust in the the system.
- Because the interface allowed to basically connect anything to anything people were confused about what action they were expected to do.

- No feedback about contributions made from peers and teachers reduced engagement.
- There was no aggregation or updating module to let participants know what is going on throughout the system.
- Poor adjustment of the tool to the use-case - Contribution was perceived more as a formality than a something that stands on its own merit.
- Students were putting the data on the Moodle forum and only then on Knownodes. Contribution was perceived more as a formality than a something that stands on its own merit.
- Lack of visual feedback, left people confused about what they were actually doing.
- The site didn't leave an impression of having a distinct identity (It wasn't "cool")

All these, together with a functional-yet-non-attractive interface, complex form inputting and troubles in navigating the system, led to a user experience that was better than using the moodle forum, but not enough of an improvement to encourage continued usage and engagement.

The challenge of getting people to contribute to the system is much more apparent now than before. Early prototyping have surfaced problems concerning the best format to input information and what kind of benefits does the system provide to motivate users.

By reading and researching around domains such as semantic annotation games(<http://semanticgames.org/>), research on the use of topic and concept maps(Benteley et al. 2011) and critique on them(Ruiz-Primo & Shavelson 1996). in education and discussing with people who worked on similar projects, notably [thisIsLike](#) and <http://app.kleenk.com/> (both of them abandoned) have made me realise that motivating users to do such an unfamiliar action as annotating connections requires redesigning the system for solving a much more concrete problem as well as being more engaging.

A second challenge is to strengthen the sense of identity and coherence of the system: This would be provided by producing a map module to navigate the system, and some design "magic"(making it cool).

A third challenge would be to focus on real-time usage and collaboration on Knownodes. Meaning to give people the ability to input content and receive updates instantly from a mobile or browser extension.

Lastly I believe the system should focus on the creation and sharing of innovative ideas, generating questions and inspiring connections.

Discussion

Following the feedback received, we've revamped the visualization and collaboration system of the software as well as changed the flow of task generation and input to suit a realistic use-case.

The feedback from my Phd advisors(Ariel and Uri), from several labs,(Stanford, Pasteur) open source companies and organisations has convinced me that the software being developed should go a step beyond merely a proof of concept or an experimentation tool, but there is a real need for an open and intuitive tool for helping people collect and explore knowledge graphs.

To do this, we had to shift beyond working with temporary interns into a more stable organisational environment.

Shortly after applying the new features, we ran into another problem. The system, initially built by interns was too complex to further enhance and development became too resource intense for such a small-scale project. In addition for this we identified that capturing knowledge was too complex a problem to complete during the time of the thesis, and it is more efficient to collaborative map human-capital. Humans usually do a good job as being more up-to-date with the state-of-the-art, with changing opportunities and projects within the organisation. The cost and maintenance of mapping humans and keeping the information up-to-date is several levels of magnitude lesser than pure knowledge and less subject to interpretation and confusion from the part of organisation members. Lastly, people have a vested interest in mapping themselves and their project, while neutrally mapping knowledge and research questions is something we were not able to motivate people to do voluntarily.

We abandoned the project and started working on Rhizi, which implemented the lessons learned.

CRI master students - collective community documentation for internship opportunity mapping

Background

Internships serve a key role in preparing students for their future careers. This is because Internships help students learn and refine skills while experiencing real-life tasks and social interactions. Additionally, internships provide a unique experience that allows students to reflect about the choices they make about a particular career.

Despite the importance of internships, finding a good one can be a daunting task. This is especially true for those seeking a career in science. Based on my own experience as well as that of other students, the choice a student makes regarding where to apply to an internship is often based on incomplete information. Labs can submit their internship proposal to an institutional database, but the coverage of available opportunities is often lacking or fragmented. The most common way students find internships is through direct recommendation from a teacher or colleague, or by applying to a lab that the student is familiar with their work. Looking online for opportunities often points to influential labs, that are flooded with requests from students. On the other hand, there are many labs that do interesting work, with very engaged PI's, that get lost like a needle in the haystack. As one of the CRI students phrased in the interviews we did "When you search for thematics[online] it's hard to find real labs, because it's always crappy articles that are talking about science but don't know anything about it, it's really hard to find concrete things about science, to really find directly the people involved in it".

The main drawback of the existing methods is that the student is exposed to a very limited pool of opportunities. This leads to students taking decision that are not based on a full understanding of their options, and for less well-connected labs to miss out on talented interns.

It is at this point that enhancing organisational awareness can provide students with a more complete picture of the options available to them. Previous generations of students before them have a direct experience of past internships, but these experiences are not stored anywhere and for the most part, not accessible to current students. What if students could learn about the internship experience of others without needing prior knowledge about who specifically to contact? On a broader perspective - How would the dynamic between labs and interns change if there was a tighter feedback loop between prospective and past interns?

This kind of problem is interesting and could be generalized to solve problems faced by millions of students around the world. Also, following the first pilot with Knownodes at the CRI ([See chapter 3.1](#)) we realized the importance of providing students with tangible value that provides them with an intrinsic motivation for contributing to the system. We also wanted to see how the interface of Rhizi, that we knew was a lot more intuitive and reliable would fare as compared to the previous Knownodes software.

Objectives

The objective of this pilot was to provide students with a platform to recommend and find internships for others based on their experience. Following requests from the CRI staff, we added additional criteria such as interests and projects to encourage students to find each other not just for internships, but also to find each other for other contexts such as clubs or special projects.

We were curious to test the following things:

- 1) Would the graph visualization provoke more participation from students compares to the Bibsyn pilot the year before?
- 2) Would the students be able to explore the different options they had without having to start off with a particular query? Would exploring the graph provide additional value to a direct query or a list based interface?
- 3) Would students gain value from viewing graph properties such as shortest paths between them and the internship or the skill they wished to acquire?
- 4) Would the data generated be valuable for future students?

Because the pilot relies on data that the students shared on 2015/2016 but has value to the students of 2016/2017 we cannot yet address questions 2,3,4 as yet. What we could test for is students engagement with the pilot, get feedback from students about their experience, and demonstrate the use-case as a proof of concept for data collection.

Methodology

Usage

To launch the use case we first defined a data schema for the nodes and their properties. For node types we choose: People, project, skill, keyword, organisation and media. Media was selected as a catch-all node type when people wanted to link to a web resource that is not directly any of the other node types.

We did a small prototyping run at the end of 2014/2015 with 1st year master (15 students) and 2nd year master students (35 students). In 2015/2016 we then asked the next generation(1st and 2nd

year) of master students to participate. This was done through email, forum messages and an in-classroom presentation. In addition, we created a [video tutorial](https://www.youtube.com/watch?v=n4pMaG24q78) (https://www.youtube.com/watch?v=n4pMaG24q78) specific to their use-case that explains the usage and rationale and created a custom landing page for students to access Rhizi (see figure 55):

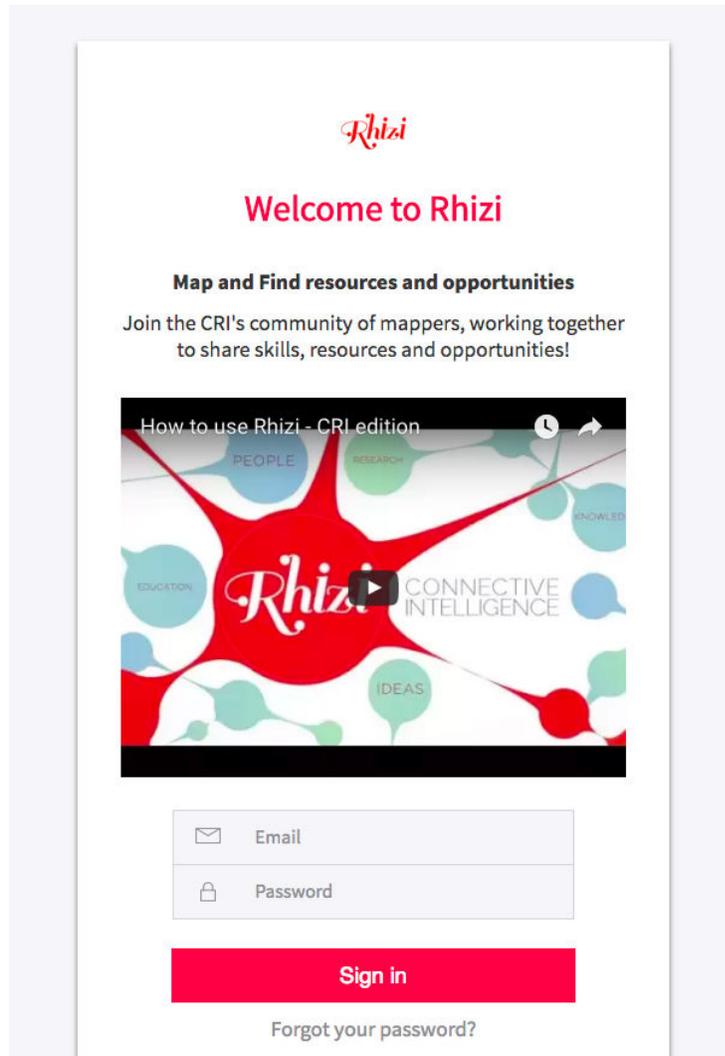


Figure 55: Rhizi CRI landing page

Students were requested to add themselves, three of their interests, three skills they have, an internship they did from the last semester, their internship supervisor, the organisation they did their internship in and up to three skills they have acquired from their internship. You can read the instruction in detail in appendix 1 - [Instructions for AIV students](#).

Other classes within the CRI also participated, notable the 1st year master program. The EdTech masters program and the 1st year Bachelor's degree students used it, but were not closely followed. A total of 6 maps were created as part of this pilot.

Feedback

While the main questions are related to students using Rhizi in 2016/2017, we did extended video interviews with 10 2015/2016 CRI students using semi-structured interviews.

Results

Resulting maps

Upon request, I can provide login details for those who want to explore the map here:

<http://cri.rhizi.net/index>

Below are a few examples of the maps created and ways people can use them to learn about internship opportunities, who to consult with to acquire new skills and learn about other student's projects and interests. We will focus on the map where most of our attention went - the M2 2015 map (figure 56).

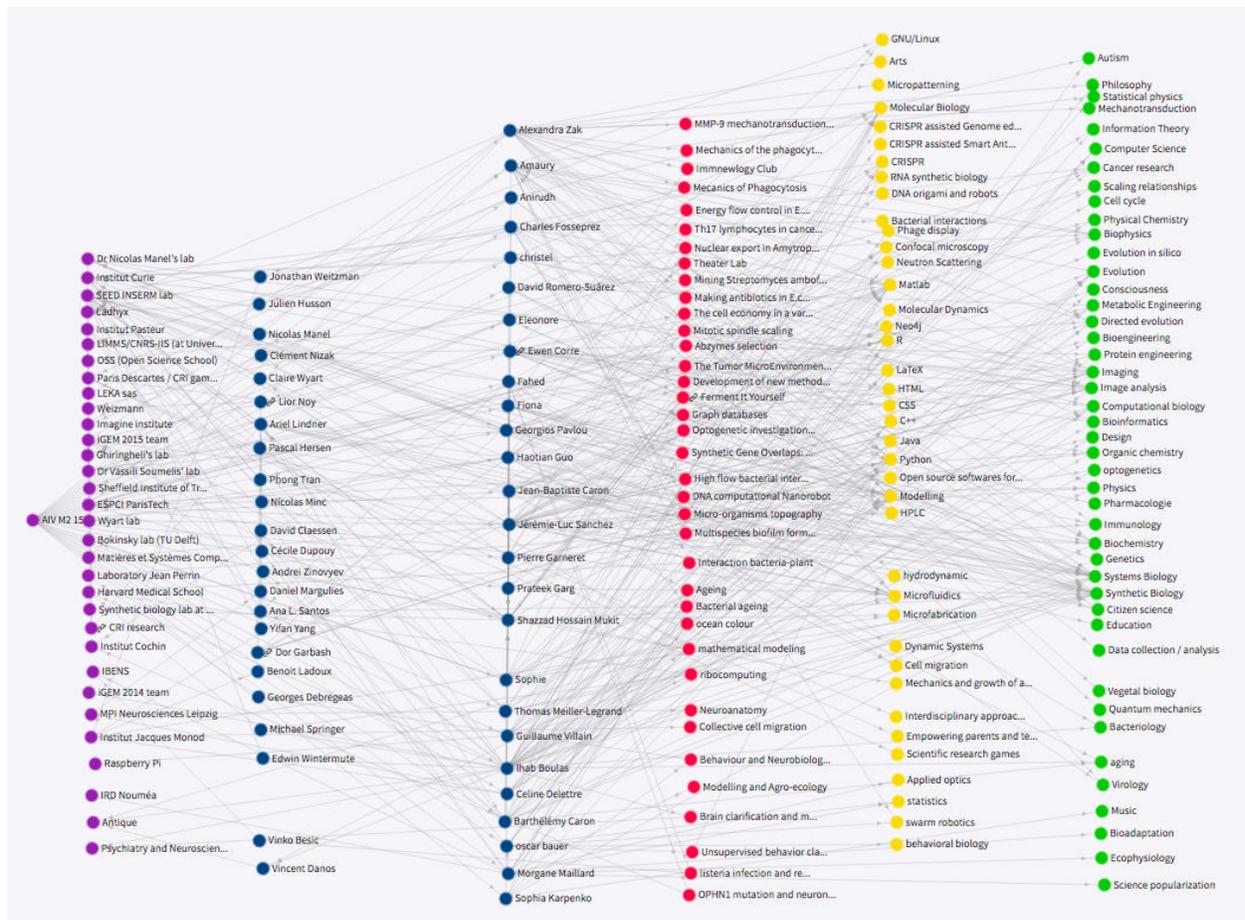


Figure 56: CRI M2 2015 map. View is arranged in custom layout as columns from left to right: organisation (purple), non students people (blue), students (blue), internships (red), skills (yellow) and keyword (green).

As a first use-case many of the students were curious about potential collaborators for projects or shared work in the class, but also just as a way to socialize with others. Below is an example of a student centric map (Figure 57):

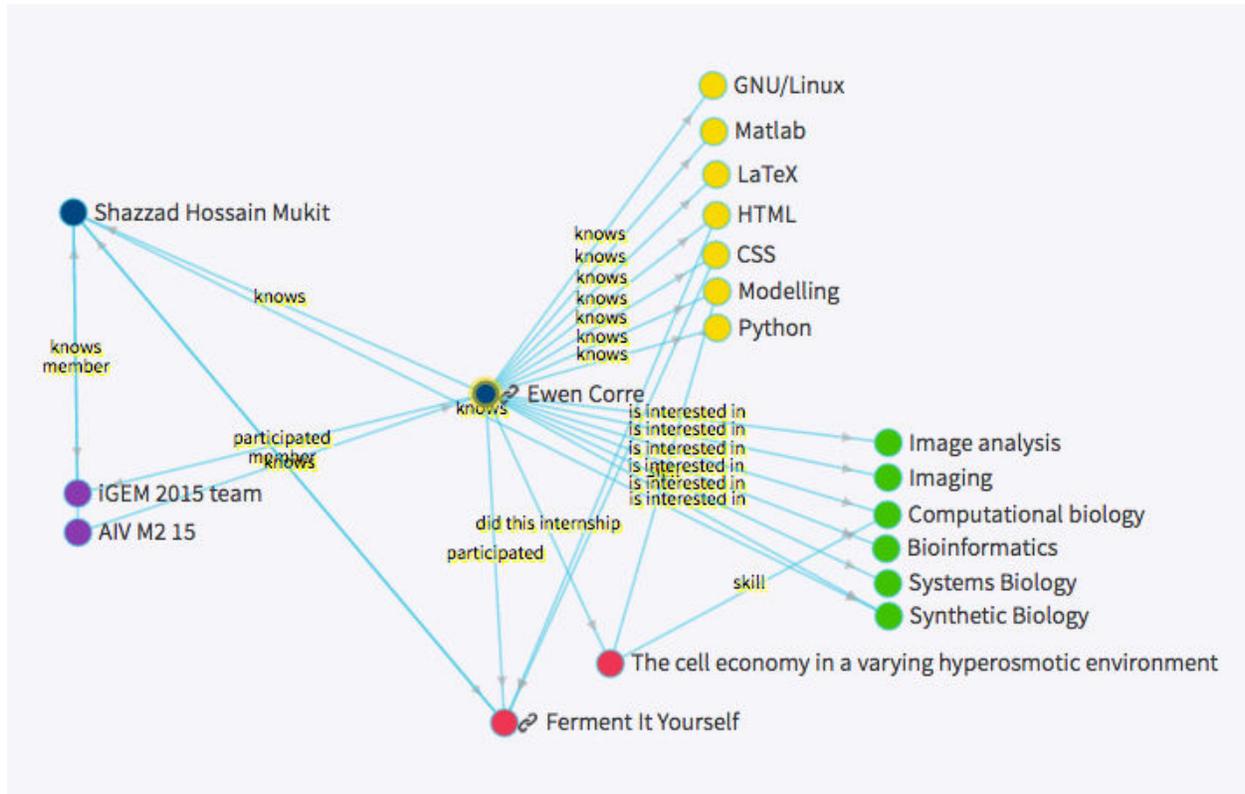


Figure 57: Ewen Corre M2 2015 Ego network.

From the map you can see technologies he mastered, scientific domains he is interested in, projects and organisation he is a part of, internships he did and students he's closely related to.

Any students at the CRI can open this map and click on a skill he wish to acquire. For example a students can click on “Modelling”, switch to “Zen-mode” to filters out other information and get the following result (Figure 58):

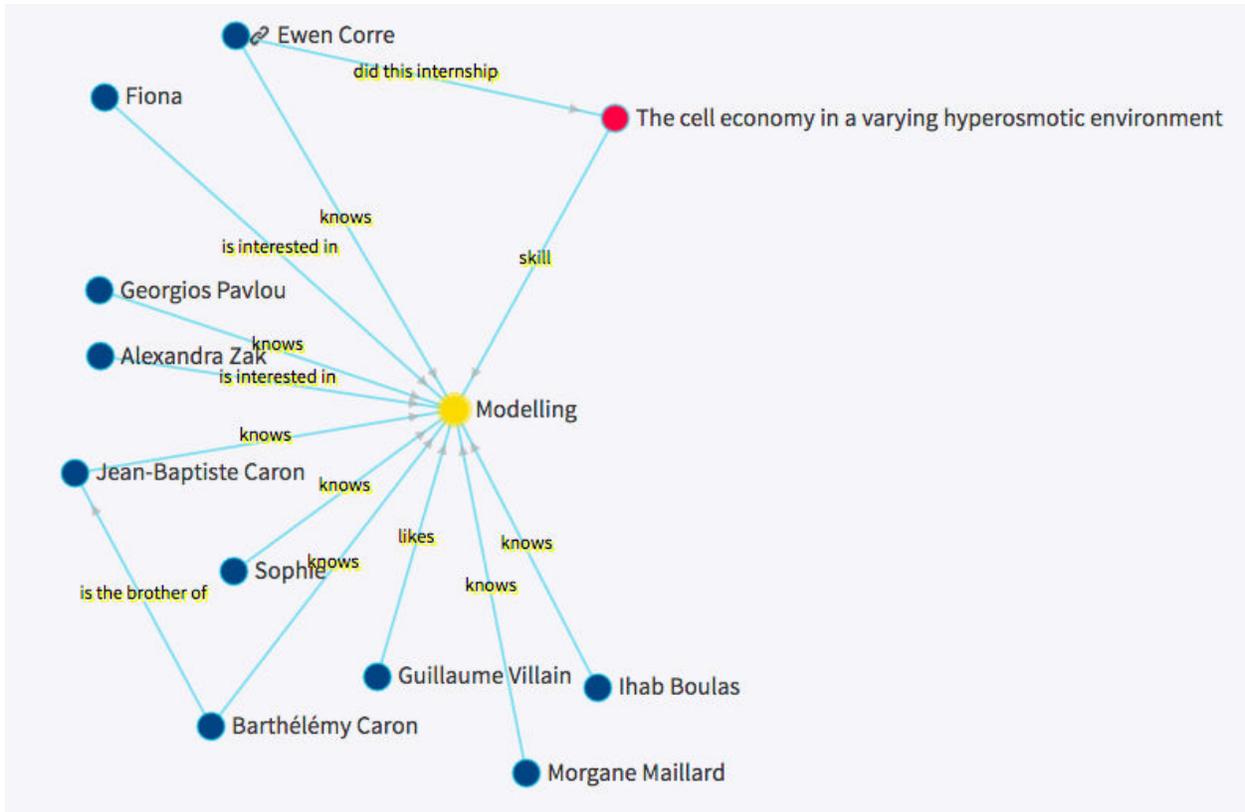


Figure 58: Clicking on the modeling skill and the resulting view.

A student can see: Which students in class and knowledgeable about Modelling and can probably refer to learning opportunities or advise on the subject. Which students are interested in the subject, therefore might be interested in co-learning with him. And an internship opportunity that has taught Modelling, and who did that internship. This given the student the knowledge who to contact to get a personal account about the quality and specifics about the internship.

Clicking on the internship itself reveals useful information about it (Figure 59):

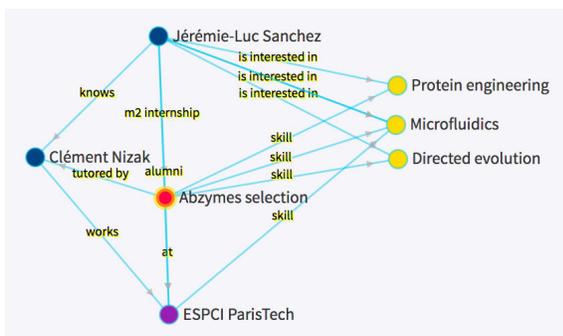


Figure 59: (left) Internship centric view (right) Infocard revealing additional information regarding the internship

A student can see which student can refer him to the internship, skills other students testified to have acquired during it, the name and details of the project tutor, and the institution in which it is hosted.

Clicking on each node also opens a card view with additional details regarding each entity such as a description or contact details.

Another interesting feature is the ability to select two entities and visualize the path between them. By using this feature a student can see how to best connect to an institution he wishes to attend (Figure 60):

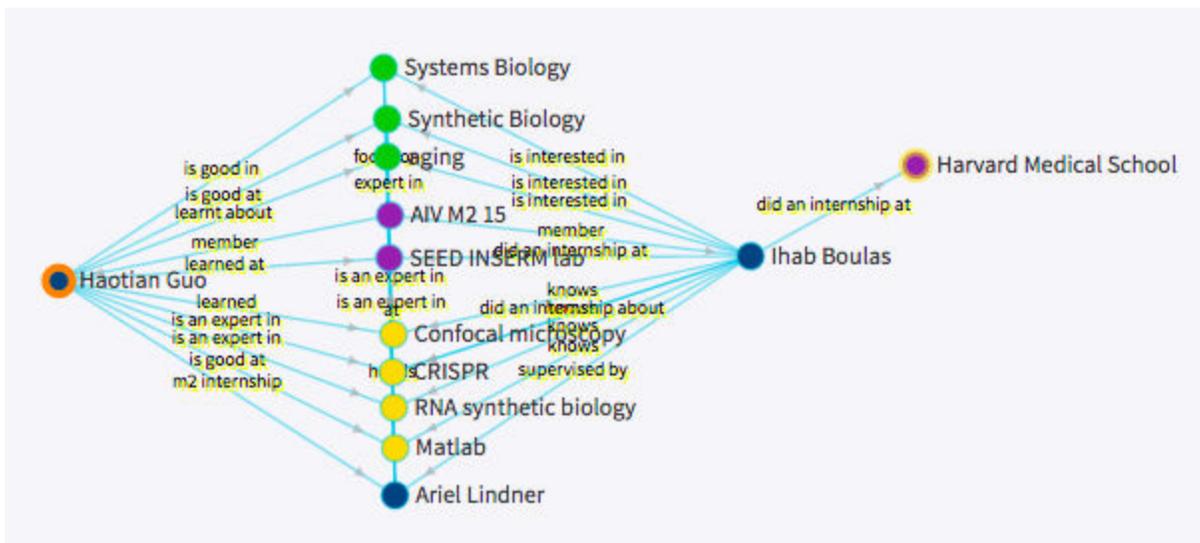


Figure 60: Shortest path visualization between a student and prospective institution.

In the example in figure 60, the student Haotian Guo is hypothetically interested to attend Harvard Medical school. Through the shortest path feature he can see there are multiple ways and contexts he can get in touch with Ihab Boulas who did an internship there. This could be through shared skills they possess, a shared supervisor, shared organisations or several keywords. He can see that Ihab is interested in Synthesis, Systems biology and Aging to all he declared proficiency. The network visualization provides not only the name of a contact, but different way to break the ice and find mutual interests within the interaction.

The table below specifies the type of content students have entered. Note that it was the AIV M2 2014 and 2015 that were explicitly prompted to enter internships according to the methods described.

Map title	Internships	Skills	People	Organisation	Keywords
AIV-M1-2014(Early prototype)	12	33	17	-	47
AIV-M1-2015	12	38	27	19	35
AIV-M2-2014	78	76	38	-	56
AIV-M2-2015	37	39	49	33	44
FdV-L1-2015	4	11	19	14	37
Master EdTech 2015	8	28	9	14	26

Figure 61: Types of nodes inputted by students for each map throughout the pilot

User participation

26 out of 35 students in the 2015 AIV M2 class created an account, and created content within the map. It is important to note that quite a lot of students were away on internships and were not even present in class during the project presentation.

Generally speaking, students filled the task requested, but often did not complete it as required. This is mostly due to the amount of work that was required of them, coupled with lack of any credit, or reward except an intrinsic motivation to use the system and help students in finding internships the year after them.

Beyond the internship mapping project, interactions with Rhizi inspired students and staff in the CRI to create additional maps. This is a notable achievement this was an unexpected result of the activity which inspired interesting use-cases:

- (1) Immunology club map: Created by a student who started a club to map biological data in a club she founded: “It's collaborative, dynamic, everyone can have access, everyone can add content, you can link whatever you want, for example, Ariel is using Rhizi for the CRI for connecting people, skills, make new connections, and we're using it for cells, molecules, so it's really broad range.” Also mentioned was the ability to create any type of connection you want without necessarily subscribe to a certain connection category: “you can create whatever link you want, any action that you want, it is in the dynamic in fact.” The map can be accessed here: <https://cri.rhizi.net/rz/immunetwork>

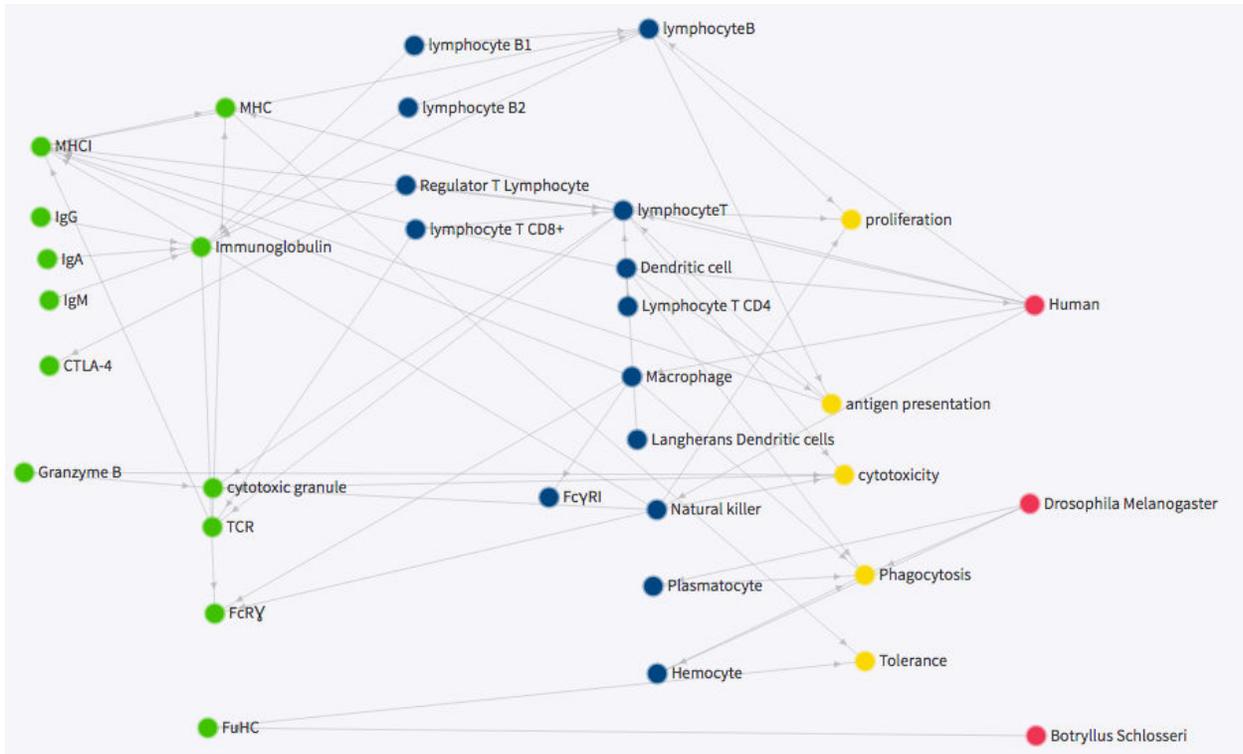


Fig 61B: Immunology club map

- (2) An CRI administrator created a map of documents and folders within the CRI shared online document repository: <https://cri.rhizi.net/rz/AIV-POC-Admin-docs-2014>
- (3) Igem project pilot started with students scaling up the work done in their classroom and applying a similar framework within the Igem competition (3.6 IGEM).
- (4) A map detailing CRI staff competencies was created under initiative of the HR manager. Details in [3.7 Other pilot projects](#).
- (5) A student created a map of a process associated to CRISPR: <https://cri.rhizi.net/rz/Ibrahim>
- (6) An anonymous user created a map of labs, researchers and key research concepts: <https://cri.rhizi.net/rz/JFD1> and <https://cri.rhizi.net/rz/JFD2>
- (7) A group working on a project proposal for open-source and collaborative drug-development created a graph of a use-case and a possible data-model for their proposal: <https://cri.rhizi.net/rz/Open%20Drug%20Discovery>
- (8) An incomplete testing map importing all skills, keywords, organisations and thesis advisors of the CRI's Doctoral students: <https://cri.rhizi.net/rz/Phd%20test>
- (9) Anonymous user has created a mapping of the “Laboratoire d'informatique de Paris 6” this includes, projects, keywords and people: <https://cri.rhizi.net/rz/Welcome%20Rhizi>
- (10) The 2016 Igem team mapped candidates competencies for the team formation stage of their project: <https://cri.rhizi.net/rz/iGEM%202016>

Semi-structured interviews

I conducted video interviews with 10 CRI students that experienced Rhizi throughout their academic year. All of them participated in the opportunity mapping tasks and two of them participated in the Rhizi Igem project (See [chapter 3.6](#)). The most notable finding from the interviews was that all students say that they would like to use it again next year, and those who were asked, said they would recommend it to their friends from other organisations. Another important result is that students reported gaining value from discovering other students interests and skills and reported practical value that would be given to them by finding people to help provide information about internship opportunities and for providing technical assistance regarding experimentation techniques.

Value of Rhizi

When asked about the value they got from the system they provided the following responses:

Conveying complexity: “The power of Rhizi is that you could do the same in principle in an excel sheet, but you will never get the same complexity in terms of connections that you can visualize and it will give you more insights I’m sure. It is much harder to look at 2-d plots or a table and to see the same complexity than what is possible with a tool like Rhizi.” another student added: “It’s already doing a little like what is happening in your head when you speak of think about someone of something, it’s the kind of web we have in the head when we speak of think about someone, but here it’s more concrete and it’s not limited to what you know, but what everyone in the connection know. It’s pretty easy to understand how it works and how you get information from it.”

Ability to give and receive help to fellow students: “To help with techniques, is to ask people working in the lab or the lab right next to us. I guess Rhizi in a way is useful because we really don’t try to find these people initially, people in the same school, in france, they don’t work in the same institute, so they are sparse. At least as a master student” Rhizi provides people a socially acceptable way to seek and give help “In this context people will not answer an unknown person, if you have facebook friends you will respect you more than a random person, you belong to the system, the chances for benefit are larger”, “It makes people collaborate. Collaborate[in the mapping] to collaborate.” In conversation people were very interested in the tool to facilitate mutual help as an example a student talked about a rare experience he had in the CRI that he wished he could have more through Rhizi: “Once i talked with an FDV student about my current project, I was using PCA for something, she told me i had the same problem I went from PCA to ICA, and it worked well I published with this tool. I didn’t know this tool at all, it’s way better than PCA for this and i’ll publish with it’ and “If i had a problem, it would be nice to click on “Java” and see who can help me”, “Usually in the lab there are a lot of technical problems with experiments, experiments always fail, but after you get

experience you can do it more quickly. If you met somebody who has experience and can give you the shortcut to do it well and fast, it would be really nice to meet someone like that.” and “You always need somebody to help, and they are not always available in the place you work”. Another value point was getting to understand what people are working on right now, rather than end of year presentations “...it’s quite late when you see someone presenting at the end of his internship, now you understand that ‘oh it’s quite nice’, but you don’t talk to him during this”

Finding interesting people in the CRI: One student testified that though a shared interest in “Transhumanism” discovered by Rhizi, students found each other and formed a group together about the subject. , another added, “For my internship and Phd project I need to be more clever in statistics and bioinformatics in general. People who can explain things very efficiently in statistics and computer science. “

Compact information: “When you are on the website on an organisation or a lab, you have to search for each information really in different corners, oh I want to contact, to see who is in the lab, oh i want to find the projects they are doing, what skill they are doing, it’s really going in all directions, here you just click on the lab and you know who is in, which are the thematics, what are the skills used, what kind of internship have been done there, that’s nice”

Learn about labs and internships: “I’m glad to have found some labs that I want collaborate with, maybe in the close future I would need to create some concrete collaboration and have some real information about what people I know are doing can be an advantage”

Showcasing the CRI: “Showcase the CRI as a big knowledge map. Show to the world what the CRI can do.”

One things that came up during the interviews was that students who received value from Rhizi went on a tried to create their own maps, tried to disseminate the software to their lab, club and generally were involved in adding data. Those who didn’t receive value were far less involved.

Students did not see a lot of value in the network visualization in itself, when directly asked, none of the 10 interviewees mentioned it gave them real insights. This will be discussed in the discussion segment.

What was interesting for students to look at?

Because of the broad scope of information available to the students, I was curious to find out which parts of it were the most interesting and important for them.

Learn more about their classmates: “What they were up to“, as one them explains: “Sometime, I have friends in the master that I highly consider, when we speak together we are not so professional so I don’t manage to understand about their exact professional life, it can be nice to spy a little about what people are doing”.

Skills, internships and projects of colleagues: “I first clicked on my friends’ nodes to see what they were capable to do, what internships they did, what people they were in contact with, to ask questions about techniques or internship to search for.” Several students were interested in finding people connected to specific internships or keywords because they wanted to learn more about related internship opportunities. Another student was interested in other students interests because he wanted to form a photography club.

Issues

When asked about things they wished to improve in the system or problems they have encountered, Our main finding were:

Onboarding issues: One student mentioned that the the system seemed complicated when he initially tried to decide whether to start using it but that it was in fact very straightforward and quick. This suggests the onboarding process can be improved. Another suggested that “In the organisation side we should have sessions dedicated to this - one hour for mapping themselves in Rhizi.” Lack accreditation for the work done makes it hard to compete with other tasks, of which students are accredited and have a more direct impact on their goals.

Ontology management issues: “I don’t really see who will do the job of linking skills together. If it’s always in the egoist way, we can’t be something else than egoist, now I understand it’s nice to link evolution and evolution in-silico, but i can’t do the job. No one will not do this if he is not paid to, or if it’s not his thesis subject”. Another issues was deciding between usage of the node category “skill” and “keyword”, especially since sometimes a given category such as “immunology” was interpreted by the student as being both.

Adapting to structuring input: “I was thinking about how to put the information in the best way... I had to feel the network in a way, I wasn’t sure what was the best way, so it took me a long time to decide”

Unclear instructions: Two students mentioned that it wasn’t clear from the instructions whether they should map their skills strictly in the context of the CRI, or it should be a more general mapping of their skills, including past skills. “I had difficulties to see if we were to put out some professional scientific knowledge in it or what you know how to do outside of the classic cursus”. This suggests that the framing of the task should have been more focused.

Readability on larger maps: One student mentioned readability issues because there was too much information in one map.

Feature requests

Generally speaking, students were pretty happy with the existing features. There were some suggestions related to:

Adding more functionality to the connections: Adding the ability to define Bi-directional connections, multiple connections between two entities and ability to draw lines between nodes rather than use the text input interface.

Keeping track of new content: Add functionality which would make it easier to see what new content was created since they viewed a map.

Add timing functionality: “For example Charles Baharud is linked to Ladic lab, he a physicist, and you can see him at the CRI in one or two weeks. It could be useful.”

Discussion

The Internship opportunity mapping project was key to the project in general, and therefore most of the discussion regarding it is in [chapter 4](#), however, below are a few pilot-specific thoughts I thought would be useful to include here.

Lessons learned

While our team did the best we could to address the challenges that were exposed in the Bibsyn project, the success was only partial. This is due to:

- **Better onboarding:** This challenge can be easily addressed as we have recently developed an import system that enables us to populate the graph from CSV data from all previous Phd students at the CRI. Asking a student to correct existing data about himself rather than create completely new information should dramatically improve the onboarding experience.
- **No accreditation or reward for participation:** This is a systematic problem we encountered in the Descartes IT pilot as well. The solution is to attach real and immediate intrinsic value to the mapping task. See chapter 4.1 for my proposal how this can be done for future iteration of the project.
- **Too many competing tasks and entities:** While there was considerable work done to simplify the tasks, there was not sufficient buy-in from the CRI staff to support including less types of data and different mapping tasks in the system. My proposal to address this challenge by focus on internships and skills rather than trying to map internships, skills, interests, organisations, tutors and keywords.

- Not gaining a big enough critical mass of content to attract student attention: This can be resolved by (i) Auto-importing information (ii) Creating a unified map of all students in the CRI, as well as a global search functionality (iii) Integrating the usage of Rhizi in a more structured way with the timeline of the CRI academic year, specifically coordination with deadlines to apply for internships.

Perspectives

One interesting potential use-case resulting from this pilot is the ability for students and their tutors to discuss student personal learning goals and setting together plans to achieve these goals by using the ego-centered map created by the students.

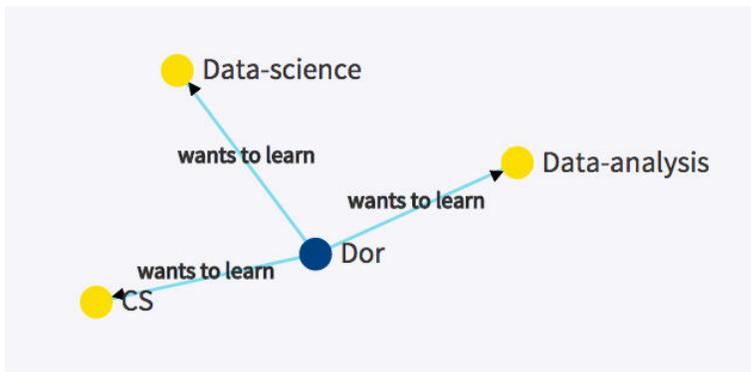


Figure 62: Mapping learning intentions

In the example in figure 62, a student, “Dor”, defined certain skills he wants to learn this year. In figure 63, you can see how together with his tutor or peer students he can set a learning plan matching learning opportunities with goals he has set for himself. This can create a “learning opportunity” market of sorts, in which students can express their goals, while people who offer internship can try and match their offerings to the students and vice-versa.

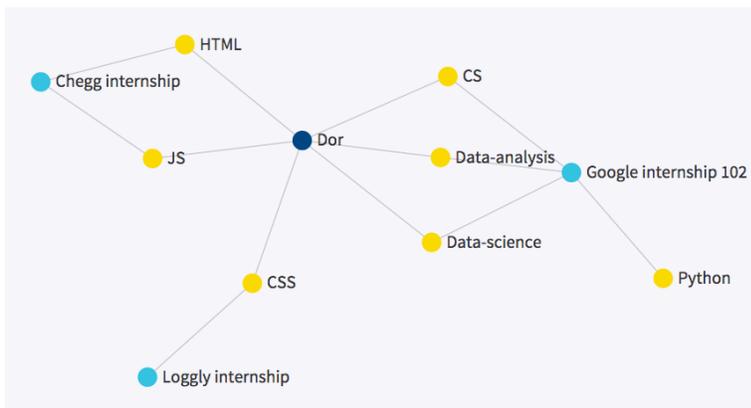


Figure 63: Mapping learning Intentions to learning opportunities

Another perspective is adding a notification system and visual feedback about relevancy of internship proposals to the students.

Scaling up to more students would potentially create even more value, as one of the students said the interview: “It could be important for bachelor students because it’s hard to know someone in the amphitheatre, in big universities”

Descartes University IT Department

Background

As a case study we chose to work with an IT department of a large university. The IT department services for the needs of ~4000 researchers, ~50,000 students and engaged within a dense network of cross-institutional R&D development. With 114 employees, 120 ongoing projects and more than 100 collaborating institutions, only several managers possess an internal model of who is an expert in what field and which projects which person is involved with. Because of the scale and variety of its responsibilities, we find that it exemplifies many of the organizational challenges facing knowledge-intensive organisations.

Objectives

Our overall objective was to lower the energy barrier for sharing knowledge capital within organisations. We focused on improving the expert identification process, and the knowledge fragmentation problem that stands at its root. Knowledge fragmentation happens when relevant pieces of information are stored in silos and therefore are difficult to aggregate together. This can take shape as informal knowledge about experience or skill is distributed across people's heads ("Carl knows how to fix the copying machine"), as unstructured information such as cv's, emails and documents, or even as structured data that is stored in disparate databases, spreadsheets, and project management software.

Accordingly, the challenge organisations face when defragmenting knowledge is two-fold: From the input side, how to efficiently input and update the skills and project-related experience of individuals? From the output side, how to reduce the cost of searching and sorting this information? In most organizations, deriving these insights requires manual and expensive human intervention or a plethora of mostly proprietary softwares.

Characterize existing expert identification issues

As a case study we chose to work with an IT department of a large university. The IT department services for the needs of ~4000 researchers, ~50,000 students and engaged within a dense network of cross-institutional R&D development. With 114 employees, 120 ongoing projects and more than 100 collaborating institutions, only several managers possess an internal model of who is an expert in what field and which projects which person is involved with. Because of the scale and variety of its responsibilities, we find that it exemplifies many of the organizational challenges facing knowledge-intensive organisations.

We first started by engaging in conversation with the head of the IT department, and define the main issues that are both important and practical to solve. Following these discussions we have set

three sub-objectives that follow a complete cycle of product development: First, to characterize existing issues that the staff of the IT department was experiencing when trying to identify experts beyond their immediate team. Second, to test the feasibility of our mapping system by deploying Rhizi within the department. Lastly, to gather feedback and conduct several tests to gain further insight and chart the future development of the system.

Our analysis of the IT department led to the identification of the following key problems:

- *High perceived cost for finding and contacting peer-experts:* Where the main causes for the perceived cost are:
 - Having to involve managers, resulting in time loss for both requester and manager as well as potentially appearing less competent in front of the manager.
 - Failure to find the expert because manager or colleague were not aware of the expert.
 - Loss of productivity as new and remote employees take a long time (1-2 years) to collaborate effectively with other department members.
 - Lastly, outsiders have no way to know who to contact when needing assistance.
- *Lack of “big picture” understanding of current projects and the involvement of others in them.* This stems from the fact that data about employee involvement is distributed across multiple databases and application without synthesis of the information in an easily accessible way.

Following this we have set the following goals for the Rhizi mapping system:

1. Achieve a real-time representation of relationships between project, people and skills inside the IT department.
2. Develop a transparent and collaborative platform that will enable IT staff to identify experts and explore the organisational map.
3. Work towards a transparent representation of the department's projects and skills that would enable even further collaboration, down the line, with other university departments and outside collaborators.

Together, we have decided that identifying sources of intelligence rather than making the knowledge explicit like in a Wiki or a knowledge base is preferable because the work of relating people to skills and projects is considerably cheaper than the creation and maintenance of a knowledge base. Incentivizing people to share their skills and keep it updated, however, is vital. This is why we thought it critical to provide an engaging input interface for the knowledge graph.

Lastly, we have articulated together the desired outcome of our collaboration: To empower staff to see the big picture, freeing individuals within the department to collaborate without relying on intermediaries. Provide staff an independent way to gain awareness of the collaboration opportunities available to them.

Based on our design goals, we expected that the use of Rhizi would affect users' experience along four dimensions.

- Lower staff's perceived cost of collaborating with others by provide a way for people to identify and contact experts

- Improve awareness to the overall organisation and their relationship to others.
- Motivate employees to update their information in system.
- Assist management to evaluate the needs of future projects and formulate training sessions accordingly.

We specify in the methods and results section 7.1 the feedback received and its analysis.

Usage

Following the goal-setting with the IT leadership, we defined principles we wanted to follow while implementing the system. Most importantly, we wanted to the system to communicate to the IT members that we see them as “citizen scientists” of their own organisation. Framing the problem in such a way that the staff is aware of what they are doing and building. We wanted to create a individual-centric model for collaboration, so an agent can change and modify his skills and experience dynamically, immediately affecting the organisation's' perception of its affordances.

Having each employee spend a few minutes mapping his own local perspective is a lot more efficient than one person “doing the rounds” and interviewing everyone and then transcribing it into a database. By empowering people to map together, we get a comprehensive understanding of the whole system, giving us not only insight, but also a human feedback layer to update this information as things change.

These principles influenced us to add the following requirements to Rhizi:

1. Collaboration features - Users have unique accounts from which they can add their contribution, each user can create and edit entities in the map.
2. Transparency - User actions are public and logged. All maps and content are accessible to anyone with a registered account.
3. Real-time - Once an actions has been made, it is immediately visible for other clients.
4. WYSIWYG(“What You See Is What You Get”) interface - Similar to the relationship between a table database and a Spreadsheet, every action on the network is represented directly as a network visualization.

Another design choice was to use the metaphor of a map to describe the actions taking place in the system. Mapping is an abstraction used to navigate physical space, and we think the same abstraction is a useful one to familiarize users with the concept of navigating resource and opportunity spaces.

These principles represent how our approach differs from existing expert mapping and network visualization platforms. Rhizi is designed to be used and understood by all the participants involved in the mapping process, rather than subjects of a scientific enquiry, from which data is sourced, later to be exclusively analysed and and valorized solely by managers and analysts.

Usage in the department

We provide several examples of how the IT department used the tool. For the sake of anonymity, we have replaced names of projects, people and organisations with numbers. Usage period was between 1.1.2016 - 1.4.2016

Sources of data within Rhizi

The data within Rhizi originated from the following sources:

1. *Direct input by IT management during interviews with staff members.*
2. *Direct input by a subset of participants who added or modified data by themselves.*

IT personnel officially started self-mapping with Rhizi one month after an initial meeting which introduced them to the tool. The self-mapping task consists of reviewing, correcting, and adding information about their skills, roles, and relationships with their peers. They were also invited to explore the data already available in order to familiarize themselves with the competency landscape of their organization.

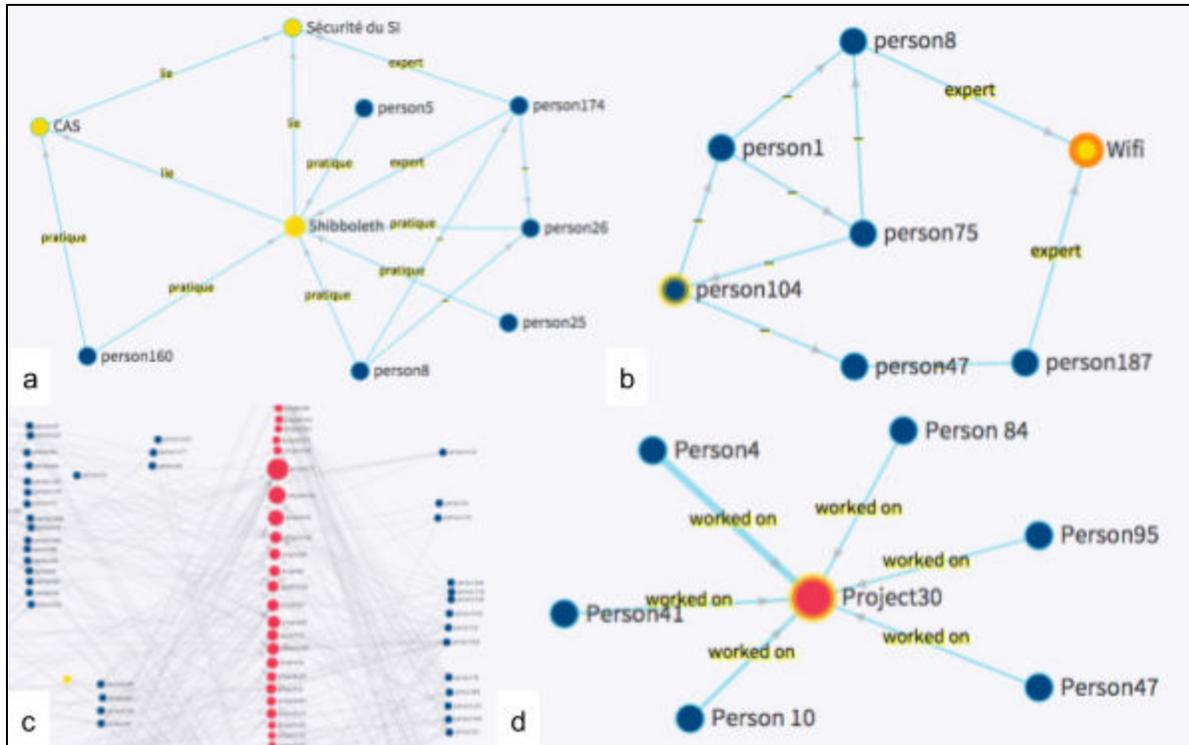
Finding competencies: Using the search interface, any staff member can type a skill he is looking for. In the example below - the employee queried the term “Shibboleth” - an Identity management technology. The result is two technologies identified as related, and several people of varied degrees of expertise. “Pratique” as in a normal practitioner and “expert”(figure 47a).

Shortest path: A person looking for Wifi expert can find multiple interlockers to connect him to a Wifi expert(figure 47b).

Project overview analysis: We have imported this data into the Rhizi system as a network of relationships between people and projects, where the width of the connections between people and project represents the amount of time they have invested and the size of the project node represents the total amount of time dedicated to the project by all employees.(figure 47c)

Staff experience: Selecting an individual staff member in the Project map, provides a visualization of how he allocates his time between projects. Link is thicker the more time is allocated.(figure 47d)

Selecting a project: shows which people are most experienced in it.



47 : (a) Query result for Shibboleth (b) Shortest path between person104 to Wifi skill (c) Project overview - node size depict total amount of work hours invested in project (d) project centered view with time-allocation visualization per staff member.

Methods

Participants

The IT department is responsible for providing IT services to a university with ~35,000 students. It has 114 employees, ~80 listed skills, ~120 ongoing projects <100 collaborating organisations and institutions. Employees are distributed across multiple offices across the city and a small percentage of them ~10% work remotely.

Each methodology was composed of a different cohort from the total of 114 IT Staff members:

The online survey was filled online and offline by a total of 19 participants. 7 of whom were women, 7 worked remotely. With regards to roles: 8 managers, 8 technical and 3 administrators. All worked for longer than a year in the department.

The expert identification quiz was filled by 13 staff members, 4 of them were women, all worked on premise, average seniority of 5.1Y, average age 39.

Semi-structured interviews were conducted with 8 staff members, 1 of whom was a women. None worked remotely. The jobs of the 8 were as follows: Technical assistance, IT administrator, System engineer(X2), webmaster/graphic designer, and manager(X3)

Online survey

In order to validate the problem of expertise identification we have sent an online survey designed to examine which, if any, aspects of peer collaboration and identification are important to members of the staff. The survey consisted of a series of statements which respondents could rate on a 7-point Likert matrix table (Krosnick, Lavrakas, & Nuri, 2000) to collect feedback about staff's experience with using Rhizi. Invitations to participate were sent through email. Out of 114 employees contacted 19 filled the survey. To make the result presentation more clear, we have merged the 7-point into a 3-point Likert matrix (Agree, neither agree nor disagree, disagree). We have also asked respondents a series of open-ended questions to provide us with some feedback about issues they had and suggestions for improving the system.

Expert identification quiz

In order to understand how much people can identify and effectively direct others to an expert, we composed a quiz asking employees to identify a specific person who can help us with 10 requests for help.

These help requests require expertise in a certain technology. A secondary objective was to understand what happened when the subject didn't manage to identify an expert by name. The expert identification quiz were asked in-person in direct dialogue with the subject and recorded as audio.

Questions asked in expert identification quiz

Can you refer me to a specific colleague or colleagues who can help me with the following requests?

- 1) Doing a complicated graphic work on Photoshop.
- 2) Help with changing settings in the Jefyco admin.
- 3) Setting an API with o365.
- 4) Setting up EZpublish to work with a smartTV.
- 5) Help for Development using .NET language.
- 6) Setting up Shibboleth support for a new management software.
- 7) Completing a project using HTML5.
- 8) Making a short training video.
- 9) Consultation about distributed computation for a university project.
- 10) Consultation about using Moodle.

Semi-structured Interviews

In contrast to the very structured expert identification quiz and survey, we used semi-structured interviews (Wengraf, 2001) to explore deeper into people's experience using Rhizi. This helped us learn more about their thoughts about collaboration, giving and receiving help and ideas regarding future implementations of the software. We also thought it would be a good opportunity to understand better how these subjects relate to their daily work-life challenges.

Rhizi-assisted expert identification quiz

We wanted to test if Rhizi provides better results for expert identification than asking employees. To this end, we have asked 5 people (out of which 2 women) who are not involved professionally with the IT department to use Rhizi to provide answers to the same 10 questions that were given to the IT staff. These sessions were made in an informal setting. It included a 2 minute demonstration of how to use the search function to search for skills inside Rhizi.

Data Gathering and analysis

In total we received 140 distinct responses to the expert identification quiz (10 responses from each participant), and 50 responses to the Rhizi assisted expert identification quiz. We also obtained 416 minutes of video recordings from the semi-structured interviews. Drawing from the interviews, we coded the responses (Burnard & Philip, 1991) to each question and tried to identify common and notable themes raised by participants. With the expert identification quiz, we coded responses when the participant wasn't certain who the expert is.

Results

Usage in the Paris Descartes University IT department

We provide several examples of how the IT department used the tool. For the sake of anonymity, we have replaced names of projects, people and organisations with numbers. Usage period was between 1.1.2016 - 1.4.2016

Finding competencies

Using the search interface, any staff member can type a skill he is looking for. In the example below - the employee queried the term "Shibboleth" - an Identity management technology. The result is two technologies identified as related, and several people of varied degrees of expertise. "Pratique" as in a normal practitioner and "expert" (figure 47a).

Shortest paths

Using this feature enables to find the short, or several shortest paths if path length is equal between two entities. For example, A person looking for Wifi expert can find multiple interlockers to connect him to a Wifi expert.(figure 47b).

Project overview analysis

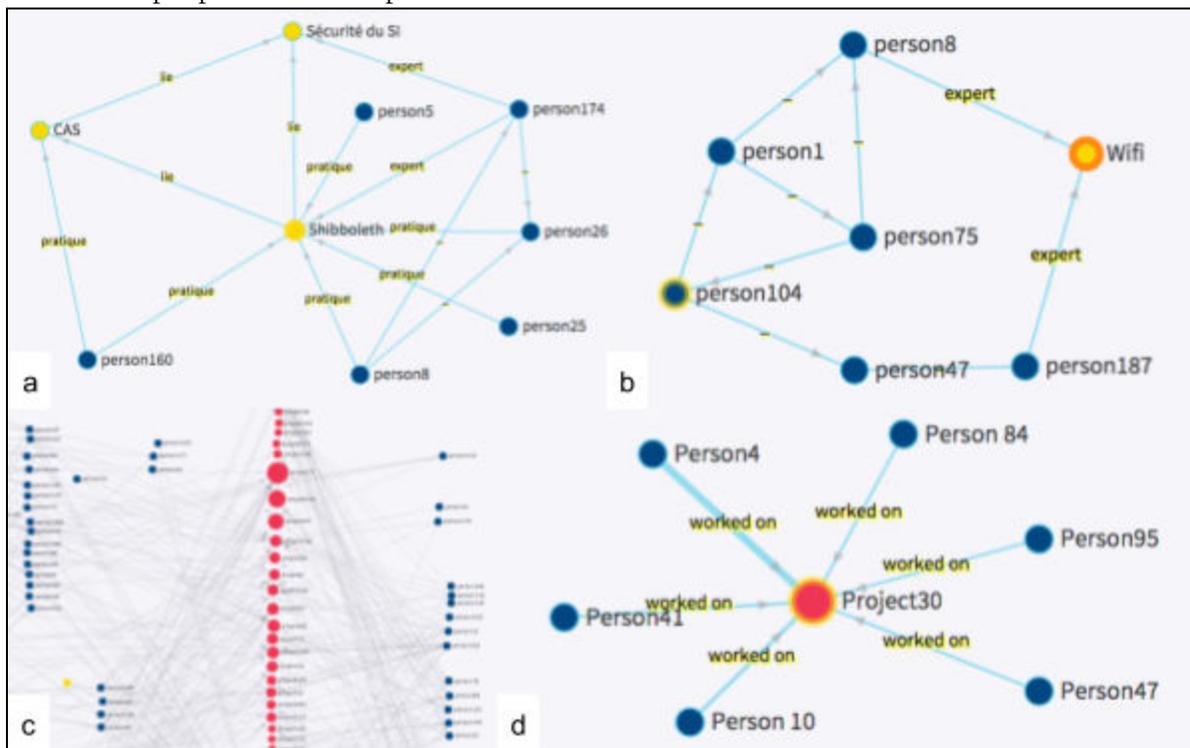
We have imported this data into the Rhizi system as a network of relationships between people and projects, where the width of the connections between people and project represents the amount of time they have invested and the size of the project node represents the total amount of time dedicated to the project by all employees.(figure 47c)

Staff experience

Selecting an individual staff member in the Project map, provides a visualization of how he allocates his time between projects. Link is thicker the more time is allocated.(figure 47d)

Selecting a project

Shows which people are most experienced in it.



47 : (a) Query result for Shibboleth (b) Shortest path between person104 to Wifi skill (c) Project overview - node size depict total amount of work hours invested in project (d) project centered view with time-allocation visualization per staff member.

Survey and Interviews

Low ability to identify experts as an underlying bottleneck for interdepartmental operations

We characterized expert identification issues and identified how they relate to the more general challenges the organisation was facing. The following subsections cover each one of our findings.

Self-reported need for expertise and project awareness : Survey results confirm our hypothesis that the IT staff indeed thinks that having better knowledge about projects, skills and collaboration are needed. The usefulness of being able to quickly identify an expert was also validated.

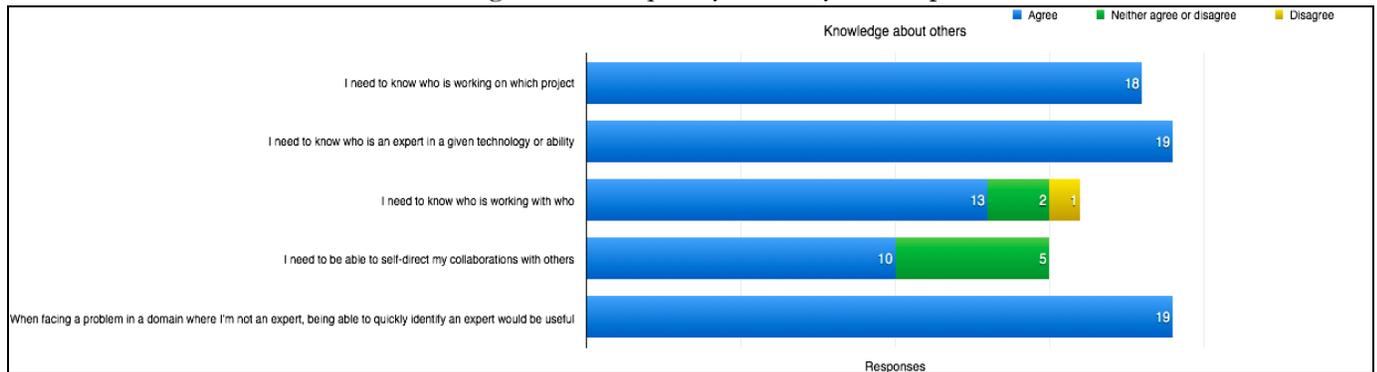


Figure 64: Survey results on questions related to needs related to expertise and project awareness

Expertise identification is cumbersome and inefficient in current practice - In contrast to the needs stated above, staff had difficulties to identify experts throughout the department. We have composed a 10 question quiz. Each question was for formulated as a request for help in a specific technical task. We asked the staff members to identify by name specific experts who they think can help us with the request. Running the test on 14 subjects, a specific name of a person was given only 44.6% of the time (58 out of 140). We didn't allow the staff involved to look at Rhizi for the answer, nor to ask their colleagues, they were free to use any other tools.

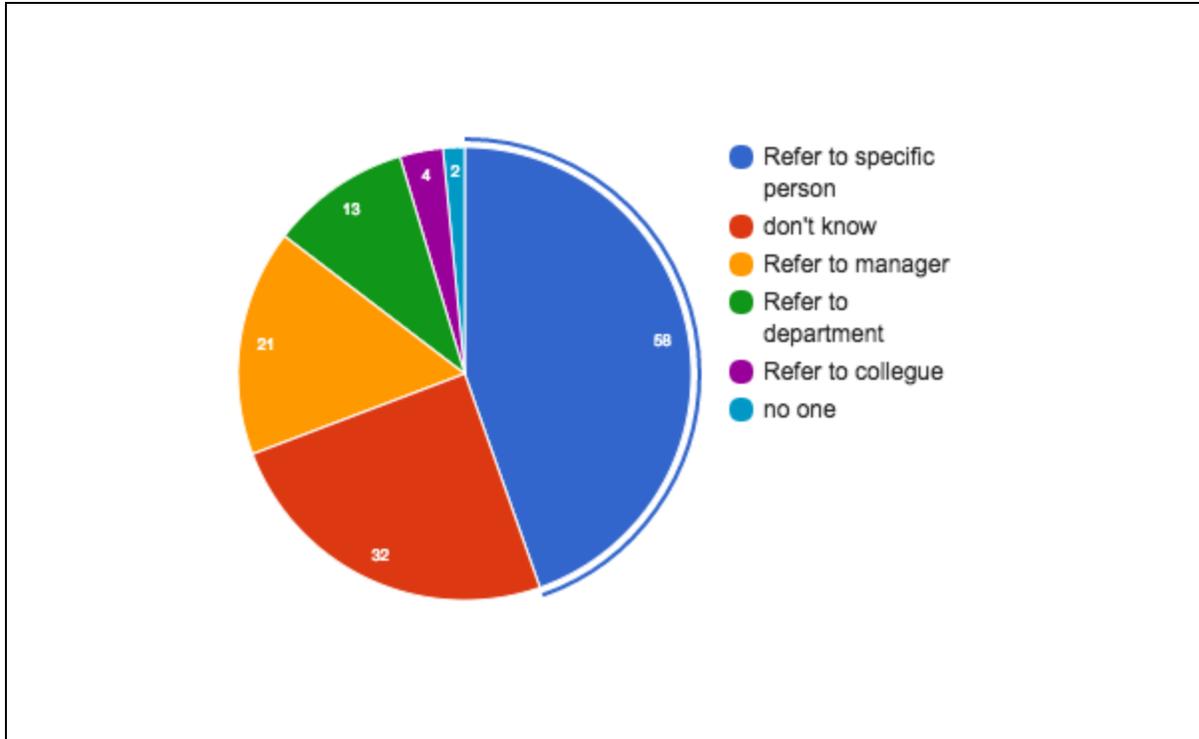


Figure 65: Results of the expert identification quiz.

We think these results indicate that expert identification is indeed an existing problem within the department. Average seniority of the recipients was 5.1Y and so they had plenty of opportunities to become familiar with other people’s activities through the bi-annual meetings and other work related communication activity such as emails and newsletters.

The large amount of “don’t know”(32 responses) and referrals to managers(21), or departments(13) also means that if a hypothetical person wished to casually ask someone for help and consulting on an issue, they were forced to initiate a more formal process of asking administrators or managers. Small number of colleagues referred(4), might suggest expertise silos, where staff can’t identify experts and often doesn’t know who can identify them. This confirms our hypothesis about the high perceived cost expert identification. See section 7.2 for validation of the Rhizi-assisted solution of this issue.

New and remote employees have a higher barriers to identifying experts: As one technician mentioned during his interview: “A new employee arrives to a local team, that describes to him their local functionality, but the global functioning of the university is very complex, and impossible to ascertain when one arrives.” On another interview, the Manager of the IT department said that having a new or remote employee understand the “who does what” in the department may take between 1-2 years.

Project overflow due to bureaucratic bottlenecks - 4 interviewees (3 of them managers)

emphasized having too many projects in proportions to allocated resources: “IT Administration is like a big company, we do the job like a big company but not with [big company] money”, “We begin a lot of projects but not finish them or after time”.

Reasons for this project overflow, such as budgetary constraints, are outside the scope of this research to address. However, two relevant underlying reasons for the expert identification problem were brought up. The first is that projects were getting delayed because of “weak links” in the bureaucratic chain: “You launch a project and it can have big delays on different reasons but mainly because administration and a lot of people that are not always involved in their work in this big administration”.

The second reason given was long delays in cross-team communication. One staff member gave an example of how a solution that should only take several hours to implement, took in practice four months to complete. “I made a request, I waited for a reply. One week, I was told, no problem, we will do it. One week later, we are in the process of doing it, how can we install the thing. I said to them, I can do it with you. ‘We don’t want you to do this with us.’ So how do you want to do this?, one week later the person in charge says, ‘ok, let [subject’s name] work with you to show you’, then nothing for a month or two, and then, ‘ok, [subject’s name], you can come over.’ they have set an appointment, for two weeks later. Four months”. When our own researcher tried to coordinate meetings within the IT team for interviews, emails were often ignored, and an alternative system of directly phoning people through their personal mobile device was often the only way for him to establish contact with the subjects.

Lack of coordination on interdepartmental projects may cause sub-optimal or duplicate solutions: “Methods are not well shared, there is a loss for the organisation, because they don’t have a record who already did such a thing. So the project is successful because it is here, but it didn’t have all the benefits it could have because it wasn’t shared globally. All project are like this, it’s really too bad that we don’t share these things. It’s a loss that is hard to measure, but nevertheless exist, because no one can retake on this project tomorrow, “ and so, it’s not only that projects take longer, but there are often duplicated efforts and sub-optimal solution due to lack of communication between the experts.

Lack of expert coordination can also lead to a “Reinventing the wheel” phenomena: “We now have one [technology name removed] platform, but it used to be that each one had it’s own [technology name removed]. In [name of department removed], each one had it’s own platform. It’s very inefficient, We see that people are not directed. It’s just in IT, it’s everywhere. The [redacted faculty name] asks to do something, and it’s separated when the university wants to do something, it works like that, it’s not just IT.”

This is in line with earlier reports as a 2014 survey of 152 Campus IT leaders in the US, that estimated that 19% of the campus IT systems are redundant, costing US universities 3.8B\$ per year(Cloud campus 2016) .

Need for expert collaboration increases in times of rapid change, or when working on cross-departmental projects: In times of change collaboration becomes crucial. “When we all face the same problem on the same time. When there’s something new. When there’s a new software to install, there is a period of time where knowledge is not well structured.” Another employee initiated and led a bottom-up effort to create a platform that is serving several university departments and requires collaboration from different and diverse fields of technical expertise. He had a lot of trouble identifying the right collaborators: “I had to meet many people to do that, but could never manage to find the time to find out who did what, who knows what and who could help me ‘carry’ this project.”

Conclusion: Difficulty to identify and directly communicate with experts leads to project overflow and lack of coordination. High perceived cost and uncertain results of trying to manually identify experts leads people to try to solve issues themselves. If the expertise is necessary for the continuation of the project, this can serve as a bottleneck that leads to unnecessarily delays. A synthesis of the barriers for effective cross-team collaboration can be seen in figure 4. The perceived cost was well established from our quiz and interviews. The second part, lack of incentives for collaboration is covered in the literature review in section 1.2 under the *Limitations* sub-section and is further covered in the discussion section.

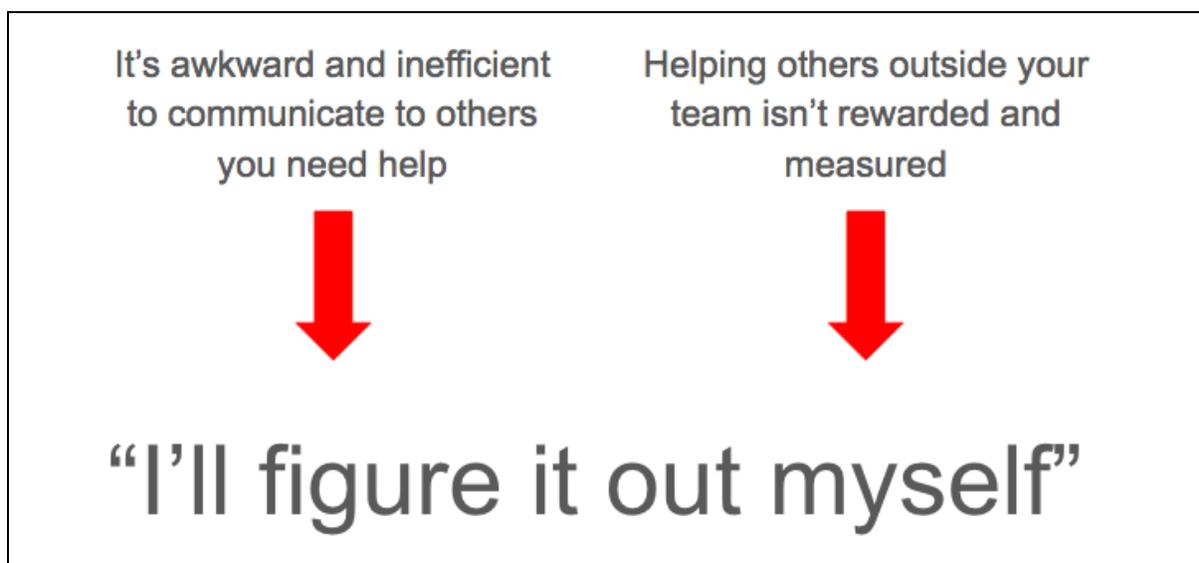


Figure 4: Simple breakdown of perceived cost and lack of incentives for organisation wide collaboration.

Validation of Rhizi as an effective solution for expert identification

Creation of a transparent visual database of expertise and project experience: Two maps were created as a result of the study case. A “Digital experts” map containing people, organisations and expertise that mapped not just the department, but key people in relationship with the department's staff. The Project map is sourced from a time reporting tool, and makes it explicit how many days each person dedicated to each project. The following table (table 2) represents number of entities mapped:

Map name	People	Projects	Skills	Organisations	Links	Edges/Nodes ratio
Digital experts	218	-	155	91	1309	2.82
Digital projects	61	125	-	-	542	2.91

figure 67: Number of entities in each of the resulting maps, according to Entity type. (True for 04.15.2016)

Employees had transparent and formal way to know who’s doing what and who possesses what skills. This allows any employee to make a informed decision about who to contact in case of blockage or if a consultation is needed.

Rhizi allows successful identification of experts: To validate the solution, we have redone the expert identification quiz questions described in section 6.5 with 5 random participants who have no association to the IT department. The outsiders success rate of naming specific expert was 100%, and often their recommendation included a list of alternative experts if the first expert could not be reached. The process took 2m on average, where the employees in the original expert identification quiz took on average ~9m to answer the 10 questions.

Usage experience beneficial for motivation for sharing expertise: Overall, survey results show that the maps created had a general positive impact on people’s ability to identify experts, and motivate staff to share their expertise and reach out for others .

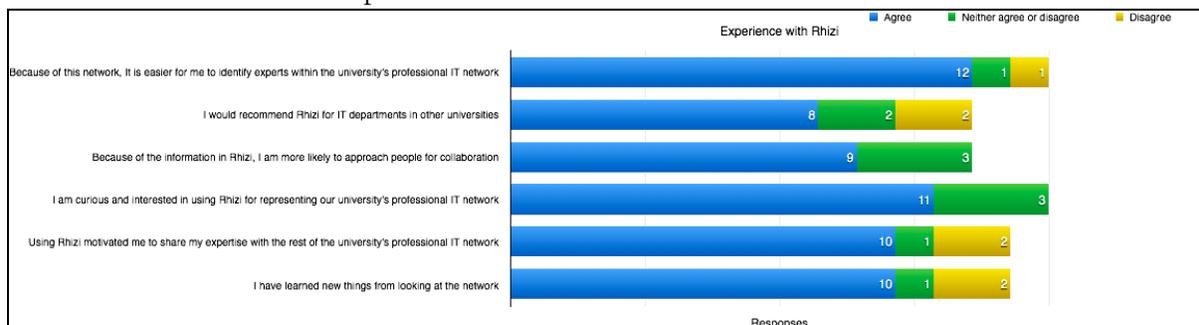


Figure 68: Survey results on questions related to experience with Rhizi.

As one respondent said in the interview: “When I started my project, I looked for people who knew [web technology]. That could help me to start. I did it by myself. Now, I look first in Rhizi.”

Understanding the organization's' direction and competency gaps : Several respondents said that the maps helped them understand where the organisation as a whole is advancing to: “I like it because I can imagine what is the direction we are taking in the university, mobile or this service or pedagogy” For another respondent the system-level view of resources allowed him to identify vulnerabilities: “[name of staff member] leaves in three months, and I know that all her competencies will disappear, that will allow me to see which people I should look for to replace those. As a manager it’s interesting, if I see the project she is working on, I can make sure she passes the information to those who replace her.”

In the digital experts map, we have identified several skills that were weak-points, indeed it would be simple to see if an employee leaves, which skills remain as “orphan nodes” and make sure to train or hire for these skills. Out of skills listed 50 were identified as weak-points having only one person marked as an expert in them.

The Project map, out of which 25 project had 1 staff member logged as involved with.

Shortcoming and further development routes identified for the Rhizi software development/deployment

Negative criticism towards the network visualization interface: The major negative criticism about the system was the complexity of the network visualization interface. While heads of the IT department quickly learned how to use it, most casual users didn’t like the unfamiliar interface, and the complexity in the map. This was also repeated when we asked as an open question to describe negative or frustrating experience with Rhizi. This also reinforces the fact that *despite* the complex interface, people still found value in the underlying information.

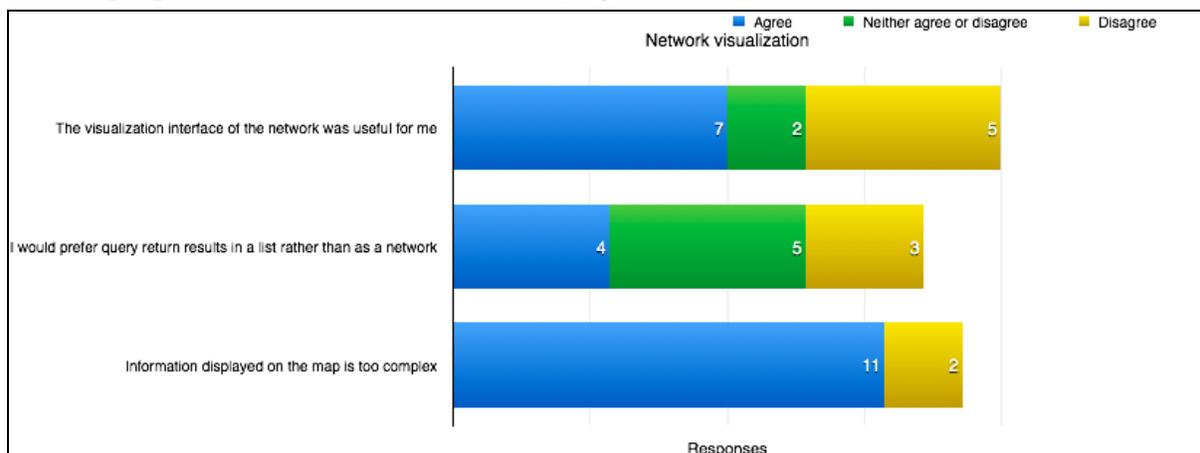


Figure 69: Survey results on questions related to network visualization interface.

Direct communication between people: Two respondents mentioned that direct contact between people without mediation of bosses or hierarchical structures as something that would be beneficial for the entire university organisation. “Who is the person I can contact I can see the organigram but you don’t know which person is better to contact. You can’t contact the super boss because his jobs is not to answer demands. So, direct contact of people”. When interviewing others about using Rhizi as a communication system, response was overall positive, especially the idea of subscribing to certain domains or fields of expertise and seeing requests for help based on tags rather than a direct contact with another person. As a next step, we would like to add the ability to contact respondents based on their fields of expertise or project involvement.

Knowing something is not the same as being motivated to use it to help others: Some respondents expressed a wish that only a subset of their skills will be public. “We all have previous lives, and probably you don’t want that people know. I don’t want that other people know that I did some shell scripting during my Phd, or understand how to build a network of TCP/IP. I’m not a computer guy. I like photography, I’d like to put that, because it can create a connection with other people who share this.” People are more motivated to help in a subset of the skills that they find interesting or relevant to their career trajectory. Other criteria for likelihood to help others included understanding the value of the help to the recipient, like knowing that an issue is urgent or blocking, or that the amount of work required isn’t too large. We think that mentioning skills people are not motivated to help with, will reduce overall trust in the system, as people will be slow or avoid responding to skills they are not motivated to help with. This should be reflected on the input and onboarding interface of the software.

What are people interested to know the most: Four people mentioned they are most interested in finding people who share their domain of expertise : “I would find it super nice if you could see who are doing the same kind of job and how to contact them and what they are working on so we could exchange information.” This supports that people are really interested in forming expert cohorts within organisations. This is further discussed in the discussion section of this paper.

Discussion

During this research project, we have discovered that expert identification has far-reaching consequences for the entire organisation: Bureaucratic bottlenecks, unnecessary project delays, and redundant solutions. We validated the problem, and validated that explicit and transparent mapping of people, projects and expertise can help even outsiders identify who can help on given issues and motivated people to discover and share expertise within their department. This improves access to the organisation’s knowledge capital across the entire employee-body. A separate finding is that while the network visualization was helpful for people to understand they were all part of an expert

network, it has also obfuscated some of the information because it was unnecessarily complex and forced casual users to adapt to an unfamiliar mode of looking at data.

These findings can further the field by acting as a proof of concept of how human computation can help with transition from collaboration practices of hierarchical organisations into a more flat and/or network-centric models. Related to known knowledge-management solutions, but not requiring the overhead of having to update a large repository of knowledge, as well as giving each person the ability to update his expertise as necessary. This could be of value to team science, because of the high throughput of humans and projects, and indeed we are planning to apply this solution not only for university IT, but for research communities as well.

Our main takeaway towards building the next generation of our collaboration tool is that we'd like to create a system to directly contact experts. This requires developing strong integration with email and other communication technologies such as real-time chat, video, and phone conferences. Another important issue is providing means to measure the impact of helping others and forming a reputation system that rewards those that unblock cross-departmental issues or save funds through mentorship and consulting.

Empowering people to directly approach experts within their organisation, can be part of the solution that both reduces the time each project takes while encouraging innovation and system-level thinking. More importantly, if university administration, governments or any large organisation want to be an attractive destination for innovative people, they must ensure that people can break down inter-team barriers and get the collaboration they need in a timely way.

IGEM

Background

The Igem (International Genetically Engineered Machine) competition is a yearly event in which over 200 teams of students from across the world compete to innovate in the field of synthetic biology. During the competition the students need to develop a project over the course of the summer, which is later assessed by experts, who rank it in accordance to various prize categories such as tracks (*e.g.*, Hardware, community lab) or special prizes such mathematical modelling, human-practices and public engagement.

Our objective in this pilot was to find out if usage of Rhizi in this context can help facilitate inter-team collaboration. The challenge was even larger, as the teams were in competition with each other. After conversing with the Igem team at the CRI, notably with Prateek Garg, we identified that there are unique values gained for teams that exchange expertise between them, notably, by exchanging experimental and computational methods, and when working on less common organisms such as fungi or mycobacterium. We agreed to work together to map the teams competing in Igem in order to facilitate expertise exchange between them.

One important challenge we faced, and a lesson learned from previous pilot projects, was making sure the CRI team itself saw the tangible value of using the software. The result was very apparent as the team detailed in their own [website](#)² the value they derived from Rhizi:

“Collaborations are an important essence of iGEM in particular and science in general. Teams in the past have shown excellent outcomes coming out of collaborations. We realized, however, that there is no user-friendly platform for iGEM teams to access information of other teams’ projects for a possible collaboration.” and “ Rhizi is a real-time collaboration tool whereby a team can share its project details... and common nodes between teams are automatically connected. For instance, a team working on *lactobacillus plantarum* would automatically be connected to all the other teams working on the same organism, and then can share vectors, promoters, protocols, and in general any advice with all these teams.”

Methods

The Rhizi team generated a [special instance](#) for them which included a domain-specific data scheme and a special landing page.

² http://2015.igem.org/Team:Paris_Bettencourt/Practices/Rhizi

They first created a [pilot map](#) that consisted of several closely related teams to test the mapping task and gather feedback.

The test teams were asked to put within the map their team name, the tracks in which they participated in the competition, the organisms they were working on, specific computational and wet-lab techniques they were using in their project and project keywords.

Following positive feedback from the test teams, the CRI team created the [actual map](#) which was the target for the mapping.

The Igem team created a [video tutorial](#) and emailed other teams in the Igem competition, asking them to input their information in Rhizi. 21 other teams voluntarily contributed to the database.

Data model

Together with the CRI Igem team, we defined the following node types to adjust to the use-case:

- Team
- Track
- Wet-lab technique
- Computational Technique/Tool
- Organism
- Project Keyword

Gathering feedback

Following the conclusion of the project, I did several video interviews with participants from the CRI Igem team to gather their feedback.

Results

By the deadline for submitting their details, 22 iGEM teams had shared their project features in Rhizi.

22 teams have successfully and voluntarily populated the maps. 70 entities of type keyword, 15 organisms, 14 tracks, 14 computational techniques, 31 wet-labs techniques were created.

Here is an example of a mapping made by the CSU Fort Collins team (figure 70):

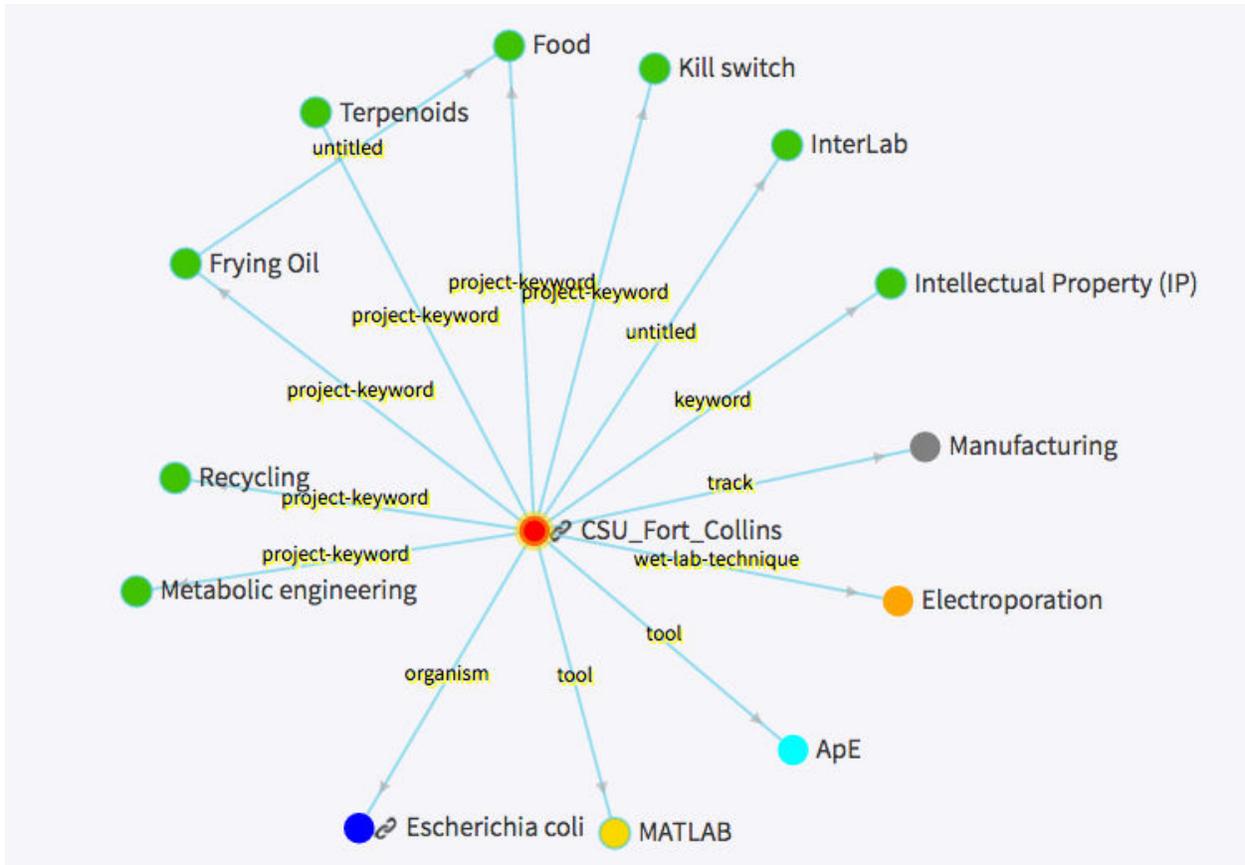


Figure 70: Team-centric view of CSU Fort Collins

When looking at this you can easily see they are working on the organism *Escherichia coli*, which could be useful for team who look for domain expertise about it. Knowledge of MATLAB, Electroporation and ApE might be useful for teams who need advice or get stuck and need to consult someone outside their core team. The list of keywords, give context regarding their project and can help others decide how to prioritize the teams to which they approach to seek help.

Clicking on the “*Saccharomyces cerevisiae*” reveals all the teams working on this organism:

“...it was cool, it was really cool. Not all the teams registered on it, because most of the teams are a bit selfish, but the teams that used it, it was very cool for us. Because we were able to use it a lot, we knew what was a project, what people were working on, we asked people for advice. It was a good experience.”

After interviewing several members of the Igem team, we found they had a very good experience, from one testimony from one of the CRI Igem students:

“So for example, the most basic stuff. If you have a protocol, but it’s not working. You don’t want to annoy your PI again and again, you can just ask other people from other team what they are doing, if for example you need to find the team who is working on yeast, and find most of the teams are working on e.coli. you need to scroll 2000 wikis, but if you can just have a map and click on yeast and find all the protocols with their names it’s way easier.”

This suggests that teams that collaborate between them have an advantage, this could produce the kind of effect that positively incentivises teams towards sharing.

Conclusion

This use-case shows a lot of potential to be generalized in other settings. Even in a competitive setting, scientific teams can help one another and increase the overall capacity of all participants to innovate and problem-solve. With more institutional support and accreditation this can be implemented in other competitions and increase overall innovation.

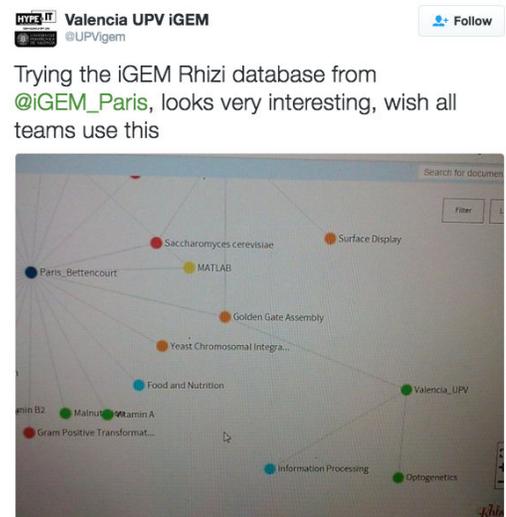


Figure 75: Rhizi Igem perspectives

Other pilot projects

Besides the pilot listed in the main segments of this chapter, there have been quite a few earlier pilots, prototypes and idea explorations we have conducted in collaboration with partners. These small experiments created a pool of potential use-cases that might be useful for others if they wish to pick up on it, provide insight how the project evolved over time, and demonstrate the scope and potential of the tools developed.

Open your Phd

As part of the [NightScience hackathon](#) organised by the CRI, I worked with participants and other Phd's to share our research by using Knownodes. By working together in the same space we turn the act of sharing our research into a collaborative and interdisciplinary experience.

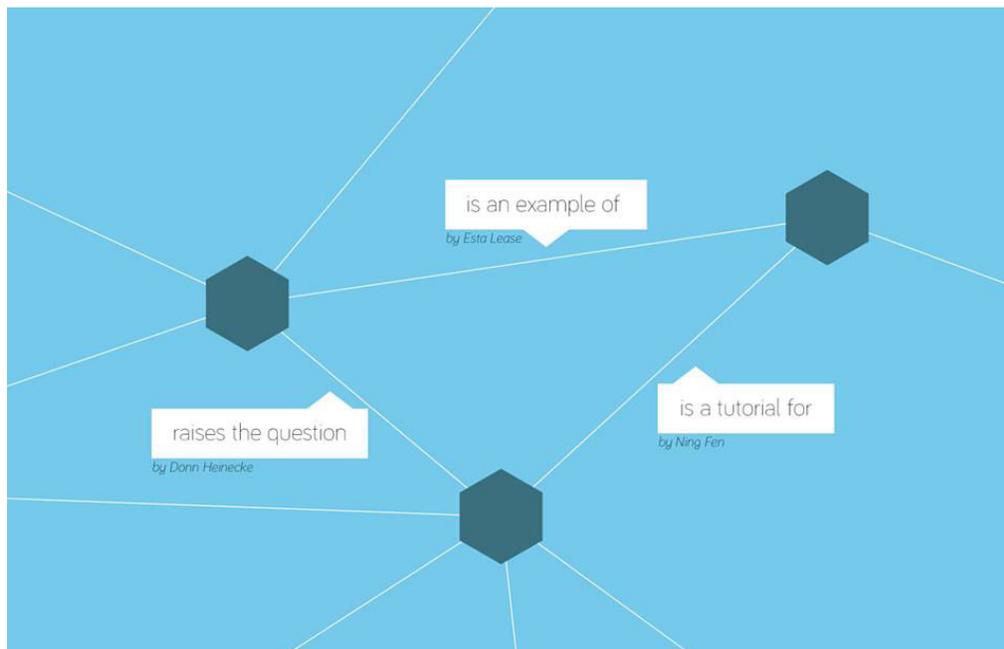


Figure 76: Knownodes concept example.

Our goal: Each participant will create an open network of his Phd research. To do that each one of us will alternated between creating his own network to giving feedback to someone else.

The week following the hackathon, a bunch of us (currently about nine people) are going to the [Performing Arts Forum](#), a former monastery in the countryside, for four days to apply what we have learned in code and design.

Results and analysis

During the event itself, it was evident people had a hard time wrapping their heads around the complexity of inputting triplets and navigating the interface. While a lot of interesting discussions were generated and useful input about UX was collected. Issues observed in the Bibsyn pilot repeated themselves. Most notably, there wasn't enough motivation for people to dedicate much time to share their knowledge.

The networked lab

An idea we rapidly mocked was to create a visual graph to help labs present their work, and for outsiders to navigate the relationships between people, projects and concepts the lab was working on:

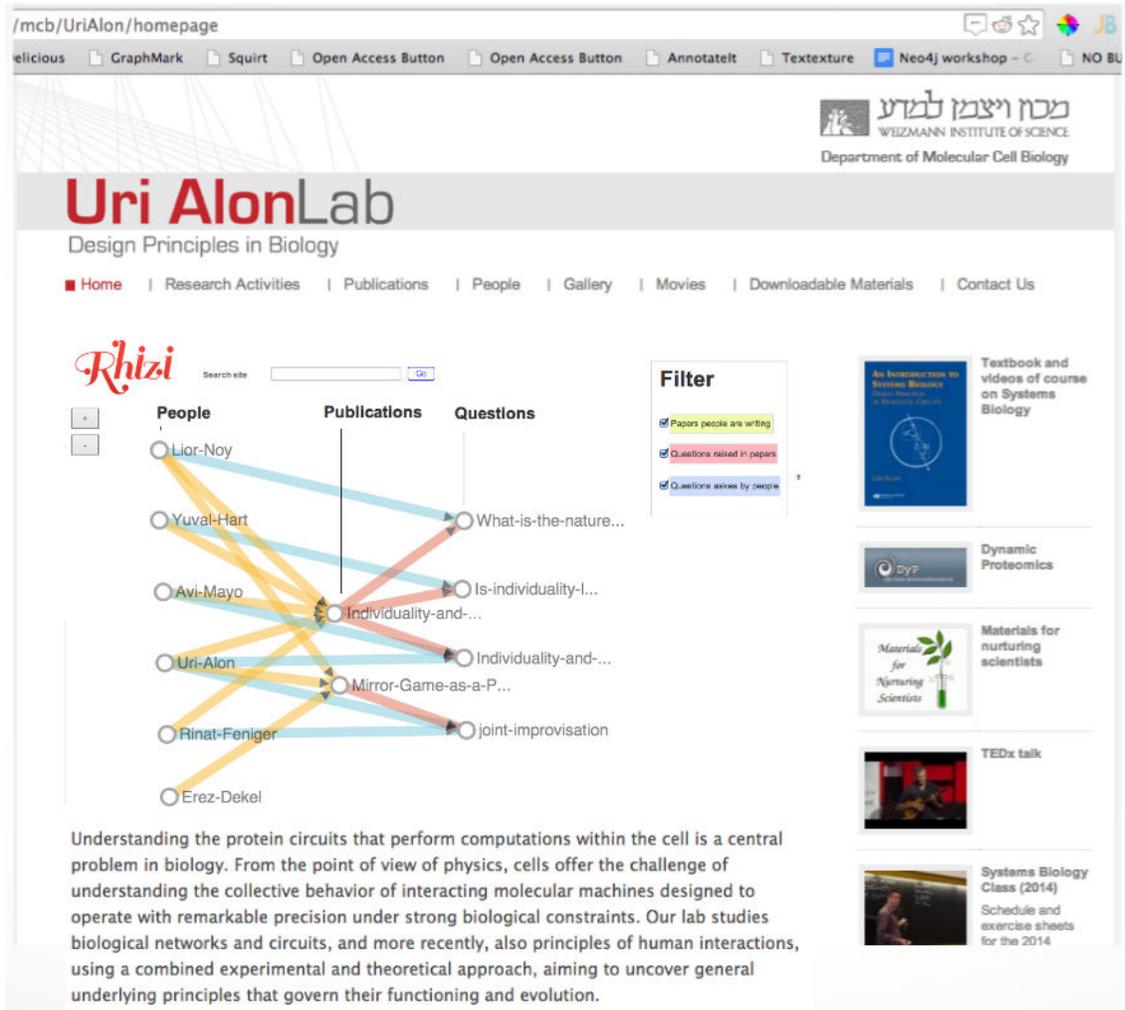


Figure 77: Mockup of the final result.

The concept for this was validated by the Computer science department in Imperial college who wanted to hire us to build them such a system. Unfortunately, we didn't want to defocus work on our main platform and didn't take the offer.

Nanopublications

Another idea we worked on and applied for an ANR(A call for projects in france) was to use Knownodes as a publishing tool for nano-publications. Attached is the abstract of an ANR proposal we have made together with David Bikard for an open Crispr system. Attached is the part related to the integration of Knownodes within that. Most of the mock-ups work is a result of conversations with David

Nanopublication of genetic interaction data

We will **develop a simple and effective interface to enable individuals to collaboratively build and explore crowd-sourced knowledge networks of scientific data and assertions.** This will enable anyone to instantly share their results using a recently developed standard known as *Nanopublication*. A nanopublication is the smallest unit of publishable information: an assertion about anything that can be uniquely identified and attributed to its author. A nanopublication has three basic elements (the *triplet*):

A unique identifier (IRI) and an integrity key.

The Assertion: a minimal unit of thought, expressing a relationship between two concepts (called the Subject and the Object) using a third concept (called the Predicate).

The Provenance: the metadata providing some context about the assertion. This includes Supporting metadata (like methods) and Attribution metadata (such as authors, institutions, time-stamps, grants, links to DOIs, URLs).

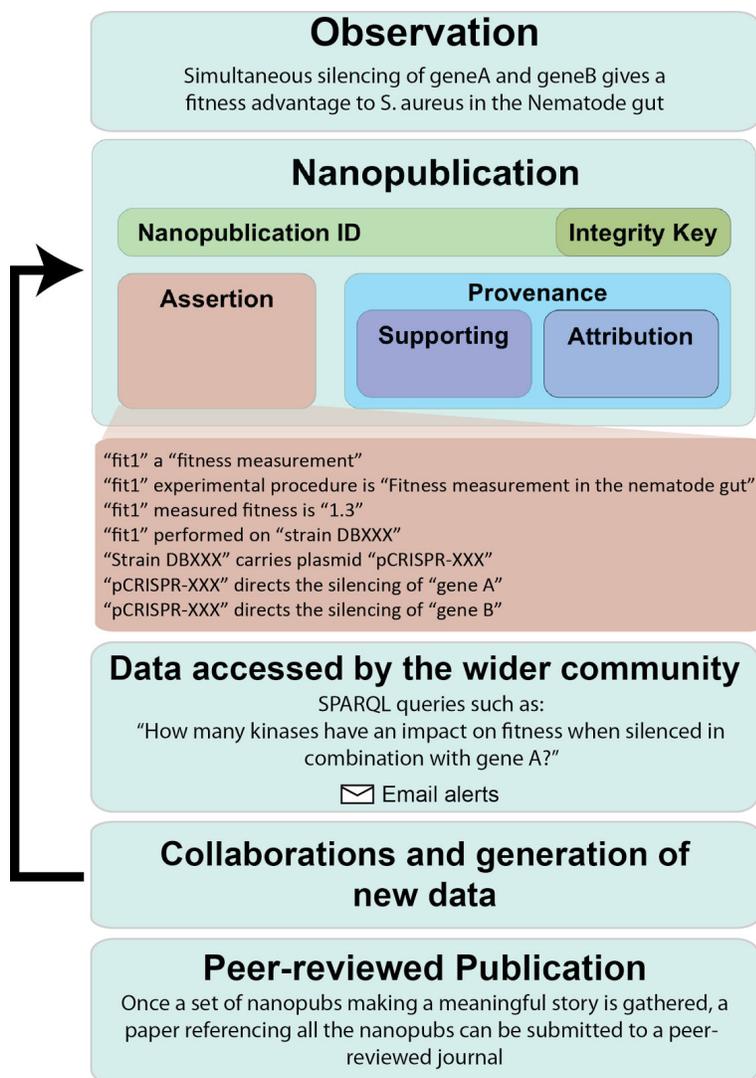


Figure 78: Data schema for nano-publications

Each nanopublication will expand the graph of facts present on the platform and will instantly be connected to the existing knowledge on the topic in a way that will enable content discovery. We will first start by building a platform to publish our genetic interaction data, as well as other knowledge gathered during the course of this project. Here is a list of the characteristics planned for our platform to guide our design efforts:

- *Author content* : Users will be able to add new factual information. This will take the form of adding a new triplet to the knowledge graph. For instance, adding a link between gene A and gene B to specify that there is a synthetic lethal interaction between the two genes.

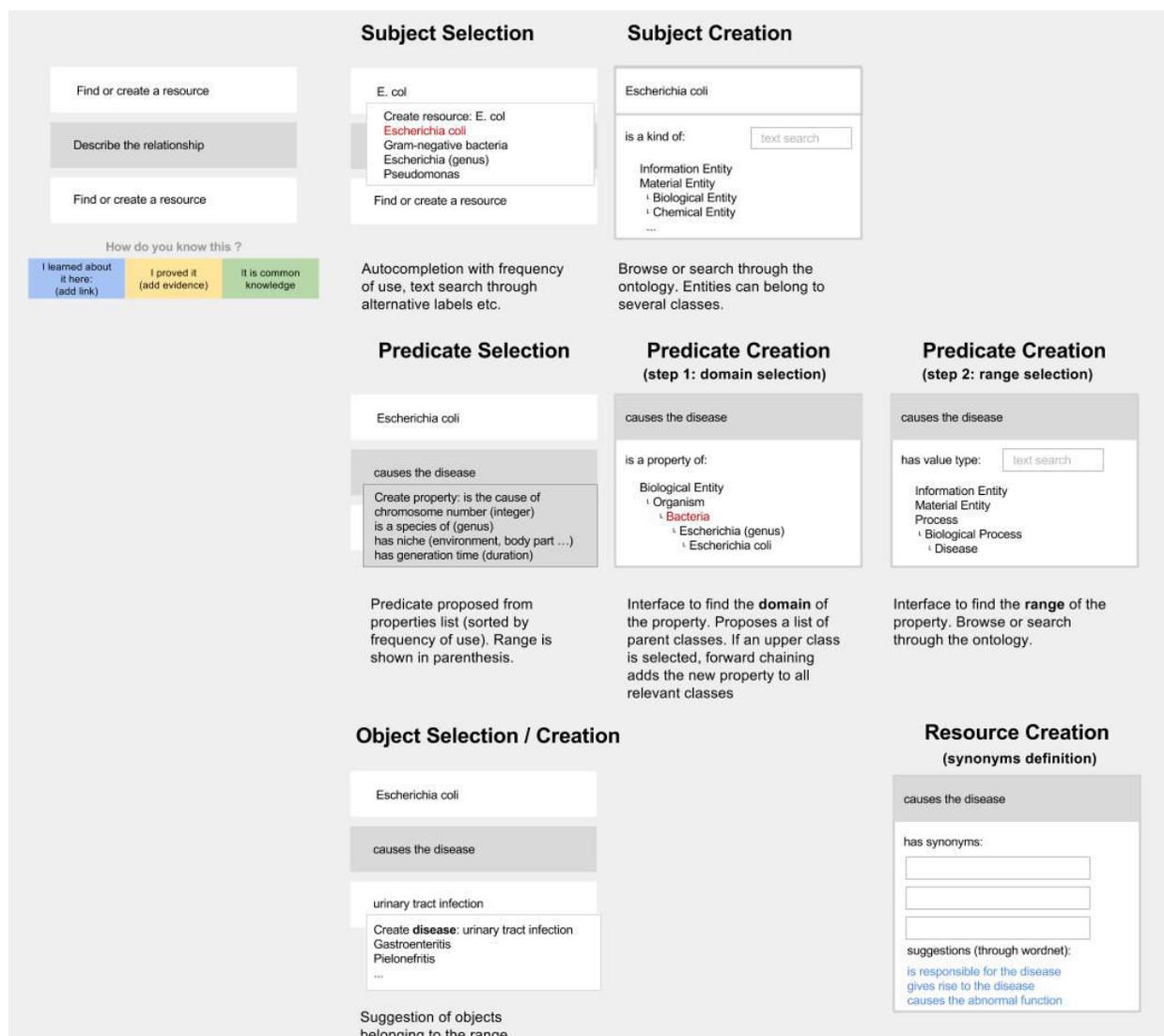


Figure x: Mock-up of input interface for Nanopub platform

- *Associated provenance of content:* (i) the author will be able to upload the raw data, experimental conditions, figures and methodology that support his specific assertion. This is where for instance we will be able to learn that the synthetic lethal interaction between A and B was observed during growth under high osmotic pressure. (ii) the provenance includes information about the author himself, who actually performed the experiment, what their affiliation is etc.
- *Discover content (search):* Searching content can be achieved by keywords, gene or protein names, accession numbers, type of relationships or even patterns in the fact graph. Content discovery can also be passive in the form of email alerts or notifications of new content from a personalized set of nodes, triplets, keywords etc.
- *Explore content (browse):* Each piece of information, each node and each edge in the network is a link to more information and will enable users to navigate through the graph. Different UI

can be built on top of the knowledge graph to answer different needs. For instance one might sometimes want to actually see the data in the form of a graph with nodes and links, but in many cases this will not be the preferred representation. Dynamic data might require a representation along a time axis, other data might require graphs, maps, scatter plots, animations, or just simple lists of properties. We will not be able to answer all these needs by ourselves but will enable users to contribute their own visualizations and interfaces to explore content in the most efficient way. A few interfaces will be developed in collaboration with the INSERM U1001-lead Rhizi community, a group of talented computer scientists developing open source interfaces for the manipulation of linked data (Rhizi.org).

The mock-up interface for Nanopub displays information for **Escherichia coli**. It includes a microscopic image of the bacterium, a taxonomic classification tree (Biological Entity > Organism > Bacteria > Escherichia (genus) > Escherichia coli), a Wikipedia URL, and a descriptive paragraph. A table lists subclasses and key properties:

Subclasses of Escherichia coli	
Escherichia coli str. K-12 substr. MG1655	
Escherichia coli O157:H7 str. Sakai	
Escherichia coli IA139	
Escherichia coli UMN026	
Escherichia coli O83:H1 str. NRG 857C	
Escherichia coli O104:H4 str. 2011C-3493	
see more	

Causes disease	Urinary tract infection, Gastroenteritis	+
Niche	Human intestine	+
Chromosome number	1	+
add a property	Find or create resource	

Figure 80: Mock-up of input interface for Nanopub platform

- Narrow content (query):** The ideal query system is one that would understand natural language. Until we have a true artificial intelligence this is unfortunately not possible. In the meantime we will use the powerful SPARQL query language to narrow content (ref). For instance, queries like: “give me all the transporters that interact with genes essential for colonization of the mouse skin” will be possible. We will design a user interface that will allow making this type of query intuitively.

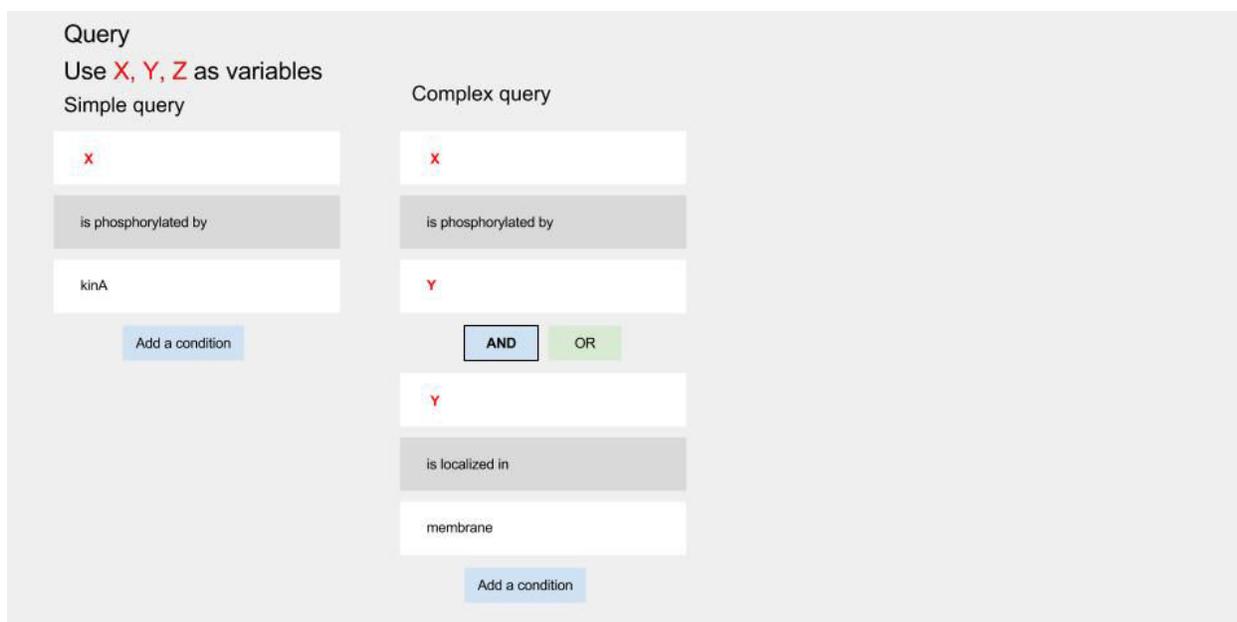


Figure 81: Query interface for Nanopub platform

- *Retrieve content (download)*: The SPARQL query language can be used to download data in the form of a RDF graph. An interface that allows making SPARQL queries in a user friendly way will therefore also be useful to retrieve content.
- *Link to content (reference)*: An important aspect of the platform is that each assertion should be easily citable. This can be achieved thanks to the unique identifier present in the nanopublication standard. Note that referencing can go both ways. Users will be able to reference nanopublication from outside the platform, and nanopublication can reference any type of content through their provenance graph, including traditional publications as well as other nanopublications.
- *Evaluate content quality*: An important caveat of an open sharing platform is how to maintain a good data quality. Several mechanisms can be implemented to ensure that data can be trusted. We will experiment in order to find which combination of mechanism provides the best outcome. For instance one can compute a confidence score for each assertion in the graph. Things that could positively influence this score include whether the assertion is supported by several independent provenances, whether a reputable user authored it, whether it is supported by a peer-reviewed publication, or whether reputable users have confirmed its quality (post-publication peer review). Author reputation itself can be computed according to his or her contribution to the graph of knowledge following mechanisms similar to what has been shown to be successful for platforms such as stackoverflow.com.

Mozilla and Open science workshops

This idea was prototyped during the Mozilla festival and later in a SPARC convention in Berlin.

The p2p science alliance is a way to make enhance collaboration by making them “open access”, meaning, make an explicit description of collaboration around a tool or research project. When we openly share our collaborations and intentions we create a basic framework to work with others.

We asked participants in the workshop to list themselves, their collaborators and projects as well create connections between people that they thought should collaborate. Within the connection description, there was space to specify why these people should collaborate.

Results

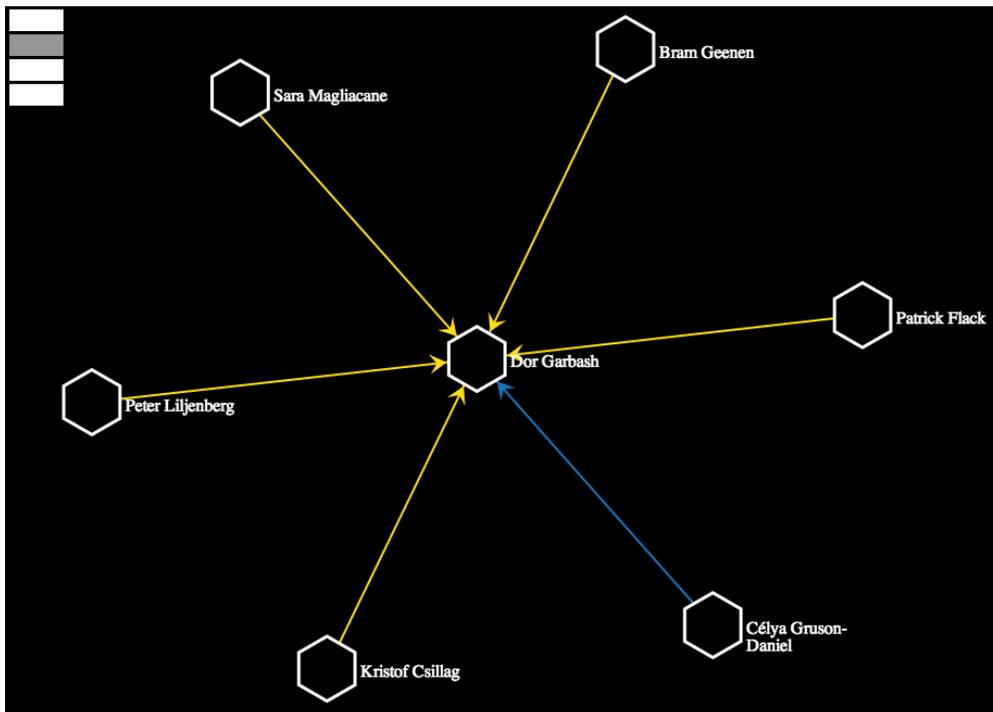


Figure 82: Example of a map interface containing existing collaboration (in blue) and potential collaboration ideas (in orange)

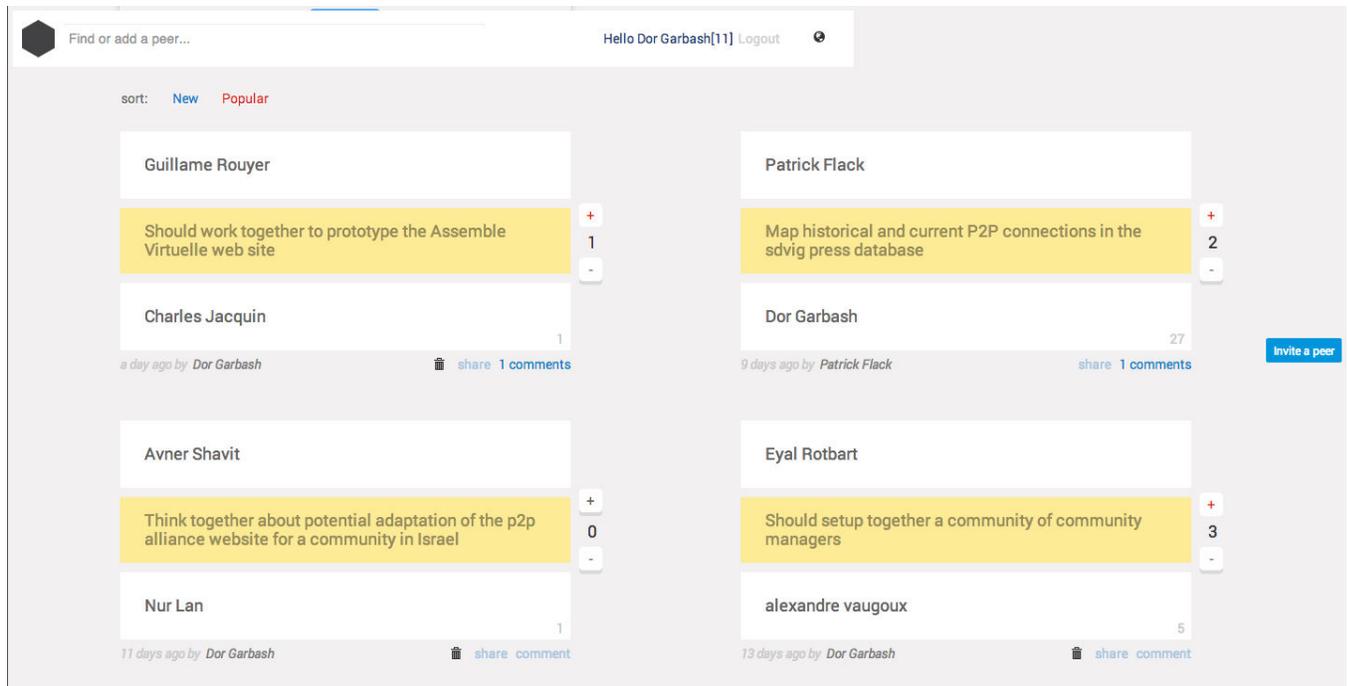


Figure 83: A list view of the most recent connections created by the system.

This was an early test of what later became Rhizi and the concept of mapping project, skills and interests to facilitate collaboration in organisations. Overall, people were excited about the visual map component that we've just added to Knownodes as well as the overall value proposal and were glad to participate.

Rhizi and Knownodes for web annotations and MOOC's

Knownodes—The MOOC companion tool for open student participation.

Despite their immense potential, MOOCs are hindered by high dropout rates, low user participation and difficulty in dealing with the variance in student competence and knowledge.

Without a teacher to closely interact with students, tools of participation, peer-review, feedback and reputation system are necessary to make the course evolvable and participatory. Our proposed tool will enhance web literacy and the collaborative skills of participants, and make user generated content interoperable between courses and platforms.

Knownodes is a simple browser extension and website that gives students the ability to connect any part of any MOOC(e.g., lesson, part of a lesson, question or task) to any other web resource or insight they have, in diverse media forms(video, image, hyperlink or text).

Being a companion tool, it integrates easily into existing MOOC technology and practices, with no need for any technological adaptation of the MOOC platform itself.

“This video reminds me of this quote”

“I have a question about this part” -
“This is the answer to your question is here”

“This is a tutorial to illustrate what the Professor said”

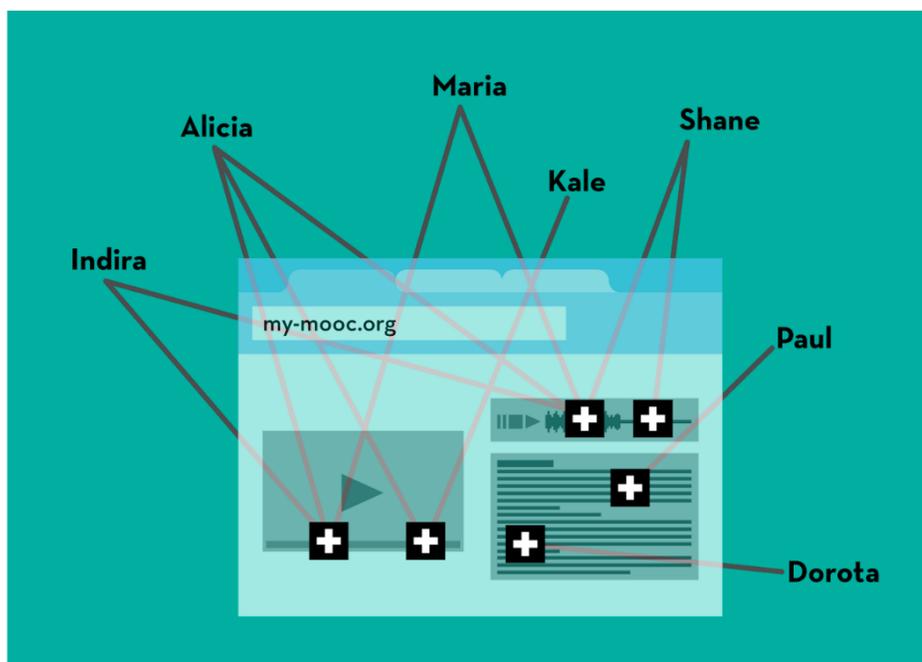


Figure 84: Illustration of usage of Knownodes within a MOOC

When a connection in Knownodes is created, it becomes an interactive entity that can be commented, peer-reviewed and shared. The connections created by the students are quickly aggregated to fellow students who can peer-review and give feedback to the contributor. Each time a connection is created, students are asked to provide a jargon-free and specific sentence that explains the connection, this adds invaluable context to both the contribution and the MOOC materials that are being connected.

The course content as seen in Knownodes starts out as simple network of class materials as created by the course facilitator and automatically generated Wikipedia content. This base can be extended and evolved by students following course tasks and free participation.

This content forms a browsable knowledge web that can be customized to display particularly relevant content to each student.

Beginner students would be instructed to enrich the knowledge graph by asking questions about the material, and by following a quizzing module that requires them to describe relationships between different materials taught in the course.

Advanced students would be challenged to connect additional material such as tutorials, wikipedia articles and additional learning materials and ideas which relate to course material.

All connections can be presented to other students from more advanced classes. By providing feedback on the younger students work, the older students benefit from learning by teaching, save the teacher's time and create a virtuous circle where students constantly teach and learn from each other.

Ultimately, a scalable, highly connected and curated knowledge graph is generated. The context given in the connection assures that material is reusable to future iterations of the course. All data generated in Knownodes is open and compliant to the W3C open annotation standards. Having crowd-sourced data on connections between knowledge and web resources would position France as a world leader in open annotation and open data technology, giving competitive advantage to french companies and organizations working in the open data eco-system.

This plan was co-constructed with representatives from HarvardX and presented in a conference in Malaga about innovations in MOOCs as well as Iannotate 2014 - a conference in San-Francisco about web-annotations.

As part of this we have prototyped a simple way to capture annotations in a Neo4j database:
<https://github.com/alon/annotator-store-neo4j>

The demo of the software can be viewed here:
<https://www.youtube.com/watch?v=yuMs9WRnSKc>

Lessons learned

After considerable efforts from our parts, the project didn't achieve maturity because HarvardX pulled away because of unrelated issues. Because annotations themselves are a bit of an experimental technology in themselves, we made the choice of focusing on our core issue, which is mapping human capital using Rhizi and not devote further resources into this channel.

Leadership workshop

During a leadership workshop conducted at the CRI, several volunteers created real-time graphs of concepts, ideas, people and projects people were mentioning during the talks. We tried to engage the workshop participants to participate in the mapping activity, but for the most part, the attempt did not succeed.

The main problems encountered during the workshop were: 1) Adding information during a real-time time event is task that requires experience with working with Rhizi interface, and more importantly, to understand how to construct phrases in such a way that they would build up a graph. It takes significant cognitive work by humans to parse the information conveyed in a regular

presentation or discussion into a graph structure, making these kind judgements in real-time proved to be too challenging the context of a busy conference happening in parallel 2) People preferred using Hackpad, a well established collaborative word-processing software as it was more intuitive and better fitted the workshop's note-taking work.

Rhizi staff skill graph

The CRI wants to establish links between CRI's employees (~40 ppl) for problem solving and coordination of training.

The objective was to change the current state where (i) Employees often don't know who to ask for help when having a software question or require an advice from an expert. (ii) Employees cannot organise training requests by themselves or as cross-organisational groups, therefore preventing people to pool requests for training sessions.

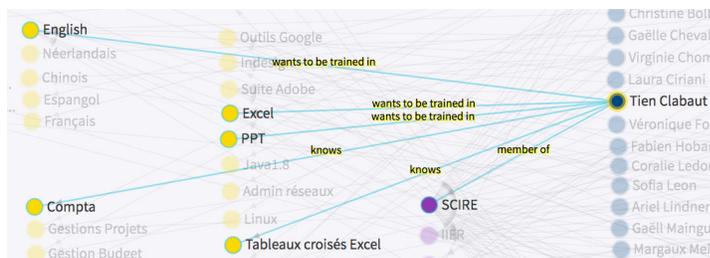
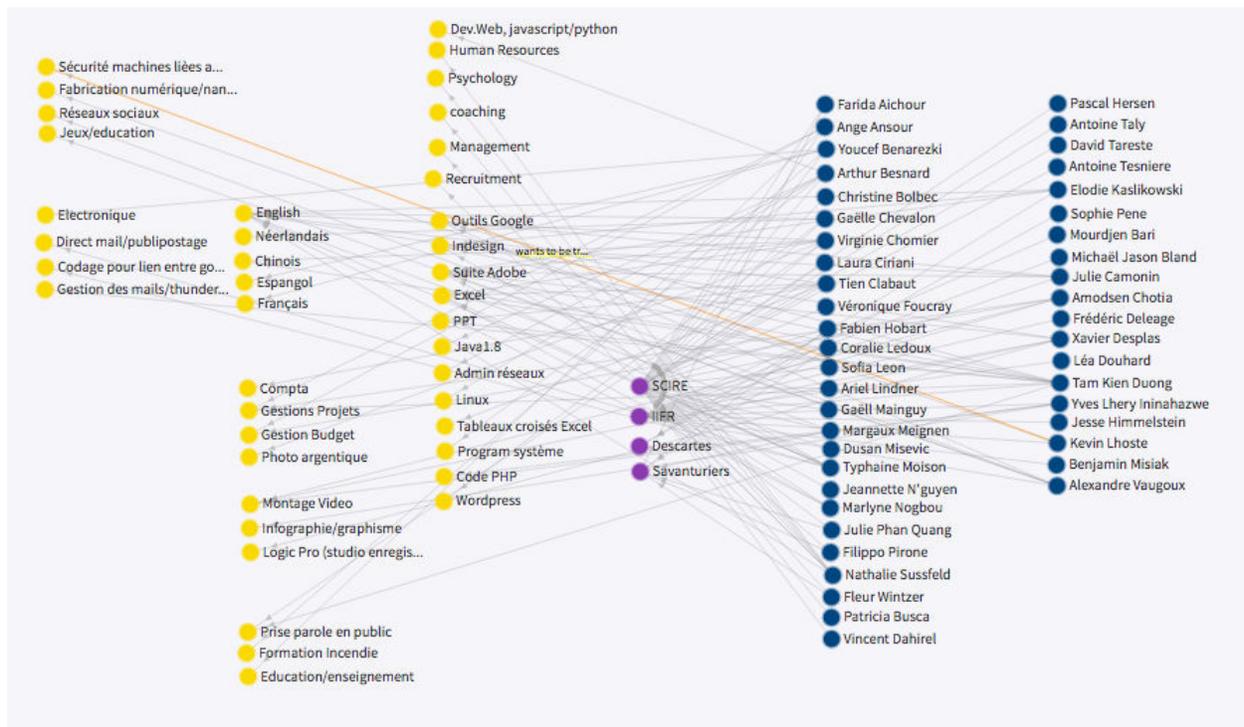


Figure 84: (top) entire CRI staff skill-graph. (right) if focusing on one person (left) When focusing on specific training session.

An initial map was created based on a csv file prepared by the HR manager. A person seeking assistance can go to Rhizi, search for a particular skill and engage directly with a colleague, or add desired training sessions or skills.

The pilot was not completed due to timing issues, but has potential because of the relative success that the Descartes IT pilot that addressed a similar set of problems.

Conclusion: Met and unmet challenges analysis

“It’s already doing a little like what is happening in your head when you speak or think about someone of something, it’s the kind of web we have in our head when we speak or think about someone, but here it’s more concrete and it’s not limited to what you know, but what everyone in the connection know. It’s pretty easy to understand how it works and how you get information from it.”

- CRI student when interviewed about his experience using Rhizi.

This chapter analyses how the challenges set in segment [1.4 Challenges summary](#) were addressed during this thesis. While I was “shooting for the stars”, and did not completely address the full set of the challenges, I do offer in chapter [4.1 Proposal for a future project](#) a complementary future project that could be expanded into a good plan of attack to tackle the challenges left unanswered.

Social challenges encountered

How can we help experts share their expertise despite their cognitive biases?

When asking an expert to document his knowledge, the cognitive biases are left mostly unaddressed because there is no feedback loop that lets the experts understand what steps in his explanation are missing. The work done with Rhizi shows that human-capital mapping can help facilitate a conversation with an expert to provide two types of context: (i) Where are there gaps in the knowledge between the expert and help-seeker (ii) What is the specific task that requires help from the expert.

Our work addresses the issue of ignorance to human capital within the organisation. Our unique contribution showing the gaps in expertise awareness before and after Rhizi in the Descartes IT pilot ([3.5 Descartes IT](#)) responded to that challenge. There is still work left to best determine how to keep people engaged in the system over multiple years and keep it up-to-date. One possible solution involves coupling expertise mapping with issue tracking. This is elaborated in the future work proposal.

How can we provide motivation for organisation members to take part in knowledge sharing activities? What incentives can we provide? What barriers can we remove?

This was another focal point in the thesis. We have managed to address these challenges fully with the Igem pilot and partially with the Descartes IT and CRI pilots. With Igem we have seen that the ability to quickly solve concrete issues related to experimental and computational techniques provided sufficient motivation for teams to participate in the mapping project. With Descartes IT

and CRI, we had partial success, which also provided a key learning for future implementations. We managed to motivate administrators, management and particularly innovative students and employees because it helped them better understand their own organisation. The visualization module also was key in attracting people to several voluntary projects and maps specified in segment [3.4 CRI - Opportunity mapping](#). However, within these two pilots we have not achieved full buy-in from the rank-and-file members.

Addressing this challenge of providing both management and rank-and-file with motivation to engage in human capital mapping has the potential to provide huge value to organisations: Logging and acknowledging collaboration work outside role definition could become a framework that enables a truly networked organisation where those who get the job done can coordinate their collaboration in a decentralized way with a self-sense of autonomy and growth. In future work proposed at segment 4.1 will briefly describe how this can be done through a simple software solution.

Another important aspect of Igem and Descartes IT use-cases is that both included a relatively small group of experts that understood how the activity helps to solve immediate and important tasks. Even within the successful pilots there is a clear distinction between those who reported getting concrete value through discovering a collaborator or internship opportunity went on to create their own map, start a Rhizi related projects, complete information for others in a Rhizi map and other “super-user” type of activities. Those who might have connected with the project’s concept only in theory without direct experience of the benefits it offered, have done only the minimal amount of effort the tasks given required.

With regards to removing barriers, the main technical and design work we did dramatically improved the graph input interface. We have lowered the effort for inputting information into a knowledge graph in such a way that non-experts learned how to do it within a few minutes (and more often seconds). This involved two major innovations: (i) Text-to-graph interface turning complex form filling into an intuitive text interface (ii) Providing real-time feedback how the text forms a graph and how it connects with existing knowledge within the system. The combination of the two created a pleasing and unique user experience that beyond removing barriers, motivated people to adapt the system to their own use-cases, e.g. what turned into the Igem pilot, the Immunology club map and many other maps that were created without encouragement from the Rhizi team.

How can we prevent the “norm of reciprocity” from inhibiting people to seek help? How can computer-mediated systems help people feel more comfortable seeking help?

When we eliminate the need to ask multiple people for help, just in order to get information who is the right expert, we reduce some of the psychological barrier for “norm of reciprocity” from requesting assistance from the intermediaries. We also reduce the amount of people we engage in

our need, thereby increasing our sense of comfort. The Rhizi map generated in the Descartes IT pilot allows people to directly seek experts, and for the CRI use-case, it allowed students to discretely contact and inquire with their peers regarding internship opportunities, and scientific techniques. As one of the CRI students testified: “You don’t want to annoy your PI again and again, you can just ask other people from other teams what they are doing”.

How can we avoid formal form filling and following rigid sets of rules that make knowledge sharing a lot less satisfying than informal ways?

How do we reduce knowledge retrieval within organisations? How do we deal with dynamic change in data, objectives, units and high price for centralization and standardization of knowledge storage and retrieval?

Instead of mapping knowledge directly, Rhizi helps map the sources of expertise within the organisation producing a *commons for sources of knowledge*. By putting human capital at the center of the mapping process, we facilitated direct and informal human contact. This also addresses the “tasks that are too broad” challenge of collaboration. By empowering members to discover and contact experts, a task which seems daunting can be well-framed while conversing with an expert.

With the work in Knownodes we have seen that categorizing connections and complex form filling was a major inhibitor to sharing the knowledge in the system. Allowing for a more free form expression using text-to-graph interface, where the connection types were not fixed to a pre-set list proved to be the right balance between structure and fluidity of experience. This is evidenced with Rhizi, where we had multiple cases of people using it “just for fun”.

How can we decouple collaboration activities that can be done in a networked way from hierarchical organisational structures?

Hierarchical structures often manifest in the form of formal meetings and work spaces. Allowing people to know about the competencies, projects and interests of people beyond these spaces also allows the decoupling of collaboration. This was seen in Igem when teams collaborated through discovering expertise in the Rhizi network, where otherwise they would go to teammates or experts within their local work places for help.

How to address Reluctance and sanctions against bottom-up information and norm sharing?

Removing intermediary brokers to connect with experts is a useful benefit that Rhizi enables. In the future project proposal, I also discuss additional ways to address this.

How can we overcome the downside of Preplanned, static division of labor?

Beyond a person’s skill set that is adapted to his current role, there are many hidden skills and interests that have been exposed through engaging in human capital mapping. The ability for others to discover these interests and skills can help engage people in work or learning that is beneficial to the organisation as a whole but goes beyond the person’s formal role.

How to reduce the high barrier to entry for collaboration? How to achieve conversational critical mass? What alternatives can be provided for organisational cultural resistance or resistance of hierarchically minded, bureaucratic, and centralized minded managers to give up control?

Forcing an organisation to achieve a goal that it does not wish to meet is nearly impossible. However, if there are specific units or managers that do not wish to participate partial participation can work as well. It was easily demonstrated in the Igem project that even when just ~10% of the teams were involved, the benefit was large enough to provide value for other teams. This fact was also stated by members of the Descartes IT. “When I started my project, I looked for people who knew [web technology]. That could help me to start. I did it by myself. Now, I look first in Rhizi.”. This did not require a buy-in from the entire institution.

How do we deal with lack of knowledge structure? How do we define ontologies for organisational awareness?

From the work done, we discovered the fine balance between allowing flexibility, which worked great for us as a discovery framework, and the unnecessary complexity that comes from having too much customisation. The general conclusion is that *less is more* when dealing in a social computing context. Ontologies must be minimally defined to achieve focused goals. The ultimate aim is to achieve immediate value and buy-in from multiple groups within the organisation and expanding from there to accommodate additional use cases and provide more value.

How do we deal with knowledge fragmentation?

While there is definitely potential for future work to adapt Rhizi to map non-human sources of knowledge such as documents and databases. With the Descartes use-case we managed to expose the entire department’s involvement with projects, technical skills and therefore managed to resolve the human-capital awareness fragmentation challenge. Using the yearly employee feedback Descartes IT does with their employees to gather data proved to be a very valuable practice for achieving that, as well as usage of the CSV import function within the CRI.

How can we guarantee a shared organisational awareness state? How to deal with information overflow?

Information overflow was a major issue in the Descart IT use-case, when maps grew to a large size. To address this I propose to provide a list-based view as shown in the future project segment. To guarantee a shared awareness it is important to really understand what information would be beneficial for employees to receive as well as to get. From interviews with CRI students and Descartes IT members, it was clear that the main type of information that would fit **both** of these requirements would be specific competencies that employees are: (i) interested to learn (ii) see the skill as something they would like to signal to other as part of their career path (iii) skills they are passionate about but lay outside the scope of their fixed role (iv) a skill they know would directly help another person in a meaningful way.

The reason why these specific types of skills should be collected is because it is important to optimise for the practical value reported by the super-users - to find learning or work opportunities, joining shared interest groups such as communities of practice (e.g. Python programming language fans or student clubs), and most prominently solving concrete challenges encountered during day to day work.

Another important conclusion that I have reached from the research is that network visualization while novel, useful and produced superior results to the list based Knownodes, did not, in itself, provide sufficient value to the casual user compared to the amount of work it took to build and solve the myriad network visualization problems associated to it. While the potential to have real-time organisational social analysis provides a real value to administrators and managers it should be reserved as a secondary feature behind focusing on well-established interfaces such as news feeds, cards and list based results.

Unmet challenges

One strategy we intended to test, but haven't received a sufficient critical mass to explore is how people could use the collaboration features to produce a 'collectively intelligent' Knowledge graph where people help in their activities to remove redundancies, correct mistakes, and customize the view. I think that the ability for a crowd of people to move information objects around in a two dimensional space and then observe how positioning can indicate relatedness of entities can be an interesting way to learn how the wisdom of the crowd can shape visualization of information.

Technical challenges encountered

Modeling the data

Within the scope of this exploratory thesis, I needed to adapt to the changing needs of stakeholders, and the different scenarios they wanted the software to accommodate for. In addition, we have been constantly improving on our own value proposition and understanding of the problem space. This has forced us to deploy significant efforts in building a highly customisable system, which I would not recommend for less-exploratory projects.

Access-control

The need to create content that should be available for a specific team or that is not transparent to outsiders requires development of an access control system. We did not manage to develop these features with the resources we had available. We did create a simple white-list of emails and email domains as a form of access-control. This feature not address, however, the need to limit access to specific maps.

Here are future specification that could be useful to this end:

- Offer a default permission menu for the system that will define the following functionalities:
 - Review list of users
 - Defining super users (people who edit the document, delete everything regardless of permissions)
 - Define default permissions on a document when it is created
 - Write permission (list of users)
 - Whether the document is public or private
- Track the creator of a document
- Offer a menu to change permissions once the document has been created
 - Write Permission (list of users)
 - Change visibility of the document (private/public setting)
 - Lists of users that can access the document

Collaborative features

Collaboration features such as commenting, voting, reputation systems, notification system, editing options and history are important to help the process become more of an engaging conversation rather than a one-off data input task.

Network visualization

There is much work on network visualization that remains to be done. Mainly the ability to create semantic-zoom to view large maps. That is the ability to show relevant information in accordance to the scale of the map.

Integration with existing tools

One option to ensure data remains up-to-date is to provide a bi-directional integration with existing knowledge and project management applications. Integration with ticketing systems, email and project management software such as Slack or Trello could provide great value as a secondary step.

Summary

The main achievement of this thesis is that it provides a social proof-of-concept of the role human capital mapping can serve within organisations. This has a truly a transformative potential that would allow workplaces in the future to become more flexible in the collaboration opportunities they offer to their members.

The other important impact of this thesis was that it was a prototype, at least within the french academic system, of a “Lean Phd” incorporating strategies used by startups to accelerate discovery

of social computing applications. I hope that the work will help others learn and adapt when applying these principles within their own projects.

Proposal for a future project

As a final synthesis following the conclusion chapter, I attach here a proposal for simple application that would address some of the unmet challenges set before, and specifically the following challenges:

How can we provide collaboration systems with measurements outside scope of defined role or task?

How to lessen the psychological barriers of seeking help from others?

How to bridge the time transition cost of moving from a self-serving to collaborative behavioural model?

How to rank experts? Can an expert be objectively assessed? How can the information remain up-to-date?

How to overcome the “Knowledge is power” perception that inhibits employees from sharing knowledge?

Following the exploration process I followed, the best way to promote organisational awareness is to couple it with directly helping organisation members to solve work related issues. When a member encounters a problem he can't resolve by himself, or s/he wishes to improve the quality of his work by consulting an expert, he should be able to identify the people in the organisation who are most likely to help him and make direct contact with them or with minimal dependencies on middlemen. This eliminates friction and avoids some of the perceived risks and disadvantages of asking for help.

The application I suggest to build channels requests from organisation members to the peers most likely to help them. Collaboration is measured based on feedback by the person who opened the ticket, this creates an incentive system that provides status and rewards for those who share and collaborate.

The interface itself is extremely simple and consists of (i) An input interface for issuing tickets (ii) A personal stream of tickets, sorted according to skills and projects listed in the user's profile. (iii) A dashboard presenting top contributors according to the skills and project keywords they have

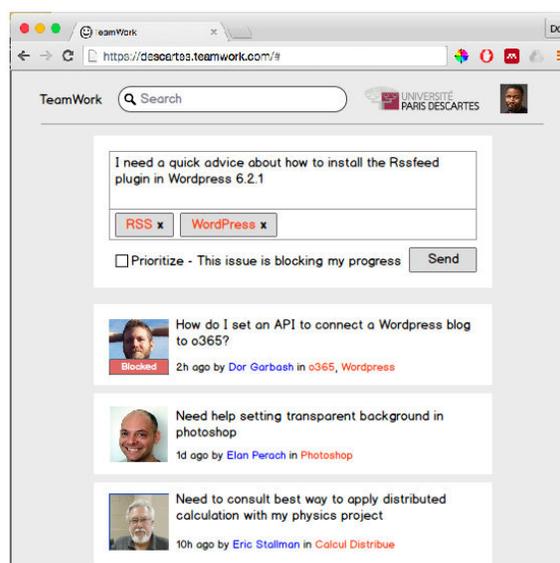


Figure 85: Future project proposal mockup

Bibliography

- Abrams, D. & Hogg, M. (Eds.) (1990). *Social Identity Theory: Constructive and Critical Advances*. New York: Springer-Verlag.
- Abrams, L. C., Cross, R., Lesser, E., & Levin, D. Z. (2003). Nurturing interpersonal trust in knowledge-sharing networks. *The Academy of Management Executive*, 17 (4), 64–77.
- Ackerman, M. S. (2000). The intellectual challenge of CSCW: the gap between social requirements and technical feasibility. *Human-computer interaction*, 15(2), 179-203.
- Adelson, B. (1984). When novices surpass experts: The difficulty of a task may increase with expertise. *Journal of Experimental Psychology: Learning, Memory, & Cognition* 10: 483- 495.
- Agarwal, N. K., Xu, Y., & Poo, D. C. C. (2011). A context-based investigation into source use by information seekers. *Journal of the American Society for Information Science and Technology*, 62(6), 1087–1104.
- Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50(2), 179–211.
- Alavi, M., & Leidner, D. E. (2001). Review: Knowledge management and knowledge management systems: Conceptual foundations and research issues. *MIS quarterly*, 107-136.
- American Productivity and Quality Center and Andersen, A. (1995), “ The knowledge management Assessment Tool(KMAT) ”, available at : <http://www.apqc.org/km/>.
- Anantamula.V.S. (2008),“ Leadership role in making effective use of KM”, *journal of information and knowledge management systems*, Vol 38 No 4.
- Arthur, J. B., & Huntley, C. L. (2005). Ramping up the organizational learning curve: Assessing the impact of deliberate learning on organizational performance under gainsharing. *Academy of Management Journal*, 48(6), 1159–1170.
- Auster, E., & Choo, C. W. (1994). CEOs, information, and decision making: Scanning the environment for strategic advantage. *Library Trends*, 43(2), 206–225.

- Babcock, P. (2004). Shedding light on knowledge management. *HR Magazine*, 49(5), 46–50.
- Baker, D. P., Day, R., & Salas, E. (2006). Teamwork as an essential component of high-reliability organizations. *Health services research*, 41(4p2), 1576-1598.
- Baldwin, N. S., & Rice, R. E. (1997). Information-seeking behavior of securities analysts: Individual and institutional influences, information sources and channels, and outcomes. *Journal of the American Society for Information Science and Technology*, 48(8), 674–693.
- Balog, K., Azzopardi, L., and De Rijke, M. 2006. Formal models for expert finding in enterprise corpora. In *Proceedings of ACM SIGIR*. 43–50. <http://dl.acm.org/citation.cfm?id=1148181>
- Beck, K. (2016), “Paint Drip Image.”From <https://www.facebook.com/notes/kent-beck/paint-drip-people/1226700000696195>. Retrieved June 8, 2016.
- Beck, K., Grenning, J., Martin, R.C., Beedle, M., Highsmith, J., Mellor, S., van Bennekum, A., Hunt, A., Schwaber, K., Cockburn, A., Jeffries, R., Sutherland, J., Cunningham, W., Kern, J., Thomas, D., Fowler, M., Marick, B. (2001). "Principles behind the Agile Manifesto". Agile Alliance. Archived from the original on 14 June 2010. Retrieved 6 June 2016.
- Becker, G. S. (1975), “Human Capital”, Chicago University Press, Chicago, IL.
- Beckman, T. (1999). The current state of knowledge management. *Knowledge management handbook*.
- Bernstein, M. S., Ackerman, M. S., Chi, E. H., & Miller, R. C. (2011, May). The trouble with social computing systems research. In *CHI'11 Extended Abstracts on Human Factors in Computing Systems* (pp. 389-398). ACM.
- Blau, P. M. (1964). *Exchange and power in social life*. New York: Wiley.
- Bock, G. -W., & Kim, Y. -G. (2002). Breaking the myths of rewards: An exploratory study of attitudes about knowledge sharing. *Information Resources Management Journal*, 15(2), 14–21.
- Bock, G. -W., Zmud, R. W., Kim, Y. -G., & Lee, J. -N. (2005). Behavioral intention formation in knowledge sharing: Examining the roles of extrinsic motivators, social-psychological forces, and organizational climate. *MIS Quarterly*, 29(1), 87–111.

- Borgatti, S. P., & Cross, R. (2003). A relational view of information seeking and learning in social networks. *Management Science*, 49(4), 432–445.
- Börner, K., Conlon, M., Corson-Rikert, J., & Ding, Y. (2012). VIVO: A semantic approach to scholarly networking and discovery. *Synthesis lectures on the Semantic Web: theory and technology*, 7(1), 1-178.
- Börner, K., Contractor, N., Falk-Krzesinski, H. J., Fiore, S. M., Hall, K. L., Keyton, J., ... & Uzzi, B. (2010). A multi-level systems perspective for the science of team science. *Science Translational Medicine*, 2(49).
- Brehm, J.W. (1966). *A Theory of Psychological Reactance*. New York: Academic Press.
- Brin, S., & Page, L. (2012). The anatomy of a large-scale hypertextual web search engine (reprint). *Computer networks*, 56(18), 3825-3833.
- Brooking, A. (1996). *Intellectual capital*. Cengage Learning EMEA.
- Brooks, A. W., Gino, F., & Schweitzer, M. E. (2015). Smart people ask for (my) advice: Seeking advice boosts perceptions of competence. *Management Science*, 61(6), 1421-1435.
- Burt, R. S. (2001). Closure as Social Capital. *Social capital: Theory and research*, 31-55.
- Cabrera, A., Collins, W. C., & Salgado, J. F. (2006). Determinants of individual engagement in knowledge sharing. *International Journal of Human Resource Management*, 17(2), 245–264.
- Camerer, C., Loewenstein, G. & Weber, M. (1989). The curse of knowledge in economic settings: An experimental analysis. *Journal of Political Economy*, 97, 1232-1254.
- Camillus, J. C. (2008). Strategy as a wicked problem. *Harvard business review*, 86(5), 98.
- Carniero, A. (2000). How does knowledge management influence innovation and competitiveness? *Journal of Knowledge Management*, 4(2):87-98.
- Casciaro, T., & Lobo, M. S. (2005). Competent jerks, lovable fools, and the formation of social networks. *Harvard Business Review*, 83(6), 92–99.
- Castells, M. (2010). *The Rise of the Network Society*, Second edition, Wiley-Blackwell.

- Ceci, S. J., & Liker, J. (1986). A day at the races: A study of IQ, expertise, and cognitive complexity. *Journal of Experimental Psychology: General*, 115(3).
- Chakrabarti, A. K., Feineman, S., & Fuentesvilla, W. (1983). Characteristics of sources, channels, and contents for scientific and technical information systems in industrial R and D. *IEEE Transactions on Engineering Management*, 30(2), 83–88.
- Chang, T. J., Yeh, S. P., & Yeh, I. J. (2007). The effects of joint reward system in new product development. *International Journal of Manpower*, 28(3/4), 276–297
- Chatzkel, J. L. (2003). *Knowledge Capital: How Knowledge-Based Enterprises Really Get Built: How Knowledge-Based Enterprises Really Get Built*. Oxford University Press, USA.
- Chi, E. H. (2009, July). A position paper on 'living laboratories': Rethinking ecological designs and experimentation in human-computer interaction. In *International Conference on Human-Computer Interaction* (pp. 597-605). Springer Berlin Heidelberg.
- Child, J., Faulkner, D., & Tallman, S. B. (2005). *Cooperative strategy*. Oxford University Press, USA.
- Ching-Yung Lin, Nan Cao, Shi Xia Liu, Spiros Papadimitriou, Jimeng Sun, and Xifeng Yan. (2009). SmallBlue: Social Network Analysis for Expertise Search and Collective Intelligence. In *Proceedings of IEEE ICDE*. 1483–1486. DOI: <http://dx.doi.org/10.1109/ICDE.2009.140>
- Chiu, C. -M., Hsu, M. -H., Wang, E., & T. G. (2006). Understanding knowledge sharing in virtual communities: An integration of social capital and social cognitive theories. *Decision Support Systems*, 42(3), 1872–1888.
- Chklovski, T. (2003, October). Learner: a system for acquiring commonsense knowledge by analogy. In *Proceedings of the 2nd international conference on Knowledge capture* (pp. 4-12). ACM.
- Chklovski, T. (2005, October). Collecting paraphrase corpora from volunteer contributors. In *Proceedings of the 3rd international conference on Knowledge capture* (pp. 115-120). ACM.
- Choo, C. W. (1994). Perception and use of information sources by chief executives in environmental scanning. *Library & Information Science Research*, 16(1), 23–40.
- Christensen, E. W., & Bailey, J. R. (1997). A source accessibility effect on media selection. *Management Communication Quarterly*, 10(3), 373–387.

- Civan, A., McDonald, D. W., Unruh, K. T., & Pratt, W. (2009). Locating patient expertise in everyday life. In Proceedings of the GROUP 2009 conference on supporting group work (pp. 291–300). New York: ACM Press
- Cloud Campus: The Software-Defined College. (n.d.). Retrieved May 18, 2016, from <https://www.meritalk.com/study/cloud-campus/>
- Coleman, D. (1999). Groupware: collaboration and knowledge sharing. Knowledge Management Handbook, CRC Press, Boca Raton, FL, pages 12-15.
- Collins, C. J., & Smith, K. G. (2006). Knowledge exchange and combination: The role of human resource practices in the performance of high-technology firms. *Academy of Management Journal*, 49(3), 544–560
- Connelly, C. E., & Kelloway, E. K. (2003). Predictors of employees' perceptions of knowledge sharing cultures. *Leadership & Organization Development Journal*, 24(5/6), 294–301.
- Constant, D., Kiesler, S., & Sproull, L. (1994). What's mine is ours, or is it? A study of attitudes about information sharing. *Information Systems Research*, 5(4), 400–421
- Cool, C., & Xie, H. (2000). Patterns of information use, avoidance and evaluation in a corporate engineering environment. In Proceedings of the ASIS 2000 annual meeting of the American society for information science (pp. 462–472). Medford, NJ: Information Today.
- Cooper, S., Khatib, F., Treuille, A., Barbero, J., Lee, J., Beenen, M., ... & Popović, Z. (2010). Predicting protein structures with a multiplayer online game. *Nature*, 466(7307), 756-760.
- Criscuolo, P., Salter, A., & Sheehan, T. (2007). Making knowledge visible: Using expert yellow pages to map capabilities in professional services firms. *Research Policy*, 36(10), 1603-1619.
- Cross, R., & Cummings, J. N. (2004). Tie and network correlates of individual performance in knowledge-intensive work. *Academy of Management Journal*, 47(6), 928–937.
- Cross, R., & Sproull, L. (2004). More than an answer: Information relationships for actionable knowledge. *Organization Science*, 15(4), 446–462.
- Cross, R.L., Martin, R.D. and Weiss, L.M. 2006. Mapping the value of employee collaboration. *McKinsey Quarterly*, August 2006.

- Culnan, M. J. (1983). Environmental scanning: The effects of task complexity and source accessibility on information gathering behavior. *Decision Sciences*, 14(2), 194–206.
- Cummings, J. N. (2004). Work groups, structural diversity, and knowledge sharing in a global organization. *Management science*, 50(3), 352-364.
- Dalkir, K. (2005), “Knowledge management in theory & practice”, Elsevier Butterworth Heinemann.
- Davenport, T. H. and Prusak, L. (2000). *Working Knowledge: How Organisations Manage What They Know*. Cambridge: Harvard Business School Press.
- McDonald, D.W. and Ackerman, M.S. 1998. Just talk to me: a field study of expertise location. In *Proceedings of the ACM conference on Computer supported cooperative work*. 315–324. <http://dl.acm.org/citation.cfm?id=289506>
- De Long, D. W., & Fahey, L. (2000). Diagnosing cultural barriers to knowledge management. *Academy of Management Executive*, 14(4), 113–127.
- de Vries, R. E., van den Hooff, B., & de Ridder, J. A. (2006). Explaining knowledge sharing: The role of team communication styles, job satisfaction, and performance beliefs. *Communication Research*, 33(2), 115–135.
- Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Plenum.
- Deci, E. L., & Ryan, R. M. (2000). The “what” and “why” of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, 11(4), 227–268.
- Deloitte. (2016). MS2016: One foot out the door. From <http://www2.deloitte.com/global/en/pages/about-deloitte/articles/gx-millennials-one-foot-out-the-door.html#report>, retrieved June 13, 2016.
- DePaulo, B. M., & Fisher, J. D. (1980). The costs of asking for help. *Basic and Applied Social Psychology*, 1(1), 23–35
- Du Plessis, M. (2007) ,“ Knowledge Management: what makes complex implementations successful? ”, *Journal of Knowledge Management*, Vol.11 No.2, pp.91-101.

- Elect, empower, enhance – TOP 3 influencer involvement scenarios. (n.d.). Retrieved July 21, 2016, from <http://blog.orgmapper.com/2016/07/05/elect-empower-enhance-top-3-influencer-involvement-scenarios/>
- Emerson, R. M. (1981). Social exchange theory. In M. Rosenberg & R.H. Turner (Eds.), *Social psychology: Sociological perspectives* NY: Basic Books, Inc.
- Evans, M. and Ali, N. (2013). “Bridging Knowledge Management Life Cycle Theory and Practice.” *Proceedings of the 10th International Conference on Intellectual Capital, Knowledge Management & Organisational Learning, Washington*
- Fahey, L., & Prusak, L. (1998). The eleven deadliest sins of knowledge management. *California Management Review*, 40, 265–276.10.2307/41165954[CrossRef].
- Fensel, D. (2001). *Ontologies: Dynamic networks of formally represented meaning*, <http://sw-portal.derit.at/papers/publications/network.pdf>
- Fishbein, M. & Wexler, J. (2014). *What Product Managers Need to Know About Rapid Prototyping. Alpha UX.*
- Fruchterman, T. M. J., & Reingold, E. M. (1991). Graph drawing by force-directed placement. *Software: Practice & Experience*, 21(11), 1129–1164.
- Furnas, G. W., Landauer, T. K., Gomez, L. M., & Dumais, S. T. (1987). The Vocabulary Problem in Human-System Communication. *Communications of the ACM*, 30(11), pp. 964–971.
- Gagné, M. (2009). A model of knowledge-sharing motivation. *Human Resource Management*, 48(4), 571.
- Galinsky, A. D., & Kray, L. J. (2004). From thinking about what might have been to sharing what we know: The effects of counterfactual mind-sets on information sharing in groups. *Journal of Experimental Social Psychology*, 40(5), 606-618.
- Galton, Francis (1907-03-07). "Vox Populi" (PDF). *Nature*.doi:10.1038/075450a0
- Garbash, D. A. (2015, August 17). *Knownodes*. Retrieved September 10, 2015, from https://www.youtube.com/watch?v=LBt5vsHh_h0

- Gerstberger, P. G., & Allen, T. J. (1968). Criteria used by research and development engineers in the selection of an information source. *Journal of Applied Psychology*, 52(4), 272–279.
- Ghoshal, S. and Gratton, L. (2002) “Integrating the Enterprise,” *MIT Sloan Management Review* 44, no. 1: 31–38.
- Giacomin, J. (2014). What Is Human Centered Design? *The Design Journal*, 17(4), 606-623.
- Ginsburg, M. and Kambil, A. (1999). Annotate: A Web-based knowledge management support system for document collections. In *System Sciences, 1999. HICSS-32. Proceedings of the 32nd Annual Hawaii International Conference on System sciences*.
- Gitomer, D. H. (1988). Individual differences in technical troubleshooting. *Human Performance*, 1, 111-131.
- Glenn, J.C. (2008) *State of the Future*, Chapter 5. The Millennium Project, Washington, DC.
- Gloet, M. and Terziovski, M. (2004). Exploring the relationship between knowledge management practices and innovation performance. *Journal of manufacturing Technology Management* 15(5):402-409.
- Gobet, F., & Simon, H. A. (1998). Expert chess memory: Revisiting the chunking hypothesis. *Memory*, 6, 225-255
- Guimera, R., Arenas, A., & Diaz-Guilera, A. (2001). Communication and optimal hierarchical networks. *Physica A: Statistical Mechanics and its Applications*, 299(1), 247-252.
- Haklay, M. (2015). "Citizen Science and Policy: A European Perspective" (PDF). Woodrow Wilson International Center for Scholars. p. 11. Retrieved 3 June 2016.
- Hall, H., & Graham, D. (2004). Creation and recreation: motivating collaboration to generate knowledge capital in online communities. *International Journal of Information Management*, 24(3), 235-246.
- Hansen, M. (2002) “Turning the Lone Star Into a Real Team Player,” *Financial Times*, Aug. 7, 11–13.

- Hansen, M. T. (1999). The search-transfer problem: The role of weak ties in sharing knowledge across organization subunits. *Administrative Science Quarterly*, 44(1), 82–111.
- Hansen, M. T. (2002). Knowledge network: Explaining effective knowledge sharing in multiunit companies. *Organization Science*, 13(3), 232–248.
- Hansen, M. T., & Nohria, N. (2004). How to build collaborative advantage. *MIT Sloan Management Review*, 46(1), 22.
- Hansen, M. T., Mors, M. L., & Lovas, B. (2005). Knowledge sharing in organizations: Multiple networks, multiple phases. *Academy of Management Journal*, 48(5), 776–793.
- Hansen, M. T., Nohria, N., & Tierney, T. (1999). What's your strategy for managing knowledge? *Harvard Business Review*, 77(2), 106–116.
- Haq, M.u. (1996) *Reflections on Human Development*. Delhi: Oxford University Press.
- Helms, R., Diemer, D., & Lichtenstein, S. (2011). Exploring barriers in expertise seeking: Why don't they ask an expert? In *PACIS 2011: Proceedings of the 15th Pacific Asia conference on information systems* (pp. 77:01–77:13). Atlanta, GA: AIS Press
- Hepp, M. (2008). Ontologies: State of the art, business potential, and grand challenges. In *Ontology Management* (pp. 3-22). Springer US.
- Hertzum, M. (2002). The importance of trust in software engineers' assessment and choice of information sources. *Information & Organization*, 12(1), 1–18
- Hertzum, M. (2014). Expertise seeking: A review. *Information Processing & Management*, 50(5), 775-795.
- Hibbard, J. (1997). Knowing what we know. URL: <http://www.informationweek.com/653/53iukno.htm> (visited on 05/30/2013).
- Hill, R. A., & Dunbar, R. I. M. (2003). Social network size in humans. *Human Nature*, 14(1), 53–72.
- Hinds, P. (1999). The curse of expertise: The effects of expertise and debiasing methods on predictions of novice performance. *Journal of Experimental Psychology: Applied*, 5: 205-221.

- Hinds, P. J., & Pfeffer, J. (2003). Why organizations don't "know what they know": Cognitive and motivational factors affecting the transfer of expertise. *Sharing expertise: Beyond knowledge management*, 3-26.
- Hinds, P., Burgess, D., Pfeffer, J. & Moore, P. (2001). Laying the groundwork: The role of organizational context in knowledge sharing and seeking. Working paper. Stanford University.
- Hirsh, S., & Dinkelacker, J. (2004). Seeking information in order to produce information: An empirical study at Hewlett Packard Labs. *Journal of the American Society for Information Science and Technology*, 55(9), 807–817.
- Hollingshead, A., Fulk, J. & Monge, P. (in press). Fostering intranet knowledge-sharing: An integration of transactive memory and public goods approaches. In P. Hinds and S. Kiesler (Eds.), *Working Across Distance*. Boston, MA: MIT Press.
- Holsapple, C. and Whinston, A.(1987). Knowledge-based organizations. *Information society*, 5(2):77-90.
- HOW4 – One platform for different needs. (n.d.). Retrieved July 22, 2016, from <http://www.how-4.com/Applications>
<https://summaries.com/Platinum/Collaboration.pdf>
- IDC (2012). *The Digital Universe in 2020: Big Data, Bigger Digital Shadows, and Biggest Growth in the Far East*, December 2012, sponsored by EMC
- Illich, I. (1973). *Deschooling society*. Harmondsworth, Middlesex.
- Jarvenpaa, S. L., & Staples, D. S. (2000). The use of collaborative electronic media for information sharing: An exploratory study of determinants. *The Journal of Strategic Information Systems*, 9(2–3), 129–154.
- Jarvenpaa, S. L., & Staples, D. S. (2001). Exploring perceptions of organizational ownership of information and expertise. *Journal of Management Information Systems*, 18(1), 151–183.
- Jones, M. C. (2005). Tacit knowledge sharing during ERP implementation: A multi-site case study. *Information Resources Management Journal*, 18(2), 1–23.

- Julien, H., & Michels, D. (2000). Source selection among information seekers: Ideals and realities. *Canadian Journal of Information and Library Science*, 25(1), 1–18
- Kankanhalli, A., Tan, B. C. Y., & Wei, K. -K. (2005). Contributing knowledge to electronic knowledge repositories: An empirical investigation. *MIS Quarterly*, 29(1), 113–143
- Kautz, H., Selman, B., and Shah, M. (1997). Referral Web: Combining Social Networks and Collaborative Filtering. *Commun. ACM* 40, 3 (March 1997), 63–65. DOI: <http://dx.doi.org/10.1145/245108.245123>
- Kim, S. and Lee, H. (2006), “The impact of organizational context and information technology on employee knowledge-sharing capabilities”, *Public Administration Review*, Vol. 66 No. 3, pp. 370-385.
- Kim, S., & Lee, H. (2006). The impact of organizational context and information technology on employee knowledge-sharing capabilities. *Public Administration Review*, 66(3), 370–385.
- Kim, W. C., & Mauborgne, R. (1998). Procedural justice, strategic decision making, and the knowledge economy. *Strategic Management Journal*, 19(4), 323–338.
- King, W. R., Marks Jr, P. V., & McCoy, S. (2002). The most important issues in knowledge management. *Communications of the ACM*, 45(9), 93-97.
- Kramer, R. (1991). Intergroup relations and organizational dilemmas: The role of categorization processes. *Research in Organizational Behavior*, 13, 191-228.
- Kubo, I., Saka, A., & Pam, S. L. (2001). Behind the scenes of knowledge sharing in a Japanese bank. *Human Resource Development International*, 4(4), 465–485.
- Kulkarni, U. R., Ravindran, S., & Freeze, R. (2006). A knowledge management success model: Theoretical development and empirical validation. *Journal of Management Information Systems*, 23(3), 309–347
- Kwok, S. H., & Gao, S. (2005). Attitude towards knowledge sharing behavior. *The Journal of Computer Information Systems*, 46(2), 45–51.
- Langer, E. J. & Imber, L. G. (1979). When practice makes imperfect: Debilitating effects of overlearning. *Journal of Personality and Social Psychology*, 37, 2014-2024.

- Larson, J. R., & Harmon, V. M. (2007). Recalling shared vs. unshared information mentioned during group discussion: Toward understanding differential repetition rates. *Group Processes & Intergroup Relations*, 10(3), 311–322.
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge, UK: Cambridge University Press
- Lee, F. (2002). The social costs of seeking help. *Journal of Applied Behavioral Science*, 38(1), 17–35
- Lee, J. -H., Kim, Y. -G., & Kim, M. -Y. (2006). Effects of managerial drivers and climate maturity on knowledge-management performance: Empirical validation. *Information Resources Management Journal*, 19(3), 48–60.
- Lenat, D. B., Guha, R. V., Pittman, K., Pratt, D., & Shepherd, M. (1990). Cyc: toward programs with common sense. *Communications of the ACM*, 33(8), 30-49.
- Leonard, D. & Sensiper, S. (1998). The role of tacit knowledge in group innovation. *California Management Review*, 40 (3), 112-132.
- Levin, D. Z., & Cross, R. (2004). The strength of weak ties you can trust: The mediating role of trust in effective knowledge transfer. *Management Science*, 50, 1477–1490.
- Liao, L. -F. (2006). A learning organization perspective on knowledge-sharing behavior and firm innovation. *Human Systems Management*, 25(4), 227.
- Liebowitz, J. (2003). A knowledge management strategy for the Jason organization: A case study. *Journal of Computer Information Systems*, 44(2), 1–5.
- Liebowitz, J., & Megbolugbe, I. (2003). A set of frameworks to aid the project manager in conceptualizing and implementing knowledge management initiatives. *International Journal of Project Management*, 21(3), 189–198.
- Lin, C. -P. (2007a). To share or not to share: Modeling knowledge sharing using exchange ideology as a moderator. *Personnel Review*, 36(3), 457–475.
- Lin, C. -P. (2007b). To share or not to share: Modeling tacit knowledge sharing, its mediators and antecedents. *Journal of Business Ethics*, 70(4), 411–428.

- Lin, H. -F. (2007c). Effects of extrinsic and intrinsic motivation on employee knowledge sharing intentions. *Journal of Information Science*, 33(2), 135–149.
- Lin, H. -F. (2007d). Knowledge sharing and firm innovation capability: An empirical study. *International Journal of Manpower*, 28(3/4), 315–332.
- Lin, H. -F., & Lee, G. -G. (2004). Perceptions of senior managers toward knowledge-sharing behaviour. *Management Decision*, 42(1/2), 108–125.
- Lin, H. -F., & Lee, G. -G. (2006). Effects of socio-technical factors on organizational intention to encourage knowledge sharing. *Management Decision*, 44(1), 74–88.
- Lin, H.F. (2007), “Effects of extrinsic and intrinsic motivations on employee intentions to share knowledge”, *Journal of Information Science*, Vol. 33 No. 2, pp. 135-149.
- Lu, L., & Yuan, Y. C. (2011). Shall I google it or ask the competent villain down the hall? The moderating role of information need in information source selection. *Journal of the American Society for Information Science and Technology*, 62(1), 133–145.
- Lynd, R.S. (1939) Knowledge for what?
- Maddocks, J. & Beaney, M. 2002. See the invisible and intangible. *Knowledge Management*, March, 16-17.
- Marmer, M., Herrmann, B. L., Dogrultan, E., Berman, R., Eesley, C., & Blank, S. (2011). Startup genome report extra: Premature scaling. *Startup Genome*,10.
- Maybury, M.T. (2006). Expert Finding Systems. Mitre Technical Report. URL: http://www.mitre.org/work/tech_papers/tech_papers_06/06_1115/06_1115.pdf (visited on 10/06/12).
- McDonald, D. W., & Ackerman, M. S. (1998). Just talk to me: A field study of expertise location. In *Proceedings of the CSCW'98 conference on computer supported cooperative work* (pp. 315–324). New York: ACM Press.
- McKinnon, J. L., Harrison, G. L., Chow, C. W., & Wu, A. (2003). Organizational culture: Association with commitment, job satisfaction, propensity to remain, and information sharing in Taiwan. *International Journal of Business Studies*, 11(1), 25–44.

- Mengistu H, Huizinga J, Mouret J-B, Clune J (2016) The Evolutionary Origins of Hierarchy. *PLoS Comput Biol* 12(6): e1004829. doi:10.1371/journal.pcbi.1004829
- Mesmer-Magnus, J. R., & DeChurch, L. A. (2009). Information sharing and team performance: A meta-analysis. *Journal of Applied Psychology*, 94, 535–546
- Mintzberg, H., Jorgensen, J., Dougherty, D., & Westley, F. (1996). Some surprising things about collaboration—Knowing how people connect makes it work better. *Organizational Dynamics*, 25(1), 60-71.
- Mohamed, M., Stankosky, M. and Murray, A. (2006) "Knowledge management and information technology: can they work in perfect harmony?", *Journal of Knowledge Management*, Vol. 10 No. 3, pp.103 – 116.
- Momtazi Saeedeh and Felix Naumann. 2013. Topic modeling for expert finding using LDA. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 3, 5 (2013), 346–353. DOI: <http://dx.doi.org/10.1002/widm.1102>
- Nevo, D., Benbasat, I. and Wand, Y. (2009). Who Knows What? URL: <http://online.wsj.com/article/SB10001424052970203946904574302032097910314.html> (visited on 10/30/2014).
- Nevo, D., Benbasat, I., & Wand, Y. (2012). The knowledge demands of expertise seekers in two different contexts: Knowledge allocation versus knowledge retrieval. *Decision Support Systems*, 53(3), 482–489.
- Nielsen, M. (2011), *Reinventing Discovery: The New Era of Networked Science*. Princeton University Press.
- Nonaka, I. (1991). The Knowledge Creating Company, *Harvard Business Review*. November/December: 96-104
- Norman, D. A. (2005). Human-centered design considered harmful. *Interactions*, 12(4), 14-19.
- O'Dell, C., & Jackson Grayson, C. (1998). If only we knew what we know: Identification and transfer of internal best practices. *California Management Review*, 40(3), 154-174.

- O'Reilly, C. A. (1982). Variations in decision makers' use of information sources: The impact of quality and accessibility of information. *Academy of Management Journal*, 25(4), 756–771.
- Organizational network analysis. (2016, July 11). In Wikipedia, The Free Encyclopedia. Retrieved 21:30, July 21, 2016, from https://en.wikipedia.org/w/index.php?title=Organizational_network_analysis&oldid=729302738
- Oshri, I., Van Fenema, P., Kotlarsky, J. (2008). Knowledge transfer in globally distributed teams: the role of transactive memory, *Information Systems Journal*, 18 (2008) 593- 616.
- Pan, S. and H. Scarbrough (1998). A socio-technical view of knowledge-sharing at Buckman Laboratories. *Journal of Knowledge Management*, 2, 55-56.
- Parise, S., Cross, R. and Davenport, T.H. (2005). “It’s Not What but Who You Know: How Organizational Network Analysis Can Help Address Knowledge Loss Crises.” *Lost Knowledge Roundtable, The Network Roundtable at the University of Virginia*
- Paul, S. A. (2016, May). Find an Expert: Designing Expert Selection Interfaces for Formal Help-Giving. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (pp. 3038-3048). ACM.
- Pfeffer, J. & Sutton, R. (2000). *The Knowing-Doing Gap: How Smart Companies Turn Knowledge Into Action*. Boston: Harvard Business School Press
- Pischke, Steve (2015). *Lectures in Labor Economics, Chapter 1*, <http://econ.lse.ac.uk/staff/spischke/ec533/Acemoglu%20Autor%20chapter%201.pdf>
- Plous, Scott (1993). *The Psychology of Judgment and Decision Making*. p. 233.
- Polanyi, M (1966). *The Tacit Dimension*. New York, NY: Doubleday
- Postmes, T., Spears, R., & Cihangir, S. (2001). Quality of decision making and group norms. *Journal of personality and social psychology*, 80(6), 918.
- Probst, G., Raub, S. and Romhardt, K. (2000). *Managing Knowledge: Building Blocks for Success*. Chichester: John Wiley & Sons Ltd.

- Prusak, L. (2011). Building a collaborative enterprise.
- Quigley, J., Peck, D. R., Rutter, S., & Williams, E. M. (2002). Making choices: Factors in the selection of information resources among science faculty at the University of Michigan: Results of a survey conducted July–September, 2000. *Issues in Science & Technology Librarianship*, 34.
- Rašul, J., Vu kšić, V. B. and Štemberger, M. I. (2012), “THE IMPACT OF KNOWLEDGE MANAGEMENT ON ORGANISATIONAL PERFORMANCE”, *ECONOMIC AND BUSINESS REVIEW*, Vol. 14 No. 2, pp.147-168.
- Reagans, R., & McEvily, B. (2003). Network structure and knowledge transfer: The effects of cohesion and range. *Administrative Science Quarterly*, 48(2), 240–267.
- Reichling, T. and Wulf, V.. 2009. Expert Recommender Systems in Practice: Evaluating Semi-automatic Profile Generation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09)*. ACM, New York, NY, USA, 59–68. DOI: <http://dx.doi.org/10.1145/1518701.1518712>
- Ries, E. (2011). *The lean startup: How today's entrepreneurs use continuous innovation to create radically successful businesses*. Crown Books.
- Rittel, H. (1973). "Dilemmas in a General Theory of Planning." *Policy Sciences*: 155-169.
- Robinson, M. A. (2010). An empirical analysis of engineers' information behaviors. *Journal of the American Society for Information Science and Technology*, 61(4), 640–658.
- Ruggill, J., & McAllister, K. (2006). The wicked problem of collaboration. *M/C Journal: A Journal of Media and Culture*, 9, 16.
- Rupp, D. E., & Cropanzano, R. (2002). The mediating effects of social exchange relationships in predicting workplace outcomes from multifoci organizational justice. *Organizational Behavior & Human Decision Processes*, 89(1), 925–946.
- Ruppel, C. P., & Harrington, S. J. (2001). Sharing knowledge through intranets: A study of organizational culture and intranet implementation. *IEEE Transactions on Professional Communication*, 44(1), 37–52.

- Sawng, Y. W., Kim, S. H., & Han, H. -S. (2006). R&D group characteristics and knowledge management activities: A comparison between ventures and large firms. *International Journal of Technology Management*, 35(1–4), 241–261.
- Schepers, P., & van den Berg, P. T. (2007). Social factors of work-environment creativity. *Journal of Business and Psychology*, 21(3), 407–428
- Schubmehl & Vesset. June 2014. The Knowledge Quotient: Unlocking the Hidden Value of Information Using Search and Content Analytics. Coveo white paper.
- Scott, W. R., & Davis, G. F. (2015). *Organizations and organizing: Rational, natural and open systems perspectives*. Routledge.
- Shami, N. S., Ehrlich, K., Gay, G., & Hancock, J. T. (2009). Making sense of strangers' expertise from signals in digital artifacts. In *Proceedings of the CHI2009 conference on human factors in computing systems* (pp. 69–78). New York: ACM Press.
- Sharratt, M. and Usoro, A. (2003), "Understanding knowledge-sharing in online communities of practice", *Electronic Journal of Knowledge Management*, Vol. 1 No. 2, pp. 187-196.
- Shepperd, James A.; Patrick Carroll; Jodi Grace; Meredith Terry (2002). "Exploring the Causes of Comparative Optimism" (PDF). *Psychologica Belgica* 42: 65–98. CiteSeerX: 10.1.1.507.9932
- Siemens, E., Balasubramanian, S., & Roth, A. V. (2007). Incentives that induce task-related effort, helping, and knowledge sharing in workgroups. *Management Science*, 53(10), 1533–1550
- Singh, P., Lin, T., Mueller, E. T., Lim, G., Perkins, T., & Zhu, W. L. (2002, October). Open Mind Common Sense: Knowledge acquisition from the general public. In *OTM Confederated International Conferences" On the Move to Meaningful Internet Systems"* (pp. 1223-1237). Springer Berlin Heidelberg.
- Siorpaes, K., & Hepp, M. (2008). Games with a purpose for the semantic web. *IEEE Intelligent Systems*, 23(3), 50-60.
- Socientize Project (2013). *Green Paper on Citizen Science: Citizen Science for Europe - Towards a better society of empowered citizens and enhanced research* (PDF). Socientize consortium.
- Srivastava, A., Bartol, K. M., & Locke, E. A. (2006). Empowering leadership in management teams: Effects on knowledge sharing, efficacy, and performance. *Academy of Management Journal*,

49(6), 1239–1251.

Stasser, G., & Titus, W. (1987). Effects of information load and percentage of shared information on the dissemination of unshared information during group discussion. *Journal of Personality and Social Psychology*, 53(1), 81–93.

Stasser, G., & Titus, W. (2003). Hidden profile: A brief history. *Psychological Inquiry*, 14, 304–313.

Stasser, G., Garold, S., & William, T. (1985). Pooling of unshared information in group decision making: Biased information sampling during discussion. *Journal of Personality and Social Psychology*, 48(6), 1467–1478.

Stokols, D., Hall, K. L., Taylor, B. K., & Moser, R. P. (2008). The science of team science: overview of the field and introduction to the supplement. *American journal of preventive medicine*, 35(2), S77-S89.

STRATEGIC DIRECTION.(2007), “Managing knowledge management ”, STRATEGIC DIRECTION, VOL. 23 NO. 7, pp. 24-26.

Streeter, L.A. and Lochbaum, K.E. 1988. An expert/expert-locating system based on automatic representation of semantic structure. In *Proceedings of the Fourth Conference on Artificial Intelligence Applications*. 345–350. DOI: <http://dx.doi.org/10.1109/CAIA.1988.196129>

Sveiby, Karl Erik (1997). "The Intangible Asset Monitor". *Journal of Human Resource Casting and Accounting* 2 (1).

Tan, H. H., & Zhao, B. (2003). Individual- and perceived contextual-level antecedents of individual technical information inquiry in organizations. *Journal of Psychology*, 137(6), 597–621.

The Future of Corporate IT (2013). 2013-2017 CEB Enterprise Architecture Leadership Council.

Thomas-Hunt, M. C., Ogden, T. Y., & Neale, M. A. (2003). Who's really sharing? Effects of social and expert status on knowledge exchange within groups. *Management Science*, 49(4), 464–477.

Tow, W. N. F. H., Venable, J. R., & Dell, P. (2015). Developing a Theory of Knowledge Identification Effectiveness in Knowledge Management. PACIS 2015 Proceedings. Paper 85. <http://aisel.aisnet.org/pacis2015/85>

- Tsai, W. (2002), "Social structure of 'coopetition' within a multiunit organization: coordination, competition, and intraorganizational knowledge sharing", *Organization Science*, Vol. 13 No. 2, pp. 179-190.
- Valmohammadi, C., & Ahmadi, M. (2015). The impact of knowledge management practices on organizational performance: A balanced scorecard approach. *Journal of Enterprise Information Management*, 28(1), 131-159.
- Valmohammadi, Ch. (2010), "Investigation and Assessment of Critical Success Factors of Knowledge Management Implementation in Iranian Small-to-Medium Sized Enterprises", *Journal of Applied Sciences*, Vol. 10(19), pp. 2290-2296.
- Vancouver, J. B., & Morrison, E. W. (1995). Feedback inquiry: The effect of source attributes and individual differences. *Organizational Behavior and Human Decision Processes*, 62(3), 276–285.
- Von Ahn, L., Kedia, M., & Blum, M. (2006, April). Verbosity: a game for collecting common-sense facts. In *Proceedings of the SIGCHI conference on Human Factors in computing systems* (pp. 75-78). ACM.
- W., Kim, Y. -G., & Lee, J. -N. (2005). Behavioral intention formation in knowledge sharing: Examining the roles of extrinsic motivators, social-psychological forces, and organizational climate. *MIS Quarterly*, 29(1), 87–111.
- Wang, C. -C. (2004). The influence of ethical and self-interest concerns on knowledge sharing intentions among managers: An empirical study. *International Journal of Management*, 21(3), 370–381.
- Wang, S., & Noe, R. A. (2010). Knowledge sharing: A review and directions for future research. *Human Resource Management Review*, 20(2), 115-131.
- Wasko, M. M., & Faraj, S. (2000). "It is what one does": Why people participate and help others in electronic communities of practice. *The Journal of Strategic Information Systems*, 9(2–3), 155–173.
- Wasko, M. M., & Faraj, S. (2005). Why should I share? Examining social capital and knowledge contribution in electronic networks of practice. *MIS Quarterly*, 29(1), 35–57.

- Watts, D. (1999). Networks, Dynamics, and the Small-World Phenomenon. *American Journal of Sociology*, 105(2), 493-527. doi:1. Retrieved from <http://www.jstor.org/stable/10.1086/210318> doi:1
- Weber, F., Dauphin, E., Fuschini, R., Haarmann, J., Katzung, A. and Wunram, M. (2007). "Expertise transfer: A case study about knowledge retention at Airbus." In S.K. Pawar, K.-D. Thoben & M. Pallot (Eds.), *ICE 2007 Proceedings*. France: Sophia-Antipolis.
- Wexler, M.N. (2001). "The who, what and why of knowledge mapping." *Journal of Knowledge Management* 5(3), 249-263.
- What is Evernote Business Expertise Discovery? (n.d.). Retrieved July 27, 2016, from <https://help.evernote.com/hc/en-us/articles/209005477-What-is-Evernote-Business-Expertise-Discovery->
- What is ONA? (n.d.). Retrieved July 22, 2016, from http://www.robcross.org/network_ona.htm
- Wilkinson, M. A. (2001). Information sources used by lawyers in problem-solving: An empirical exploration. *Library & Information Science Research*, 23(3), 257–276.
- Willem, A., & Scarbrough, H. (2006). Social capital and political bias in knowledge sharing: An exploratory study. *Human Relations*, 59(10), 1343–1370.
- Wong, K.Y, Aspinwall, E. (2005), " An empirical study of the important factors for knowledge-management adoption in the SME sector ", *Journal of Knowledge Management*, Vol.9. No.3, pp. 64-82.
- Woolley, A.W.; Aggarwal, I.; Malone, T.W. (2015-12-01)."Collective Intelligence and Group Performance". *Current Directions in Psychological Science* 24 (6): 420–424.
- Woolley, Anita Williams; Chabris, Christopher F.; Pentland, Alex; Hashmi, Nada; Malone, Thomas W. (2010-10-29). "Evidence for a Collective Intelligence Factor in the Performance of Human Groups". *Science* 330(6004): 686–688
- Woudstra, L., & Hooff, B. v. d (2008). Inside the source selection process: Selection criteria for human information sources. *Information Processing & Management*, 44(3), 1267–1278.

- Wu, W. -L., Hsu, B. -F., & Yeh, R. -S. (2007). Fostering the determinants of knowledge transfer: A team-level analysis. *Journal of Information Science*, 33(3), 326–339.
- Xu, Y., Tan, C. Y., & Yang, L. (2006). Who will you ask? An empirical study of interpersonal task information seeking. *Journal of the American Society for Information Science and Technology*, 57(12), 1666–1677.
- Yang, C., & Chen, L. -C. (2007). Can organizational knowledge capabilities affect knowledge sharing behavior? *Journal of Information Science*, 33(1), 95–109
- Yimam-Seid & Kobsa, Expert Finding Systems for Organizations: Problem and Domain Analysis and the DEMOIR Approach. *Journal of Organizational Computing & Electronic Commerce* 13 (1), 2003.
- Yuan, Y. C., Rickard, L. N., Xia, L., & Scherer, C. (2011). The interplay between interpersonal and electronic resources in knowledge seeking among co-located and distributed employees. *Journal of the American Society for Information Science and Technology*, 62(3), 535–549.
- Zack, M. H. (1999), “Developing a knowledge strategy”, *California Management Review*, Spring, Vol.41 No. 3, pp. 125-145.
- Zack, M., McKeen, J. and Singh, S. (2009), “knowledge management and organizational performance: an exploratory analysis”, *journal of knowledge management*, Vol. 13.No 6,pp.392-409.

Appendix 1: Instructions for AIV students Internship mapping

We would like to finalize next week the first map of the CRI. To this end, there are 6 tasks to fulfill by Tuesday April 5th, as detailed below. Some of you had partially started the process, most of you are long behind:)

I will be at the Meet-up room in Charles V on Friday at 14:00 (before your Scientific Communication class) to help out/debug/solve any problems you may be having with the rhizi. Don't hesitate to join me with your computers. it should not take more than 10 minutes of your time to do this correctly!

Your tasks:

1. Follow the tutorial, [posted here \(http://www.moodle-cri.org/mod/page/view.php?id=7537\)](http://www.moodle-cri.org/mod/page/view.php?id=7537)

2. Please create an account at <http://cri.rhizi.net/>. Either go to the [AIV M1 2015 page](#) directly, or choose the file from the 'Open' menu...

3. Add (i) yourself (including family name!) (ii) 3 interests

example: "Jane Kern is interested in molecular biology and synthetic biology and evolution"

[notes:

(a) there is a double space separation between terms

(b) when the cursor is on a term, you can toggle the node type (person/keyword/... by pressing <shift>+<up/down arrow>. Please use "Person" for your name and "Keyword" for your interest as node types.

(c) Once you press <enter>, click on your node and add your email, website (personal, LinkedIn etc) and a short description in the pop-up window. Anything you enter will be saved; there's no 'save' button.]

4. Add (i) Your Internship (as project node), your supervisor (as person node) and where you do your internship (as organization node).

example:

Jane Kern m1 internship Cloning GRT4 gene supervised by Lydia Amsen at Curie Institute UMR12

[note: Once you press <enter>, click on the new nodes and add more information as the URL of the lab, of your supervisor and a short description of your project!]

5. Enter 3 of your skills:

example:

"Jane Kern is expert in Medicinal Chemistry and Virology"

"Jane Kern knows about genetics"

6. Add up to 3 skills you are acquiring during your internship

example:

"Jane Kern learned cloning and multiple alignment and primer design working on Cloning GRT4 gene"

Appendix 2: Rhizi User's manual

[Tutorial Video](#)

[Registration](#)

[Login](#)

[Logout](#)

[Navigation between Rz-docs](#)

[Create a new Rz-doc](#)

[Delete an Rz-doc](#)

[Search Rz-doc](#)

[Select node/edge inside an Rz-doc](#)

[Select multiple nodes/edges](#)

[Deselect nodes/edges](#)

[Node/Edge creation](#)

[Modifying Node/Edge type](#)

[Create only edge](#)

[Edit Node/Edge](#)

[Add thumbnail image to a node](#)

[Delete node/edge](#)

[Change layout](#)

[Changing node position](#)

[Exploring content inside an Rz-doc](#)

[Accessing Urls](#)

[Filtering nodes](#)

[Find shortest path](#)

[Tutorial Video](#)

Registration

In landing page, click “**Register**” button. Fill in the form with your first name, last name, display name, email, password and password confirmation. You will receive a validation request to the email you've specified. Click on the link received in the email. Click on the link appearing in the validation

page. Go to landing page and you can login your accou

Sign up

	First name	
	Last name	*
	Rhizi username	*
	Email	
	password	
	Re-type password	

Sign Up

Login

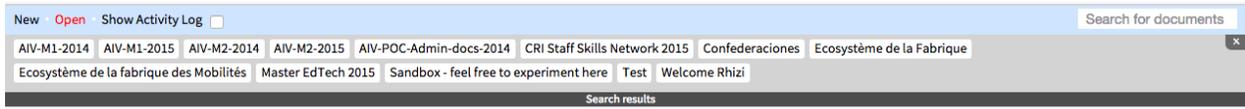
Fill in the registered email and password and click on the “Login” button. If you’ve forgot your password, click on the “Forgot password?” button. You will be asked to fill in your email address. A confirmation email will be sent to you email address. Click on the link in the mail. You will be redirected to a change password page. Type in and retype a new password.

Logout

Click the “logout” button in the top-bar.

Navigation between Rz-docs

After logging-in, a default Rz-doc will open. To show list of available Rz-doc click on the “open” button in the top bar. To open a doc, click on it’s name.



Create a new Rz-doc

In the top-bar, Click on “New” button. An input box will appear, type in it the Rz-doc title and press “**submit**” to create the doc. The new empty doc will open.

Delete an Rz-doc

In browser address bar, type: [Rhizi-Url]/api/rzdoc/[doctitle]/delete

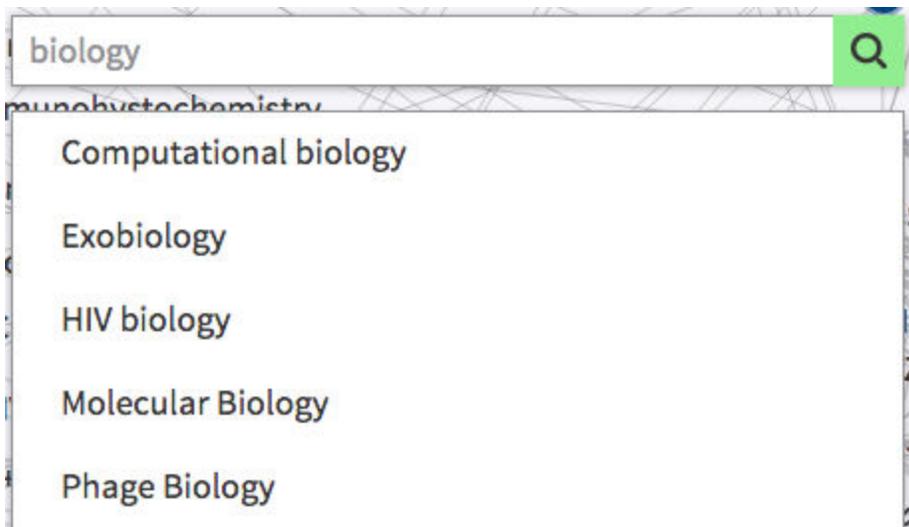
Search Rz-doc

Users can find and select specific entities in the map by using open text search. This search looks within the node title and a auto-complete function offers a list of possible results.

fin

In the top-bar type name of the rz-doc you are searching for, click the “Submit” button next to it. Click on the Rz-doc you wish to open.

Select node/edge inside an Rz-doc



Option 1: Type the node name in the search bar. Press up/down to toggle between the different auto-complete options. Press return or click on the magnifying glass icon next to it to select it.

Option 2: Click on a node/edge in the map.

Once a node/edge is selected, they will have a yellow ring around it.

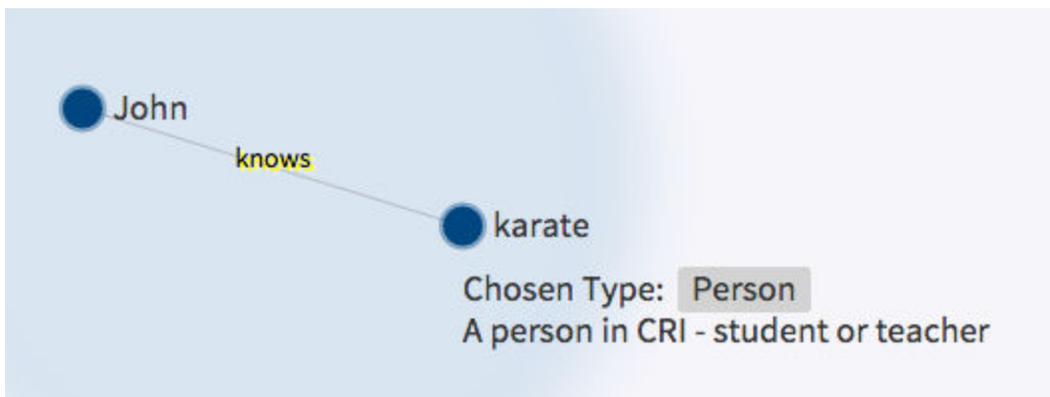
Select multiple nodes/edges

Shift+click on several nodes/edges in the map. You can select all nodes and edges by pressing control+A.

Deselect nodes/edges

Shift+click on a selected node/edge to deselect it. Click on an empty portion of the map to deselect all.

Node/Edge creation



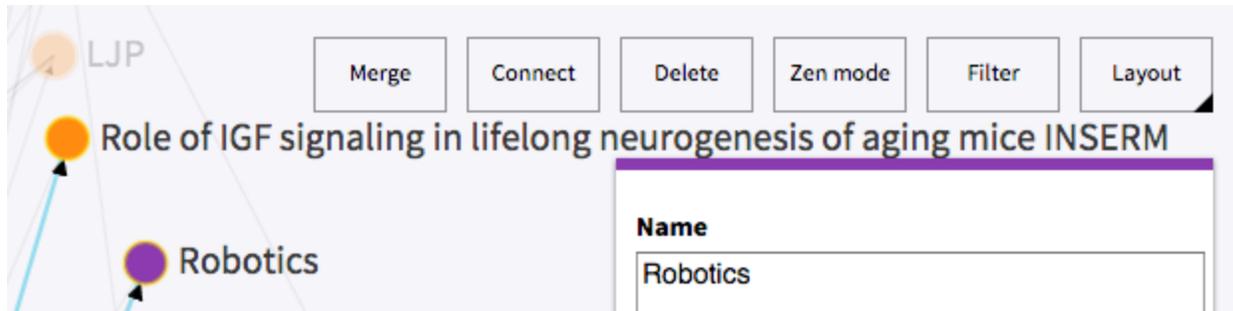
Use double-space to form a network. Typing: “John [double-space] knows [double-space]Karate” will generate a preview of the network:

To submit the network you have created, press return or click the “+” button next to the input bar.

Modifying Node/Edge type

While in preview: Press shift up/down to change node type.

Create only edge



Option 1: Using the input bar, type a sentence with two already existing nodes.

Option 2: Select two nodes with shift+click and press the “Connect” button. An Edge named “**untitled**” will be created between the nodes you’ve selected to the first node you’ve selected. *(A pop-up asking you to define edge title should allow you to set a certain title to all edges selected)*

Edit Node/Edge

Option 1: To rename node/edge title, click on the node/edge title as appears in the map, modify the text in the generated input box. Press return to submit change.

Option 2: Click on the node/edge in the map, A property box will open, change the appropriate fields in the property box.

Add thumbnail image to a node

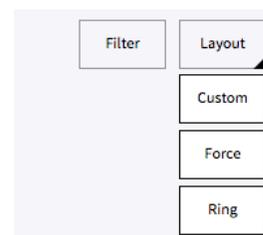
Click on the node and paste a valid jpeg or image file in the “Image-url” property box in the property menu. *(Should have upload + image crop option)*

Delete node/edge

Click on the node/edge you wish to delete. Click on the delete button in the tool bar.

Change layout

Click the “Layout” button, select the layout from the drop-down.



Changing node position

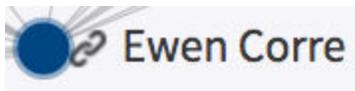
When in layout “Custom”, drag drop a node to fix its position elsewhere. When in other layouts, the position change is temporary, will not be seen by others and will not be saved.

Exploring content inside an Rz-doc

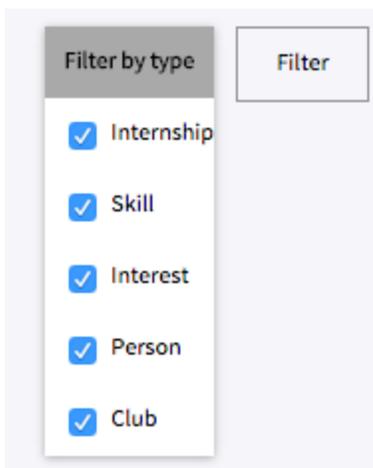
To view node/edge properties, click on node/edge. To close the property window, click on an empty space in the map.

Accessing Urls

To directly access a URL a node is pointing to, click on the clipboard button appearing next the node.



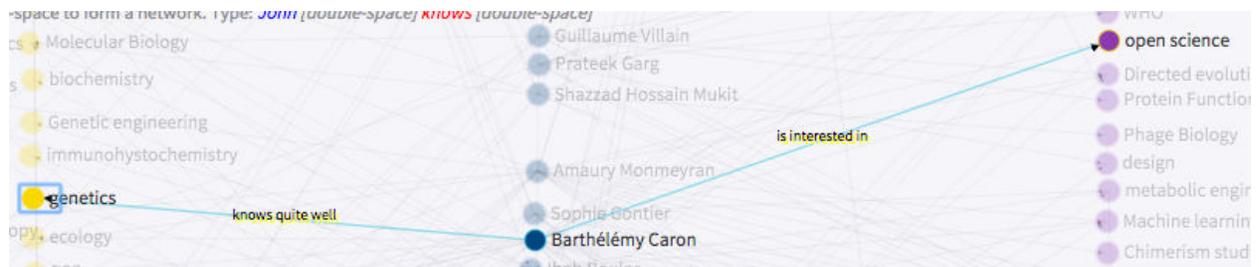
Filtering nodes



Click the “Filter” button. Tick/Untick the nodes you wish to filter. Another way to filter nodes is to selected one or more nodes and click the “Zen mode” button, this will filter out all unselected and unhighlighted nodes. To cancel “Zen mode” click on an empty space in the map.

View change log in a specific Rz-doc: Tick the “show activity log” in the top bar to open X number of recent actions made in the Rz-doc.

Find shortest path



To detect shortest paths between nodes. Click-shift to select two or more nodes to highlight the shortest path(s) between the nodes and the first-selected node.

Appendix 3: Bottom-up participatory online skills, expertise and project mapping within organizations

Dor Garbash^{1,2,3,*}, Alon Levy³, Ariel B. Lindner^{1,2,*}

¹ Centre for Research and Interdisciplinarity, Paris Descartes University

² Systems Engineering and Evolution Dynamics team, Institut de la Science et la Recherche Biomedicale (INSERM) U1001

³ Rhizi Inc.

* Corresponding authors: dor@rhizi.com, ariel.lindner@inserm.fr

Abstract

Members of organizations often suffer from lack of awareness about their organisation's knowledge capital - the competencies and project-experience of their colleagues. This communication bottleneck limits the innovation and collaboration capacity of the organisation as a whole. Information about knowledge capital is fragmented and often informal. This is aggravated for dispersed organizations that do not share a common physical space or that are forced to deal with a large variety of projects.

How can we lower the energy barrier for sharing knowledge capital within organisations? To address this challenge, we have reframed the problem as a human computation challenge. And so, we developed Rhizi, a real-time and collaborative resource mapping interface. Rhizi is a web application that translates text-based assertions made by users into a visual knowledge graph. The assertions are stored as structured data that is simple to query, explore and update. Final product of the process is a set of transparent "organisational maps".

We deployed Rhizi in an information technology (IT) department of a leading French University to qualify its impact. Department members used Rhizi to map and query its members, projects, skills and the time each person allocated to each project through a time reporting tool. Interviews and surveys conducted with the IT staff demonstrated the benefits, challenges and opportunities of the system. Furthermore, we have conducted several expert identification quizzes to show the contrast between knowledge with or without the final database produced by the IT department. Our pilot study suggests that expert identification capacity is increased considerably, has positive effect on collaboration motivation within the department and addressed some of the department self-described major issues. Our approach has potential to contribute towards flattening organizations and empowering individuals within them. This can be further strengthened by further simplifying the user interface, adding notifications and recommendations and increased network effect via deployment throughout the university. This pilot can further lead to its tailoring and use in other settings, both informal and sparse communities (*e.g.*, online citizen science initiatives) as well as organisations at large as NGOs, governmental and commercial entities.

1. INTRODUCTION

1.1 The challenge of shared awareness within organisations

Relationships between people, projects and skills in knowledge-intensive organizations are by nature complex and dynamic. It is often difficult and rather unpractical for a single individual (perhaps rarely, an

HR manager) to understand these relationships in detail, let alone harness their potential. Furthermore, Individuals within organisations are facing the natural limit to the size of their “organic” social networks. Dunbar’s number paradigm suggests the social limit to the awareness that one can have about others is around 150 individuals(Hill & Dunbar, 2003). This problem becomes even more acute when considering the increasing numbers of those that work remotely or at separate locations than the central office.

1.2 Knowledge management - state of the art

Knowledge Management is a broad term that is being used in different contexts for various functions (Scardamalia et al. 2010). In the specific context of this work we refer to the principle of treating knowledge as a community property that can be directly interacted with, rather than as siloed mental content inside individuals heads.(Scardamalia et al. 1994). We briefly address recent approaches taken to meet with the challenge of increasing awareness within organisations that influenced our own approach. For in-depth review on knowledge management see (Alavi, Maryam, & Leidner, 2001).

Much of the current knowledge management relies on multiplicity of IT tools to formalize existing knowledge in organisation that include wikis, organigrams, newsletters, HR directories, project management tools, and document repositories such as Google drive or Dropbox with varying access within the organization. These technologies often offer great convenience to individuals and can help organisations aggregate, analyse and share information more effectively than even before. On the other hand, they introduce their own set of constraints such as difficulty to acquire critical mass of adoption and time cost of contributing and accessing information (Garcia-Perez and Ayres 2010)(King, Marks, & Scott, 2002).

Expert profiling, skills and competency management is a subcategory in knowledge management. While it is relatively simple for individuals to identify experts within their immediate working environment or team, the process of identifying people from more distant circles of interaction is significantly more complex. Previous work to identify experts within organisations were based on methods such as keyword text mining(Dijk, 2009), co-word and proximity analysis(Criscuolo, Paola, Ammon, & Tony, 2007), social network analysis, citation networks and Linked Open data representation((Börner, Conlon, & Corson-Rikert, 2012) and extracting data from LinkedIn, emails and project management software (<https://corp.whoknows.com/>). These tools often serve an external researcher (auditor), or they abstract away so much of the actual representation, that the people who are being documented lose awareness of what is happening. In our previous work on mapping knowledge, [Knownodes](#) (Garbash et al, 2015), we noticed the acuteness of the problem. In the context of a university course, students who were a part of a team mapping research subjects to publications, had in fact, no good awareness that they were building a network together. This lack of awareness hindered both exploration and active participation in the construction of knowledge.

Network centric and flat organisations. A more holistic approach to the challenge of expert awareness is to replace traditional platform-based, hierarchical and top-down decision making processes with the paradigm of the network-centric organisation. Network-centric organisations are organisations that have established a communication network that allows their members to effectively communicate between them in order to facilitate agility and responsiveness to change. In these organisations, it is recognised that operators “in the field” have a better ability to make real-time decisions than those at the top.

Network-centric organisation therefore insure an IT infrastructure and culture that lowers the energy barriers for communication between the agents in the system (Crawford et. al 2009). The network-centric paradigm has received widespread adoption in military organisations and heralded as paradigm that “will prove to be the most important RMA[revolution in military affairs] in the past 200 years”(Cebrowski et al. 1998).

A key concept in Network-centric organisations is the creation of a shared awareness of the battlespace (in case of warfare)(Wilson, 2011) which increases the possibility of synergy leading to better decision making. Commercial companies may similarly wish to share shared awareness of the business-space. How should we adjust our sales tactics according to client feedback? This is typical of large organisations that have a relatively small and stable set of goals and need to deal with the shifting reality around them.

Flat organisations and practices are set to replace hierarchy pyramids by organic network of tasks rather than fixed titles and roles. Top-down management, where overarching knowledge retention is a power coin often only integrated at the top of the pyramid, is superseded by transparency. Thus, improving collaboration and access to experts and wisdom within organisations is a standing challenge for many such organisations (B. Stipelman et al. 2010) that adopt the concept of organisational transparency. Empowering employees to self organise through discussion and peer-review. More radical models such as Tiel and holacratic organizations provide evidence of the feasibility of flat, self-managed and transparent structures(Laloux, 2014). Such principles are common in volunteer organizations spanning from hackerspaces to NGOs.

Science of team science. The organisational context inside academic institutions is very different, characterized by an extremely fragmented and mostly decentralized environment with rapid transitions of students and researchers. Projects start, pivot, and end as new discoveries and problems are unveiled. Science of team science (SciTS) encompasses strategies to understand and evaluate processes and outcomes of collaborative research(Stokols, Hall, Taylor, & Moser, 2008). For example, providing evidence that demonstrates that impactful research comes from inter-institutional teams(Börner et al., 2010)

Human computation and resource mapping³. Modern human organisations experience a high throughput of projects and of humans that make the cost of tracking available resources and attention difficult. Human computation mostly focuses on the act of collecting information and mapping it so it can be processed by a computer, leading to an immense amounts of data could be in turn be structured by humans and computers.(Quinn & Bederson 2011) The human collaboration act of placing hyperlinks on html documents was the data on which google applied its famous PageRank algorithm to sort search results(Brin, Sergey, & Lawrence, 1998).

Human computation approaches for solving information fragmentation problem has a rich history such as using semantic-games to structure meta-data(Siorpaes & Hepp, 2008). Methods in human computation are often used to solve important scientific challenges such as diverse as protein folding(Cooper et al.,

³ A note about the use usage of the terms “map”, “graph” and “network”. While in many cases these terms are interchangeable, we have chose the term “map” to indicate any user facing functionality, and graph when discussing data storage in the database.

2010; Lenat, Guha, Karen, Dexter, & Mary, 1990)) and training AI's(Lenat et al., 1990). Knowledge Collection from Volunteer Contributors (KCVC) set the goal to advance artificial intelligence research by using humans to build large databases of common sense facts. Several efforts have demonstrated ways of using volunteer contributors to provide such facts, either by using games such as Verbosity(Von Ahn et al. 2006), 1001 Paraphrases(Chklovski & Timothy, 2005), or plain volunteerism such as Learner(Chklovski, T. 2003) and Open Mind Common Sense(Singh et al., 2002).

Graph visualization approaches. Graphs are data models where nodes represent entities and edges represent the relations between entities. Much work has been done in recent years to visualize relationships between entities in the form of networks. As noted by Jacob Moreno, one of the pioneers of the social network analysis approach: “A process of charting has been devised by the sociometrists, the sociogram, which is more than merely a method of presentation. It is first of all a method of exploration. It makes possible the exploration of sociometric facts”(Blake & Moreno, 1954).

Much work has been done in recent years to visualize relationships between entities in the form of networks, for the sake of brevity we will merely mention some that have directly influenced our work, which are: Gephi(Bastian et al, 2009), <http://linkurio.us/> and [Kumu](#) as tools for creating, editing and analyzing graph data using an interactive visual interface.

Limitations. Beyond technology, both cognitive and motivational concerns were put forth as limiting knowledge sharing. Cognitive issues often follow from the different levels of abstraction used by expert and novice to describe a given system. Motivational limitations arise from multiple sources, including (i) in-house competition that sets organisational units against each other, competition for promotions and raises, measuring performance against each other and tendencies for employees to identify themselves with the unit or work team they are on leading to in-group bias (Abrams and Hogg 1990), (ii) the burden of rigid, explicit protocols for knowledge sharing that disincentivize their use as well as lack of incentive (direct positive feedback) to overcome the perceived cost (Hinds and Pfeffer 2003). (iii) sanctions and norms against sharing across hierarchical boundaries, lack of trust in the organisation. In this context, Hind and Pfeffer suggest that mapping expertise rather than knowledge is highly advantageous as it facilitates interpersonal connections around interest foci as opposed to disembodied, decontextualized knowledge.

Mapping knowledge in learning organisations. To this end, we focused in this work on harnessing human computation towards enhancing creativity and learning within organisations by providing them with a bottom-up capacity to map the expertise within them. We refer to Learning Organisations as defined social entities that are based on identifying people with knowledge, on learning opportunities and on conducting coordination of training activities. Creativity in such organisations consists of building awareness for the different opportunities and resources available and how they are interconnected in unique ways to each other (forming a denser network) allowing discovery and exploration of creative

opportunities. Therefore members of the organisation become both mappers and explorers of their organisation's opportunity space.

We designed a novel web-based tool, Rhizi, as a system that harnesses the processing abilities of humans to create and update labeled connections between people, interests and skills. This enables computation of shortest paths to facilitate direct interaction between members of the organisation and provides a transparent and efficient mean to connect outside people to relevant members of the organisation. These paths could not have been computed without the restructuring of the data by humans. Rather, humans and computers are co-paving paths for expert awareness. Rhizi has elements of crowdsourcing as it has a level of human choice, but the end result is based on computed queries and their representation. Using Rhizi, humans create a mass of computational data, and while there is an element of pre-existing tasks, the value is often in the ability of each human to ask individual questions regarding the data. The results are relevant to the community interested to "know itself" and can be aggregated into larger groups (the university body, the research body).

2. Characterization of challenge within a concrete study case - University IT department

Our overall objective was to lower the energy barrier for sharing knowledge capital within organisations. We focused on improving the expert identification process, and the knowledge fragmentation problem that stands at its root. Knowledge fragmentation happens when relevant pieces of information are stored in silos and therefore are difficult to aggregate together. This can take shape as informal knowledge about experience or skill is distributed across people's heads ("Carl knows how to fix the copying machine"), as unstructured information such as cv's, emails and documents, or even as structured data that is stored in disparate databases, spreadsheets, and project management software.

Accordingly, the challenge organisations face when defragmenting knowledge is two-fold: From the input side, how to efficiently input and update the skills and project-related experience of individuals? From the output side, how to reduce the cost of searching and sorting this information?

In most organizations, deriving these insights requires manual and expensive human intervention or a plethora of mostly proprietary softwares.

2.1 Characterize existing expert identification issues

As a case study we chose to work with an IT department of a large university. The IT department services for the needs of ~4000 researchers, ~50,000 students and engaged within a dense network of cross-institutional R&D development. With 114 employees, 120 ongoing projects and more than 100 collaborating institutions, only several managers possess an internal model of who is an expert in what field and which projects which person is involved with. Because of the scale and variety of its responsibilities, we find that it exemplifies many of the organizational challenges facing knowledge-intensive organisations.

We first started by engaging in conversation with the head of the IT department, and define the main issues that are both important and practical to solve. Following these discussions we have set three sub-objectives that follow a complete cycle of product development: First, to characterize existing issues that the staff of the IT department was experiencing when trying to identify experts beyond their immediate team. Second, to test the feasibility of our mapping system by deploying Rhizi within the

department. Lastly, to gather feedback and conduct several tests to gain further insight and chart the future development of the system.

Our analysis of of the IT department lead to the identification of the following key problems:

- (1) *High perceived cost for finding and contacting peer-experts*: Where the main causes for the perceived cost are:
 - Having to Involve managers, resulting in time loss for both requester and manager as well as potentially appearing less competent in front of the manager.
 - Failure to find the expert because manager or colleague were not aware of the expert.
 - Loss of productivity as new and remote employees take a long time(1-2 years) to collaborate effectively with other department members.
 - Lastly, outsiders have no way to know who to contact when needing assistance.

(2) *Lack of “big picture” understanding of current projects and the involvement of others in them.* This stems from the fact that data about employee involvement is distributed across multiple databases and application without synthesis of the information in an easily accessible way.

Following this we have set the following goals for the Rhizi mapping system:

1. Achieve a real-time representation of relationships between project, people and skills inside the IT department.
2. Develop a transparent and collaborative platform that will enable IT staff to identify experts and explore the organisational map.
3. Work towards a transparent representation of the department’s projects and skills that would enable even further collaboration, down the line, with other university departments and outside collaborators.

Together, we have decided that identifying sources of intelligence rather than making the knowledge explicit like in a Wiki or a knowledge base is preferable because the work of relating people to skills and projects is considerably cheaper than the creation and maintenance of a knowledge base. Incentivizing people to share their skills and keep it updated, however, is vital. This is why we thought it critical to provide an engaging input interface for the knowledge graph.

Lastly, we have articulated together the desired outcome of our collaboration: To empower staff to see the big picture, freeing individuals within the department to collaborate without relying on intermediaries. Provide staff an independent way to gain awareness of the collaboration opportunities available to them.

2.2 Gather feedback and validate our solution

Based on our design goals, we expected that the use of Rhizi would affect users’ experience along four dimensions.

- (1) Lower staff’s perceived cost of collaborating with others by provide a way for people to identify and contact experts
- (2) Improve awareness to the overall organisation and their relationship to others.
- (3) Motivate employees to update their information in system.
- (4) Assist management to evaluate the needs of future projects and formulate training sessions accordingly.

We specify in the methods and results section 7.1 the feedback received and its analysis.

3. Design approach

Following the goal-setting with the IT leadership, we defined principles we wanted to follow while implementing the system. Most importantly, we wanted to the system to communicate to the IT members that we see them as “citizen scientists” of their own organisation. Framing the problem in such a way that the staff is aware of what they are doing and building. We wanted to create a individual-centric model for collaboration, so an agent can change and modify his skills and experience dynamically, immediately affecting the organisation's’ perception of its affordances.

Having each employee spend a few minutes mapping his own local perspective is a lot more efficient than one person “doing the rounds” and interviewing everyone and then transcribing it into a database. By empowering people to map together, we get a comprehensive understanding of the whole system, giving us not only insight, but also a human feedback layer to update this information as things change.

These principles influenced us to add the following requirements to Rhizi:

1. Collaboration features - Users have unique accounts from which they can add their contribution, each user can create and edit entities in the map.
2. Transparency - User actions are public and logged. All maps and content are accessible to anyone with a registered account.
3. Real-time - Once an actions has been made, it is immediately visible for other clients.
4. WYSIWYG(“What You See Is What You Get”) interface - Similar to the relationship between a table database and a Spreadsheet, every action on the network is represented directly as a network visualization.

Another design choice was to use the metaphor of a map to describe the actions taking place in the system. Mapping is an abstraction used to navigate physical space, and we think the same abstraction is a useful one to familiarize users with the concept of navigating resource and opportunity spaces.

These principles represent how our approach differs from existing expert mapping and network visualization platforms. Rhizi is designed to be used and understood by all the participants involved in the mapping process, rather than subjects of a scientific enquiry, from which data is sourced, later to be exclusively analysed and and valorized solely by managers and analysts.

4. Rhizi system overview

We translated the high-level requirements listed above into technical features and design requirements. The features are divided into two parts: Front-end focusing on interaction design and back-end focusing on data structure and storage. When necessary, we added additional information about the implementation.

Also, For a more dynamic overview of Rhizi’s interface, see a video tutorial for a different project involving Rhizi:

<https://www.youtube.com/watch?v=n4pMaG24q78>

4.1 Front-end

Text-to-network Input: Input in Rhizi utilizes a novel text-to-network real-time visualization. As the user types the output as it would be represented in the map is previewed in real time. If a user mentions in his assertion an entity that already exists he is able to see it's current location and its context - the other entities connected to it(see figure 1). To differentiate between nodes and links users use double-space in their assertions.

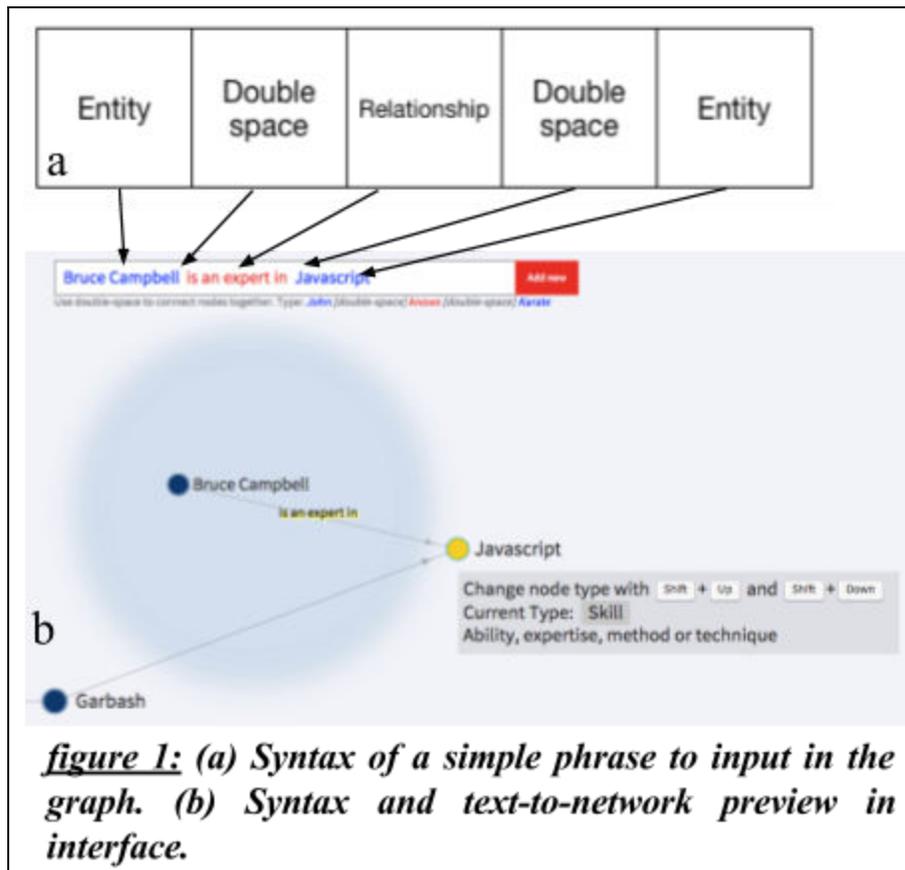


figure 1: (a) Syntax of a simple phrase to input in the graph. (b) Syntax and text-to-network preview in interface.

Search: Users can find and select specific entities in the map by using open text search.

Node and link selection: Beyond using search to select nodes, users can click with a mouse on the nodes and edges directly in the visualization. Users can also select multiple nodes and links.

Edit operations on selected nodes and links: Once a selection has been made, users can select to delete, merge and connect nodes together. Merge allows the user to merge entities that in fact represent the same thing. This tells the system to remove one node, while moving its links to the node it is merged into.

Shortest path: Once two or more nodes have been selected, clicking on the Shortest path button will highlight the minimal distance between the original node and the nodes selected.

An example of the usefulness of this function is when a user wants to know how he can reach a specific expert and the system can highlight specific people that can introduce him to the expert or shared skills and expertise they might have in common.

Zen mode: Zen-mode is a dynamic way to explore the map. It filters out all nodes and edges that are not highlighted or selected. By selecting and deselecting nodes, the user can travel along the map.

Node properties: Each node type contains properties as defined in the data schema. These properties are editable. There are several special properties that have specific effects: URL property adds a hyperlink icon next to node in the map interface that links directly to the Url's target. Image-url property will display a thumbnail of the linked image.

Filtering: Nodes can be filtered via a dropdown menu according to type.

Navigation between maps: Users can navigate or create maps by using the “New Map” or “Open Map” buttons. Each map contains a subset of the nodes and links in the database and represents a separate mapping task.

Layout: Map layout can be modified by changing the layout. The three options are: (1) Force layout (Fruchterman & Reingold, 1991) (2) Ring layout - arranging the graph according to concentric rings according to node type (3) Custom layout - which is a manual and collaborative arbitrary placement of nodes according to users (See figure 2). This allows users to get different perspectives on the data.

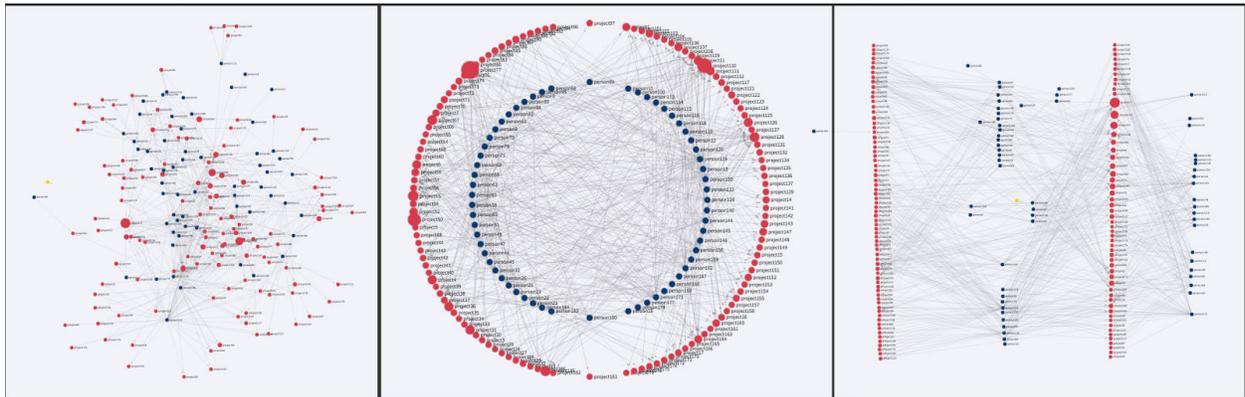


Figure 2: The different visualization layouts in Rhizi: Force layout, ring layout and custom layout.

Free assigning of connection types: From previous experience with earlier prototypes, we didn't want to burden users with selecting from an ontology of existing relationship types. We also didn't want to disrupt the flow of using our text-to-network interface and so we decided to allow users to express relationship types via freeform text.

4.2 Back-end

Graph database backend - data inside the system is stored in the graph database Neo4j. Making it easier to use graph based analysis, algorithm and applications.

Data model: Data is modelled according to the schema in figure c. Inside the Database, we have two separate graphs: A graph depicting the relationships between the entities, which have a 1-to-1 relationship with the representation shown through the UI. A second graph which represents the different maps (referred here as Rz-docs) and which specific nodes and links appear in which map and which are shared between maps.

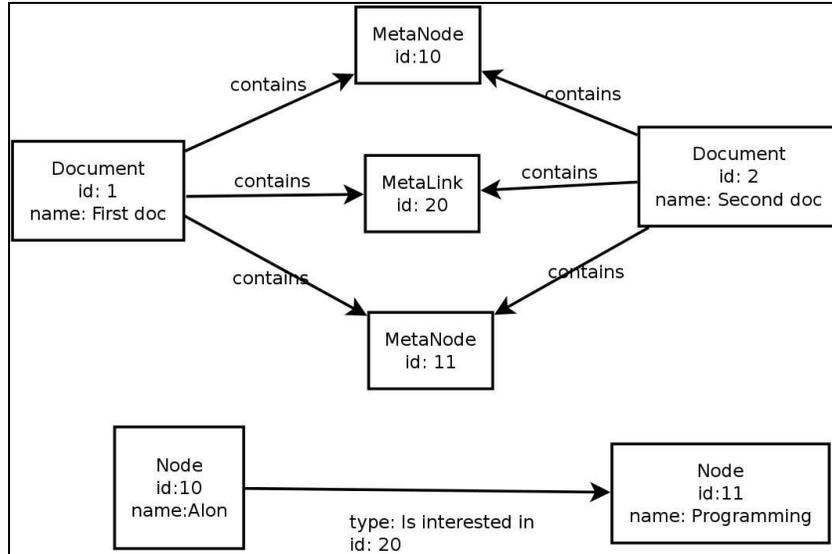


Figure 3: Rhizi data model

Real-time collaboration: Using Websockets and TCP connection, changes in maps are updated in client in real-time.

5. Implementation and usage in IT department

5.1 Data Structure

The Rhizi data schema was adjusted for the IT department according to the following table:

Node types	Node properties
Person	name, email, image, URL, description
Project	Name, Description, URL, Status, Start date, End date, country, city, address, facility
Skill	Name, Description
Keyword	Name, Description, URL
Organisation	Name, Description, URL
Media	Name, Description, URL

Table 1: Node types and their properties as defined for the IT department use-case

These types were selected to offer a good balance between structure while not hindering the user experience.

To avoid over complexifying the maps, we decided to separate the work into two maps: an Expert map and Project map. The Expert map depicts relations between staff, specific technologies, technical skills and organisations.

5.2 Sources of data within Rhizi

The data within Rhizi originated from the following sources:

1. *Direct input by IT management during interviews with staff members.*
2. *Direct input by a subset of participants who added or modified data by themselves.*

IT personnel officially started self-mapping with Rhizi one month after an initial meeting which introduced them to the tool. The self-mapping task consists of reviewing, correcting, and adding information about their skills, roles, and relationships with their peers. They were also invited to explore the data already available in order to familiarize themselves with the competency landscape of their organization.

5.3 Usage in the department

We provide several examples of how the IT department used the tool. For the sake of anonymity, we have replaced names of projects, people and organisations with numbers. Usage period was between 1.1.2016 - 1.4.2016

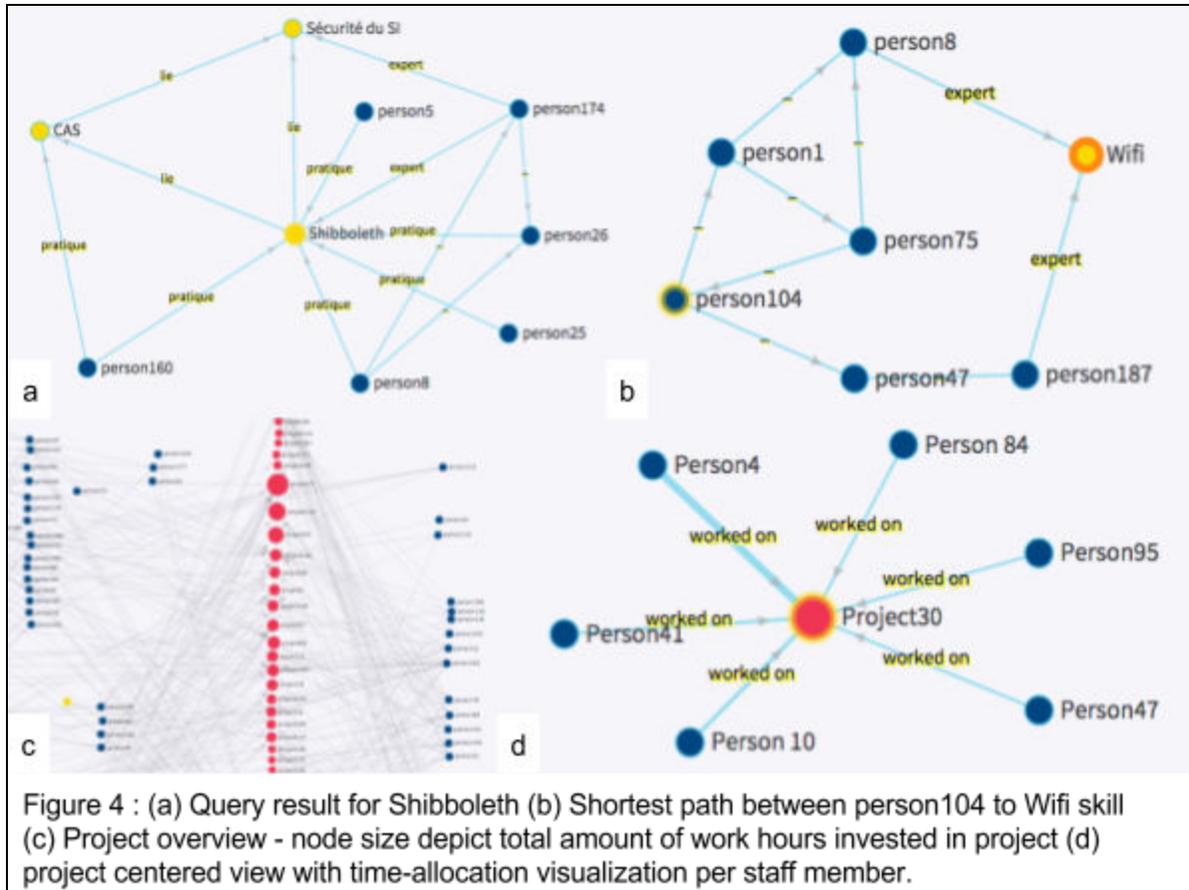
Finding competencies: Using the search interface, any staff member can type a skill he is looking for. In the example below - the employee queried the term “Shibboleth” - an Identity management technology. The result is two technologies identified as related, and several people of varied degrees of expertise. “Pratique” as in a normal practitioner and “expert”(figure 4a).

Shortest path: A person looking for Wifi expert can find multiple interlockers to connect him to a Wifi expert(figure 4b).

Project overview analysis: We have imported this data into the Rhizi system as a network of relationships between people and projects, where the width of the connections between people and project represents the amount of time they have invested and the size of the project node represents the total amount of time dedicated to the project by all employees.(figure 4c)

Staff experience: Selecting an individual staff member in the Project map, provides a visualization of how he allocates his time between projects. Link is thicker the more time is allocated.(figure 4d)

Selecting a project: shows which people are most experienced in it.



6. Experimental setup

Our study used two primary methods. First was a task designed to test whether finding relevant experts was facilitated by Rhizi, making the search faster and more effective. To this end we used an in-person expert identification quiz asking IT staff to identify experts and then re-ran the quiz with subjects who are not part of the IT staff, with the assistance of the Rhizi software. The second was a qualitative inquiry designed to validate our characterization of the problem in context of the study case, and to understand the experience of the IT staff using Rhizi. For this we conducted an online survey sent to IT staff, and a semi-structured explorative interview about the staff's current practices for finding experts and feedback about using Rhizi.

6.1 Participants

The IT department is responsible for providing IT services to a university with ~35,000 students. It has 114 employees, ~80 listed skills, ~120 ongoing projects <100 collaborating organisations and institutions. Employees are distributed across multiple offices across the city and a small percentage of them ~10% work remotely.

Each methodology was composed of a different cohort from the total of 114 IT Staff members:

The online survey was filled online and offline by a total of 19 participants. 7 of whom were women, 7 worked remotely. With regards to roles: 8 managers, 8 technical and 3 administrators. All worked for longer than a year in the department.

The expert identification quiz was filled by 13 staff members, 4 of them were women, all worked on premise, average seniority of 5.1Y, average age 39.

Semi-structured interviews were conducted with 8 staff members, 1 of whom was a women. None worked remotely. The jobs of the 8 were as follows: Technical assistance, IT administrator, System engineer(X2), webmaster/graphic designer, and manager(X3)

6.2 Online survey

In order to validate the problem of expertise identification we have sent an online survey designed to examine which, if any, aspects of peer collaboration and identification are important to members of the staff. The survey consisted of a series of statements which respondents could rate on a 7-point Likert matrix table (Krosnick, Lavrakas, & Nuri, 2000) to collect feedback about staff's experience with using Rhizi. Invitations to participate were sent through email. Out of 114 employees contacted 19 filled the survey. To make the result presentation more clear, we have merged the 7-point into a 3-point Likert matrix (Agree, neither agree nor disagree, disagree). We have also asked respondents a series of open-ended questions to provide us with some feedback about issues they had and suggestions for improving the system.

6.3 expert identification quiz

In order to understand how much people can identify and effectively direct others to an expert, we composed a quiz asking employees to identify a specific person who can help us with 10 requests for help (**Annex A**). These help requests require expertise in a certain technology. A secondary objective was to understand what happened when the subject didn't manage to identify an expert by name. The expert identification quiz were asked in-person in direct dialogue with the subject and recorded as audio.

6.4 Semi-structured Interviews

In contrast to the very structured expert identification quiz and survey, we used semi-structured interviews(Wengraf, 2001) to explore deeper into people's experience using Rhizi. This helped us learn more about their thoughts about collaboration, giving and receiving help and ideas regarding future implementations of the software. We also thought it would be a good opportunity to understand better how these subjects relate to their daily work-life challenges.

6.5 Rhizi-assisted expert identification quiz

We wanted to test if Rhizi provides better results for expert identification than asking employees. To this end, we have asked 5 people (out of which 2 women) who are not involved professionally with the IT department to use Rhizi to provide answers to the same 10 questions that were given to the IT staff. These sessions were made in an informal setting. It included a 2 minute demonstration of how to use the search function to search for skills inside Rhizi.

6.6 Data Gathering and analysis

In total we received 140 distinct responses to the expert identification quiz (10 responses from each participant), and 50 responses to the Rhizi assisted expert identification quiz. We also obtained 416 minutes of video recordings from the semi-structured interviews. Drawing from the interviews, we coded the responses (Burnard & Philip, 1991) to each question and tried to identify common and notable themes raised by participants. With the expert identification quiz, we coded responses when the participant wasn't certain who the expert is.

7. Results

7.1 Low ability to identify experts as an underlying bottleneck for interdepartmental operations

We characterized expert identification issues and identified how they relate to the more general challenges the organisation was facing. The following subsections cover each one of our findings.

Self-reported need for expertise and project awareness : Survey results confirm our hypothesis that the IT staff indeed thinks that having better knowledge about projects, skills and collaboration are needed. The usefulness of being able to quickly identify an expert was also validated.

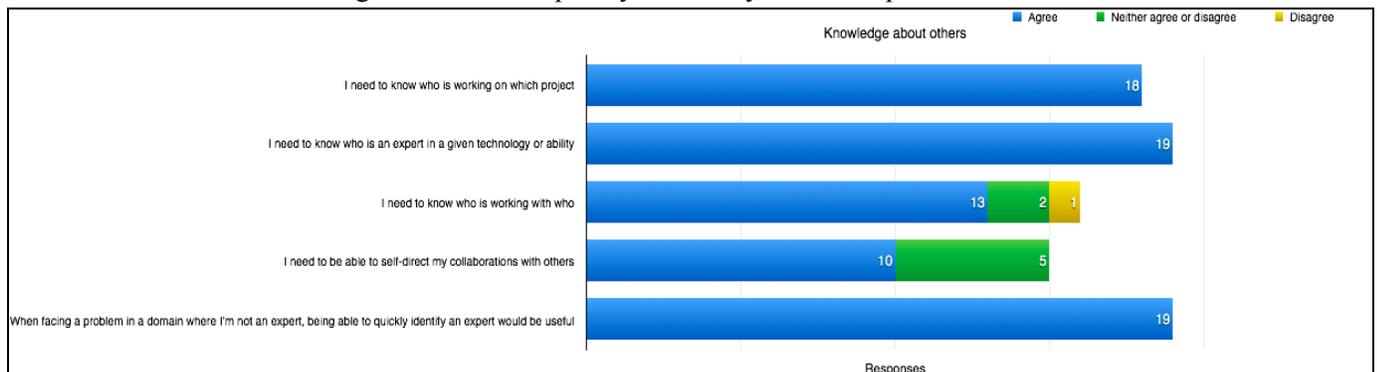


Figure 5: Survey results on questions related to needs related to expertise and project awareness

Expertise identification is cumbersome and inefficient in current practice - In contrast to the needs stated above, staff had difficulties to identify experts throughout the department. We have composed a 10 question quiz. Each question was formulated as a request for help in a specific technical task. We asked the staff members to identify by name specific experts who they think can help us with the request. Running the test on 14 subjects, a specific name of a person was given only 44.6% of the time (58 out of 140). We didn't allow the staff involved to look at Rhizi for the answer, nor to ask their colleagues, they were free to use any other tools.

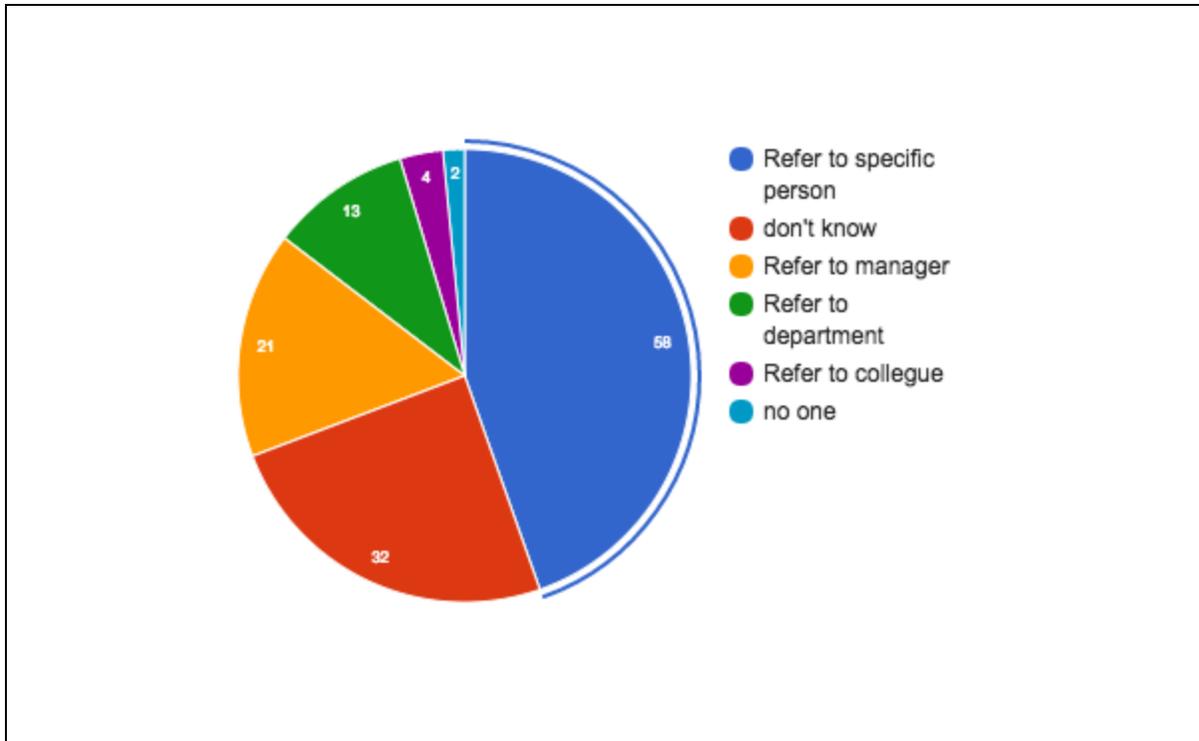


Figure 6: Results of the expert identification quiz.

We think these results indicate that expert identification is indeed an existing problem within the department. Average seniority of the recipients was 5.1Y and so they had plenty of opportunities to become familiar with other people’s activities through the bi-annual meetings and other work related communication activity such as emails and newsletters.

The large amount of “don’t know”(32 responses) and referrals to managers(21), or departments(13) also means that if a hypothetical person wished to casually ask someone for help and consulting on an issue, they were forced to initiate a more formal process of asking administrators or managers. Small number of colleagues referred(4), might suggest expertise silos, where staff can’t identify experts and often doesn’t know who can identify them. This confirms our hypothesis about the high perceived cost expert identification. See section 7.2 for validation of the Rhizi-assisted solution of this issue.

New and remote employees have a higher barriers to identifying experts: As one technician mentioned during his interview: “A new employee arrives to a local team, that describes to him their local functionality, but the global functioning of the university is very complex, and impossible to ascertain when one arrives.” On another interview, the Manager of the IT department said that having a new or remote employee understand the “who does what” in the department may take between 1-2 years.

Project overflow due to bureaucratic bottlenecks - 4 interviewees (3 of them managers) emphasized having too many projects in proportions to allocated resources: “IT Administration is like a big company, we do the job like a big company but not with [big company] money”, “We begin a lot of projects but not finish them or after time”.

Reasons for this project overflow, such as budgetary constraints, are outside the scope of this research to address. However, two relevant underlying reasons for the expert identification problem were brought up. The first is that projects were getting delayed because of “weak links” in the bureaucratic chain: “You launch a project and it can have big delays on different reasons but mainly because administration and a lot of people that are not always involved in their work in this big administration”.

The second reason given was long delays in cross-team communication. One staff member gave an example of how a solution that should only take several hours to implement, took in practice four months to complete. “I made a request, I waited for a reply. One week, I was told, no problem, we will do it. One week later, we are in the process of doing it, how can we install the thing. I said to them, I can do it with you. ‘We don’t want you to do this with us.’ So how do you want to do this?, one week later the person in charge says, ‘ok, let [subject’s name] work with you to show you’, then nothing for a month or two, and then, ‘ok, [subject’s name], you can come over.’ they have set an appointment, for two weeks later. Four months”. When our own researcher tried to coordinate meetings within the IT team for interviews, emails were often ignored, and an alternative system of directly phoning people through their personal mobile device was often the only way for him to establish contact with the subjects.

Lack of coordination on interdepartmental projects may cause sub-optimal or duplicate solutions: “Methods are not well shared, there is a loss for the organisation, because they don’t have a record who already did such a thing. So the project is successful because it is here, but it didn’t have all the benefits it could have because it wasn’t shared globally. All project are like this, it’s really too bad that we don’t share these things. It’s a loss that is hard to measure, but nevertheless exist, because no one can retake on this project tomorrow, “ and so, it’s not only that projects take longer, but there are often duplicated efforts and sub-optimal solution due to lack of communication between the experts.

Lack of expert coordination can also lead to a “Reinventing the wheel” phenomena: “We now have one [technology name removed] platform, but it used to be that each one had it’s own [technology name removed]. In [name of department removed], each one had it’s own platform. It’s very inefficient, We see that people are not directed. It’s just in IT, it’s everywhere. The [redacted faculty name] asks to do something, and it’s separated when the university wants to do something, it works like that, it’s not just IT.”

This is in line with earlier reports as a 2014 survey of 152 Campus IT leaders in the US, that estimated that 19% of the campus IT systems are redundant, costing US universities 3.8B\$ per year(Cloud campus 2016) .

Need for expert collaboration increases in times of rapid change, or working on cross-departmental projects: In times of change collaboration becomes crucial. “When we all face the same problem on the same time. When there’s something new. When there’s a new software to install, there is a period of time where knowledge is not well structured.” Another employee initiated and led a bottom-up effort to create a platform that is serving several university departments and requires collaboration from different and diverse fields of technical expertise. He had a lot of trouble identifying the right collaborators: “I had to

meet many people to do that, but could never manage to find the time to find out who did what, who knows what and who could help me ‘carry’ this project.”

Conclusion: Difficulty to identify and directly communicate with experts leads to project overflow and lack of coordination. High perceived cost and uncertain results of trying to manually identify experts leads people to try to solve issues themselves. If the expertise is necessary for the continuation of the project, this can serve as a bottleneck that leads to unnecessarily delays. A synthesis of the barriers for effective cross-team collaboration can be seen in figure 5. The perceived cost was well established from our quiz and interviews. The second part, lack of incentives for collaboration is covered in the literature review in section 1.2 under the *Limitations* sub-section and is further covered in the discussion section.

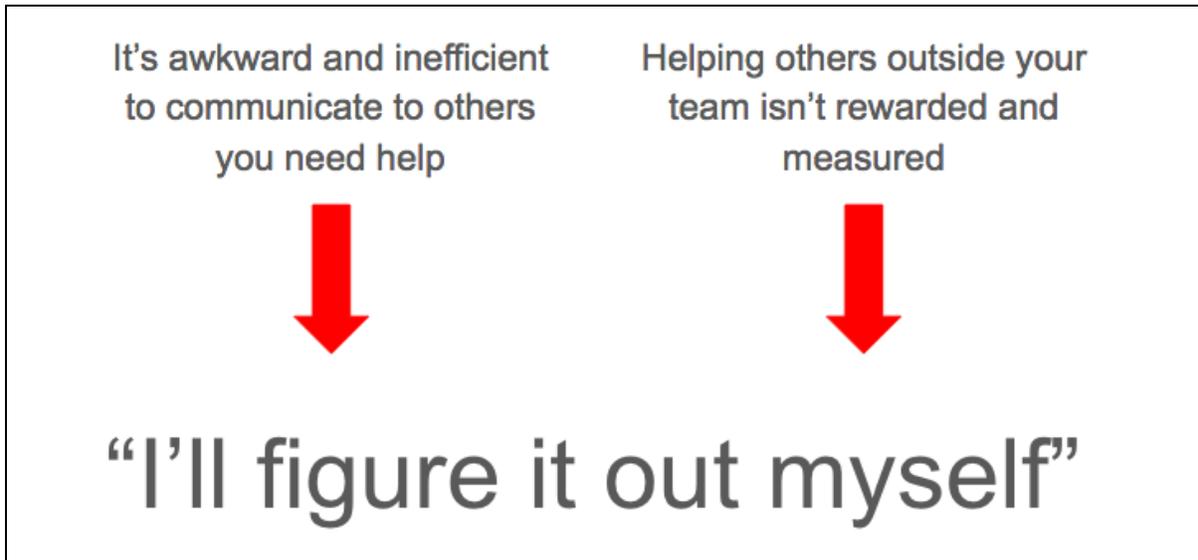


Figure 7: Simple breakdown of perceived cost and lack of incentives for organisation wide collaboration.

7.2 Validation of Rhizi as an effective solution for expert identification

Creation of a transparent visual database of expertise and project experience: Two maps were created as a result of the study case. A “Digital experts” map containing people, organisations and expertise that mapped not just the department, but key people in relationship with the department's staff. The Project map is sourced from a time reporting tool, and makes it explicit how many days each person dedicated to each project. The following table (table 2) represents number of entities mapped:

Map name	People	Projects	Skills	Organisations	Links	Edges/Nodes ratio
Digital experts	218	-	155	91	1309	2.82
Digital projects	61	125	-	-	542	2.91

Table 2: Number of entities in each of the resulting maps, according to Entity type. (True for 04.15.2016)

Employees had transparent and formal way to know who’s doing what and who possesses what skills. This allows any employee to make a informed decision about who to contact in case of blockage or if a consultation is needed.

Rhizi allows successful identification of experts: To validate the solution, we have redone the expert identification quiz questions described in section 6.5 with 5 random participants who have no association to the IT department. The outsiders success rate of naming specific expert was 100%, and often their recommendation included a list of alternative experts if the first expert could not be reached. The process took 2m on average, where the employees in the original expert identification quiz took on average ~9m to answer the 10 questions.

Usage experience beneficial for motivation for sharing expertise: Overall, survey results show that the maps created had a general positive impact on people’s ability to identify experts, and motivate staff to share their expertise and reach out for others .

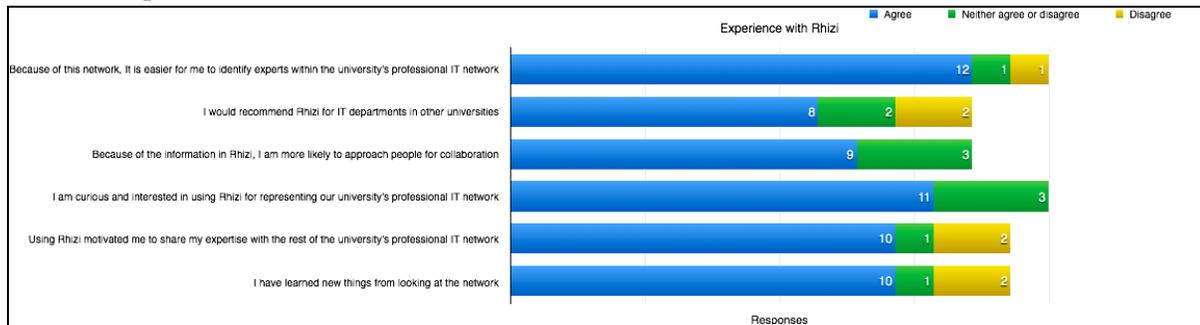


Figure 8: Survey results on questions related to experience with Rhizi.

As one respondent said in the interview: “When I started my project, I looked for people who knew [web technology]. That could help me to start. I did it by myself. Now, I look first in Rhizi.”

Understanding the organization's' direction and competency gaps : Several respondents said that the maps helped them understand where the organisation as a whole is advancing to: “I like it because I can imagine what is the direction we are taking in the university, mobile or this service or pedagogy” For another respondent the system-level view of resources allowed him to identify vulnerabilities: “[name of staff member] leaves in three months, and I know that all her competencies will disappear, that will allow me to see which people I should look for to replace those. As a manager it’s interesting, if I see the project she is working on, I can make sure she passes the information to those who replace her.”

In the digital experts map, we have identified several skills that were weak-points, indeed it would be simple to see if an employee leaves, which skills remain as “orphan nodes” and make sure to train or hire for these skills. Out of skills listed 50 were identified as weak-points having only one person marked as an expert in them.

The Project map, out of which 25 project had 1 staff member logged as involved with.

7.3 Shortcoming and further development routes identified for the Rhizi software development/deployment

Negative criticism towards the network visualization interface: The major negative criticism about the system was the complexity of the network visualization interface. While heads of the IT department quickly learned how to use it, most casual users didn’t like the unfamiliar interface, and the complexity in the map. This was also repeated when we asked as an open question to describe negative or frustrating

experience with Rhizi. This also reinforces the fact that *despite* the complex interface, people still found value in the underlying information.

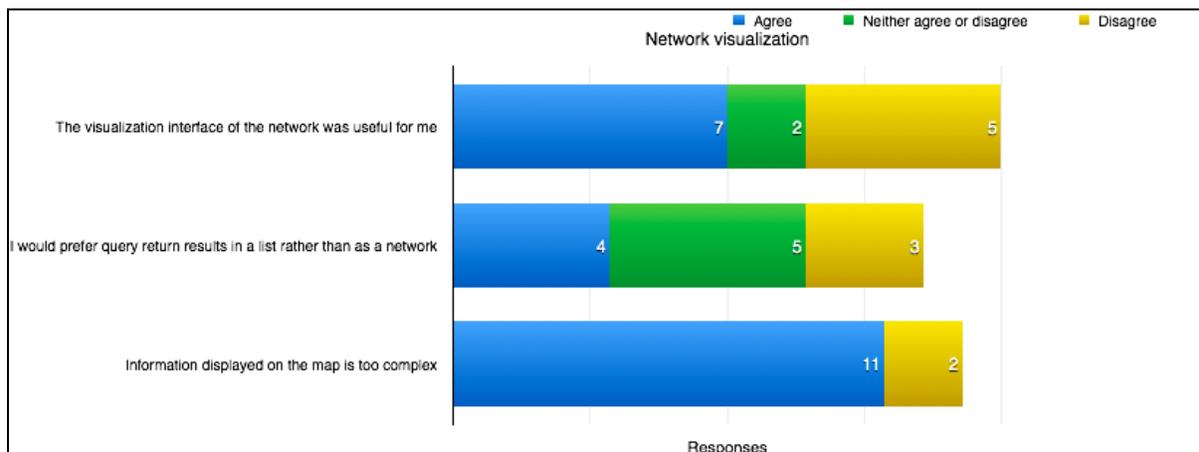


Figure 10: Survey results on questions related to network visualization interface.

Direct communication between people: Two respondents mentioned that direct contact between people without mediation of bosses or hierarchical structures as something that would be beneficial for the entire university organisation. “Who is the person I can contact I can see the organigram but you don’t know which person is better to contact. You can’t contact the super boss because his jobs is not to answer demands. So, direct contact of people”. When interviewing others about using Rhizi as a communication system, response was overall positive, especially the idea of subscribing to certain domains or fields of expertise and seeing requests for help based on tags rather than a direct contact with another person. As a next step, we would like to add the ability to contact respondents based on their fields of expertise or project involvement.

Knowing something is not the same as being motivated to use it to help others: Some respondents expressed a wish that only a subset of their skills will be public. “We all have previous lives, and probably you don’t want that people know. I don’t want that other people know that I did some shell scripting during my Phd, or understand how to build a network of TCP/IP. I’m not a computer guy. I like photography, I’d like to put that, because it can create a connection with other people who share this.” People are more motivated to help in a subset of the skills that they find interesting or relevant to their career trajectory. Other criteria for likelihood to help others included understanding the value of the help to the recipient, like knowing that an issue is urgent or blocking, or that the amount of work required isn’t too large. We think that mentioning skills people are not motivated to help with, will reduce overall trust in the system, as people will be slow or avoid responding to skills they are not motivated to help with. This should be reflected on the input and onboarding interface of the software.

What are people interested to know the most: Four people mentioned they are most interested in finding people who share their domain of expertise : “I would find it super nice if you could see who are doing the same kind of job and how to contact them and what they are working on so we could exchange information.” This supports that people are really interested in forming expert cohorts within organisations. This is further discussed in the discussion section of this paper.

8. Discussion

During this research project, we have discovered that expert identification has far-reaching consequences for the entire organisation: Bureaucratic bottlenecks, unnecessary project delays, and redundant solutions. We validated the problem, and validated that explicit and transparent mapping of people, projects and expertise can help even outsiders identify who can help on given issues and motivated people to discover and share expertise within their department. This improves access to the organisation's knowledge capital across the entire employee-body. A separate finding is that while the network visualization was helpful for people to understand they were all part of an expert network, it has also obfuscated some of the information because it was unnecessarily complex and forced casual users to adapt to an unfamiliar mode of looking at data.

These findings can further the field by acting as a proof of concept of how human computation can help with transition from collaboration practices of hierarchical organisations into a more flat and/or network-centric models. Related to known knowledge-management solutions, but not requiring the overhead of having to update a large repository of knowledge, as well as giving each person the ability to update his expertise as necessary. This could be of value to team science, because of the high throughput of humans and projects, and indeed we are planning to apply this solution not only for university IT, but for research communities as well.

Our main takeaway towards building the next generation of our collaboration tool is that we'd like to create a system to directly contact experts. This requires developing strong integration with email and other communication technologies such as real-time chat, video, and phone conferences. Another important issue is providing means to measure the impact of helping others and forming a reputation system that rewards those that unblock cross-departmental issues or save funds through mentorship and consulting.

Empowering people to directly approach experts within their organisation, can be part of the solution that both reduces the time each project takes while encouraging innovation and system-level thinking. More importantly, if university administration, governments or any large organisation want to be an attractive destination for innovative people, they must ensure that people can break down inter-team barriers and get the collaboration they need in a timely way.

Acknowledgments

This work was supported in part by Les Laboratoires Servier as a PhD fellowship to DG. We wish to thank the director and team members of the IT department under study and Amir Sagie, Ofer Lehr and Owen Cornec for contribution to code and design development.

Bibliography

Alavi, M., Maryam, A., & Leidner, D. E. (2001). Review: Knowledge Management and Knowledge

- Management Systems: Conceptual Foundations and Research Issues. *The Mississippi Quarterly*, 25(1), 107.
- Bastian, M., Heymann, S., & Jacomy, M. (2009). Gephi: an open source software for exploring and manipulating networks. *ICWSM*, 8, 361-362.
- Blake, R. R., & Moreno, J. L. (1954). Who Shall Survive? *Sociometry*, 17(1), 77.
- Börner, K., Conlon, M., & Corson-Rikert, J. (2012). *VIVO: A Semantic Approach to Scholarly Networking and Discovery*. Morgan & Claypool Publishers.
- Börner, K., Contractor, N., Falk-Krzesinski, H. J., Fiore, S. M., Hall, K. L., Keyton, J., ... Uzzi, B. (2010). A multi-level systems perspective for the science of team science. *Science Translational Medicine*, 2(49), 49cm24.
- Brin, S., Sergey, B., & Lawrence, P. (1998). The anatomy of a large-scale hypertextual Web search engine. *Computer Networks and ISDN Systems*, 30(1-7), 107-117.
- Burnard, P., & Philip, B. (1991). A method of analysing interview transcripts in qualitative research. *Nurse Education Today*, 11(6), 461-466.
- Cebrowski, A. K., & Garstka, J. J. (1998, January). [Network-centric warfare: Its origin and future](#). In *US Naval Institute Proceedings* (Vol. 124, No. 1, pp. 28-35).
- Chklovski, T., & Timothy, C. (2005). Collecting paraphrase corpora from volunteer contributors. In *Proceedings of the 3rd international conference on Knowledge capture - K-CAP '05*.
<http://doi.org/10.1145/1088622.1088644>
- Chklovski, T. (2003, October). Learner: a system for acquiring commonsense knowledge by analogy. In *Proceedings of the 2nd international conference on Knowledge capture* (pp. 4-12). ACM.
- Cloud Campus: The Software-Defined College. (n.d.). Retrieved May 18, 2016, from <https://www.meritalk.com/study/cloud-campus/>
- Cooper, S., Seth, C., Firas, K., Adrien, T., Janos, B., Jeehyung, L., ... Foldit, P. (2010). Predicting protein structures with a multiplayer online game. *Nature*, 466(7307), 756-760.
- Crawford, K., Hasan, H. M., Warne, L., & Linger, H. (2009). From traditional knowledge management in hierarchical organizations to a network centric paradigm for a changing world.

Criscuolo, P., Paola, C., Ammon, S., & Tony, S. (2007). Making knowledge visible: Using expert yellow pages to map capabilities in professional services firms. *Research Policy*, 36(10), 1603–1619.

Dijk, A., (2009), Improving expert finding and expert profiling.

Fruchterman, T. M. J., & Reingold, E. M. (1991). Graph drawing by force-directed placement. *Software: Practice & Experience*, 21(11), 1129–1164.

Garbash, D. A. (2015, August 17). Knownodes. Retrieved September 10, 2015, from https://www.youtube.com/watch?v=Lbt5vsHh_h0

Garcia-Perez, A and Ayres, R.(2010). Wikifailure: the Limitations of Technology for Knowledge Sharing” Electronic Journal of Knowledge Management Volume 8 Issue 1 (pp43 - 52), available online at www.ejkm.com

Hill, R. A., & Dunbar, R. I. M. (2003). Social network size in humans. *Human Nature* , 14(1), 53–72.

Hinds, P. J., & Pfeffer, J. (2003). Why organizations don’t “know what they know”: Cognitive and motivational factors affecting the transfer of expertise. *Sharing expertise: Beyond knowledge management*, 3-26.

Home | WhoKnows. (n.d.). Retrieved May 14, 2016, from <https://corp.whoknows.com/>

Hogg, M. A., & Abrams, D. (1990). Social motivation, self-esteem and social identity. *Social identity theory: Constructive and critical advances*, 28, 47.

King, W. R., Marks, P. V., & Scott, M. (2002). The most important issues in knowledge management. *Communications of the ACM*, 45(9), 93–97.

Krosnick, J. A., Lavrakas, P. J., & Nuri, K. (2000). Survey Research. In *Handbook of Research Methods in Social and Personality Psychology* (pp. 404–442).

Laloux, F. (2014). *Reinventing Organizations: A Guide to Creating Organizations Inspired by the Next Stage in Human Consciousness*.

Lenat, D. B., Guha, R. V., Karen, P., Dexter, P., & Mary, S. (1990). Cyc: toward programs with common sense. *Communications of the ACM*, 33(8), 30–49.

Quinn, A. J., & Bederson, B. B. (2011, May). Human computation: a survey and taxonomy of a growing field. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 1403-1412). ACM.

Scardamalia, M., & Bereiter, C. (2010). A Brief History of Knowledge Building. *Canadian Journal of Learning and Technology* 36(1). Published online at <http://www.cjlt.ca/index.php/cjlt/article/view/574>

Singh, P., Push, S., Thomas, L., Mueller, E. T., Grace, L., Travell, P., & Zhu, W. L. (2002). Open Mind Common Sense: Knowledge Acquisition from the General Public. In *Lecture Notes in Computer Science* (pp. 1223–1237).

Siorpaes, K., & Hepp, M. (2008). Games with a Purpose for the Semantic Web. *IEEE Intelligent Systems*, 23(3), 50–60.

Stipelman, B., Feng, A., Hall, K., Moser, R., Stokols, D., Nebeling, L., ... & Thomquist, M. (2010, April). The relationship between collaborative readiness and scientific productivity in the Transdisciplinary Research on Energetics and Cancer (TREC) Centers. In *Annals of Behavioral Medicine* (Vol. 39, pp. 143-143). 233 SPRING ST, NEW YORK, NY 10013 USA: SPRINGER.

Stokols, D., Hall, K. L., Taylor, B. K., & Moser, R. P. (2008). The science of team science: overview of the field and introduction to the supplement. *American Journal of Preventive Medicine*, 35(2 Suppl), S77–89.

Von Ahn, L., Kedia, M., & Blum, M. (2006, April). Verbosity: a game for collecting common-sense facts. In *Proceedings of the SIGCHI conference on Human Factors in computing systems* (pp. 75-78). ACM.

Wengraf, T. (2001). *Qualitative Research Interviewing: Biographic Narrative and Semi-Structured Methods*. SAGE.

Wilson, C. (2011). *Network Centric Operations: Background and Oversight Issues for Congress*.

Annex A Questions asked in expert identification quiz

Can you refer me to a specific colleague or colleagues who can help me with the following requests?

- 1) Doing a complicated graphic work on Photoshop.
- 2) Help with changing settings in the Jefyco admin.
- 3) Setting an API with o365.
- 4) Setting up EZpublish to work with a smartTV.
- 5) Help for Development using .NET language.
- 6) Setting up Shibboleth support for a new management software.
- 7) Completing a project using HTML5.

- 8) Making a short training video.
- 9) Consultation about distributed computation for a university project.
- 10) Consultation about using Moodle.

Appendix 4: [Product scenarios](#)

This excel spreadsheet cannot be adjusted for printing. Please go directly to the file as it appears online or contact the author for a copy.

Rhizi

Digital organisation

Client: Sorbonne Paris Cité IT personnel

Stakeholder: Eric ChereI

Start of development: Nov 1st

Start of Actual Usage: Dec 14th

OVERVIEW

Descartes University's IT department has 150 employees, but they are not fully aware of their colleagues abilities, experience and skill-sets. This causes unnecessary duplication of solutions and high-cost for finding the right expert to solve the many challenges employees face.

As a first step, Descartes would like to make it simple for each employee to query, input and explore a Rhizi containing each person's skills, projects, and teams he's involved with.

GOALS

Primary objective:

- Empower IT employees to identify experts within their group to solve issues and consult with.

Secondary objectives:

- Provide outsiders of IT department a means to visualize ongoing projects.
- Build a training plan across all SPC IT personnel
- Prepare SPC for external review of IDEX (end of 2016)

SPECIFICATIONS

Partial list of required features. TBD after the meeting Eric

- We're gonna have to force people to choose from a set list (that they can add to) for their work organization. HEGP is written differently 5 times...
- Same thing with skills ; it would be better to propose a list of existing ones that you can improve when you edit your cv.
- We probably have to use a hierarchy of organization, otherwise hundreds of people will be linked to the same node (like a hospital). I suggest teams, then lab / department, than hospital.
- Skills will probably need to be loosely linked together ; a lot of existing skills are related to psychiatry, to surgery or imagery. We could build a network or a hierarchy, both can work out.

MILESTONES

Start of development - 20.9.2015

After Signing of contract and requirements.

Launch - 15.10.2015

Launch of website to employees.

Index

Résumé (français) :

Title :

Abstract :

Mots clés (français) :

Keywords :