

Combinaison de critères par contraintes pour la Recherche d'Information Géographique

Damien Palacio

Laboratoire d'Informatique de l'Université de Pau et des Pays de l'Adour (LIUPPA)

26/11/2010



Plan

- 1 Introduction : La recherche d'information géographique dans des fonds documentaires textuels
 - Contexte : la Recherche d'Information Géographique (RIG)
 - Problématiques : Comment uniformiser, combiner, évaluer ?
 - Contributions : uniformisation générique, combinaison personnalisable et évaluations
- 2 Contribution : Vers la combinaison par contraintes de critères de recherche en RIG
- 3 Conclusion, discussions et perspectives



Plan

- 1 Introduction : La recherche d'information géographique dans des fonds documentaires textuels
- 2 Contribution : Vers la combinaison par contraintes de critères de recherche en RIG
 - Uniformisation générique
 - Recherche d'Information Géographique par combinaison de critères
 - Evaluation de systèmes de recherche d'information géographique
 - Expérimentations
- 3 Conclusion, discussions et perspectives



Plan

- 1 Introduction : La recherche d'information géographique dans des fonds documentaires textuels
- 2 Contribution : Vers la combinaison par contraintes de critères de recherche en RIG
- 3 Conclusion, discussions et perspectives
 - Bilan
 - Perspectives



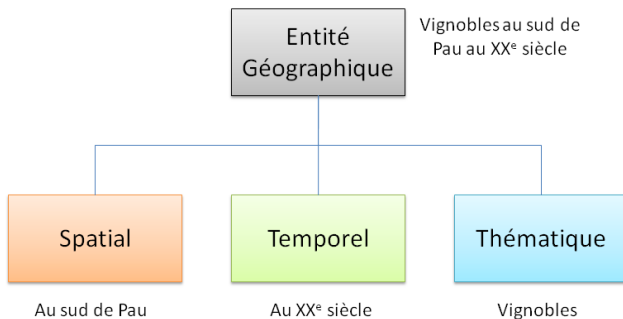
Plan

- 1 Introduction : La recherche d'information géographique dans des fonds documentaires textuels
 - Contexte : la Recherche d'Information Géographique (RIG)
 - Problématiques : Comment uniformiser, combiner, évaluer ?
 - Contributions : uniformisation générique, combinaison personnalisable et évaluations
- 2 Contribution : Vers la combinaison par contraintes de critères de recherche en RIG
- 3 Conclusion, discussions et perspectives



Contexte : la Recherche d'Information Géographique (RIG)

- L'information géographique peut être décomposée en trois facettes : le spatial, le temporel et le thématique [Use96, Gai01].



Contexte : la Recherche d'Information Géographique (RIG)

- Corpus : livres de type récits de voyages.
 - Livres numérisés et fournis par Médiathèque Intercommunale à Dimension Régionale (MIDR)
 - Plusieurs centaines de pages
 - Paragraphe : meilleur point d'entrée



Contexte : la Recherche d'Information Géographique (RIG)

- Corpus : livres de type récits de voyages.

Pendant que Russell courait le monde, une autre étoile de la pléiade, Charles Packe, apparaissait à Gavarnie, cette même année 1858. [...] A Gavarnie, il y a moins à découvrir qu'ailleurs et le Grand Cirque, d'après lui, avait depuis longtemps perdu le prestige et le charme de l'inconnu.

La tête des monts Pyrénéens s'élève, fière et peu connue, la chaîne de la Maladetta, dont le point culminant, le pic de Néthou, n'a été encore gravi par personne, et dont les approches sont défendues, presque de toutes parts, par des glaciers formidables. [...]

Contexte : la Recherche d'Information Géographique (RIG)

- Corpus : livres de type récits de voyages.

Pendant que Russell courait le monde, une autre étoile de la pléiade, Charles Packe, apparaissait à Gavarnie, cette même année 1858. [...] A Gavarnie, il y a moins à découvrir qu'ailleurs et le Grand Cirque, d'après lui, avait depuis longtemps perdu le prestige et le charme de l'inconnu.

La tête des monts Pyrénéens s'élève, fière et peu connue, la chaîne de la Maladetta, dont le point culminant, le pic de Néthou, n'a été encore gravi par personne, et dont les approches sont défendues, presque de toutes parts, par des glaciers formidables. [...]

Plan

- 1 Introduction : La recherche d'information géographique dans des fonds documentaires textuels
 - Contexte : la Recherche d'Information Géographique (RIG)
 - **Problématiques : Comment uniformiser, combiner, évaluer ?**
 - Contributions : uniformisation générique, combinaison personnalisable et évaluations
- 2 Contribution : Vers la combinaison par contraintes de critères de recherche en RIG
- 3 Conclusion, discussions et perspectives



Verrous identifiés

Problèmes

- Identifier l'information géographique :
 - Différentes facettes rarement présentes côte à côte
 - Information temporelle non présente systématiquement
 - Information thématique diffuse et ambiguë

Choix

- Traiter chaque facette indépendamment avec des traitements dédiés [CJP06, MSA05]
- Traiter le thématique avec approche par termes



Verrous identifiés

Problèmes

- Identifier l'information géographique :
 - Différentes facettes rarement présentes côte à côte
 - Information temporelle non présente systématiquement
 - Information thématique diffuse et ambiguë

Choix

- Traiter chaque facette indépendamment avec des traitements dédiés [CJP06, MSA05]
- Traiter la thématique avec approche par termes



Verrous identifiés

Problèmes

- Identifier l'information géographique :
 - Différentes facettes rarement présentes côte à côte
 - Information temporelle non présente systématiquement
 - Information thématique diffuse et ambiguë

Choix

- Traiter chaque facette indépendamment avec des traitements dédiés [CJP06, MSA05]
- Traiter la thématique avec approche par termes



Verrous identifiés

Problèmes

- Identifier l'information géographique :
 - Différentes facettes rarement présentes côte à côte
 - Information temporelle non présente systématiquement
 - Information thématique diffuse et ambiguë

Choix

- Traiter chaque facette indépendamment avec des traitements dédiés [CJP06, MSA05]
- Traiter la thématique avec approche par termes



Verrous identifiés

Problèmes

- Identifier l'information géographique :
 - Différentes facettes rarement présentes côte à côte
 - Information temporelle non présente systématiquement
 - Information thématique diffuse et ambiguë

Choix

- Traiter chaque facette indépendamment avec des traitements dédiés [CJP06, MSA05]
- Traiter la thématique avec approche par termes



Verrous identifiés

Problèmes

- Identifier l'information géographique :
 - Différentes facettes rarement présentes côte à côte
 - Information temporelle non présente systématiquement
 - Information thématique diffuse et ambiguë

Choix

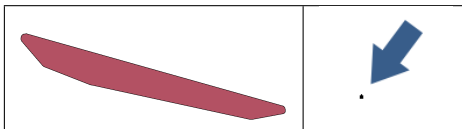
- Traiter chaque facette indépendamment avec des traitements dédiés [CJP06, MSA05]
- Traiter la thématique avec approche par termes



Index

ET_Id	Doc_Id	Texte	Interval de temps	T_Id	Terme	Fréquence
#1	D1	'1858'	01/01/1858- 31/12/1858	...		
...				#25	'inconnu'	D1,1;
				...		
				#87	'tête'	D2,1; D4,2

ES_Id	Doc_Id	Texte	Géométrie
#1	D1	'à Gavarnie'	(42.7683445932833 -0.150941405986334, ...
...			
#3	D2	'monts Pyrénéens'	(42.220343224213 2.59101516708941, ...
...			



Requêtes

- « les risques accidentels en montagne si possible liés à des balades ou randonnées »
 - 3 critères thématiques (accidents, montagne et balades).
- « documents sur les montagnes des Pyrénées entre 1800 et 1900 mais pas sur Gavarnie et si possible sans rapport avec les ascensions »
 - 1 critère temporel (entre 1800 et 1900)
 - 2 thématiques (montagnes, ascensions)
 - 2 spatiaux (Pyrénées, Gavarnie)



Requêtes

- « les risques accidentels en montagne si possible liés à des balades ou randonnées »
 - 3 critères thématiques (accidents, montagne et balades).
- « documents sur les montagnes des Pyrénées entre 1800 et 1900 mais pas sur Gavarnie et si possible sans rapport avec les ascensions »
 - 1 critère temporel (entre 1800 et 1900)
 - 2 thématiques (montagnes, ascensions)
 - 2 spatiaux (Pyrénées, Gavarnie)



Requêtes

- « les risques accidentels en montagne si possible liés à des balades ou randonnées »
 - 3 critères thématiques (accidents, montagne et balades).
- « documents sur les montagnes des Pyrénées entre 1800 et 1900 mais pas sur Gavarnie et si possible sans rapport avec les ascensions »
 - 1 critère temporel (entre 1800 et 1900)
 - 2 thématiques (montagnes, ascensions)
 - 2 spatiaux (Pyrénées, Gavarnie)



Requêtes

- « les risques accidentels en montagne si possible liés à des balades ou randonnées »
 - 3 critères thématiques (accidents, montagne et balades).
- « documents sur les montagnes des Pyrénées entre 1800 et 1900 mais pas sur Gavarnie et si possible sans rapport avec les ascensions »
 - 1 critère temporel (entre 1800 et 1900)
 - 2 thématiques (montagnes, ascensions)
 - 2 spatiaux (Pyrénées, Gavarnie)



Requêtes

- « les risques accidentels en montagne si possible liés à des balades ou randonnées »
 - 3 critères thématiques (accidents, montagne et balades).
- « documents sur les montagnes des Pyrénées entre 1800 et 1900 mais pas sur Gavarnie et si possible sans rapport avec les ascensions »
 - 1 critère temporel (entre 1800 et 1900)
 - 2 thématiques (montagnes, ascensions)
 - 2 spatiaux (Pyrénées, Gavarnie)



Requêtes

- « les risques accidentels en montagne si possible liés à des balades ou randonnées »
 - 3 critères thématiques (accidents, montagne et balades).
- « documents sur les montagnes des Pyrénées entre 1800 et 1900 mais pas sur Gavarnie et si possible sans rapport avec les ascensions »
 - 1 critère temporel (entre 1800 et 1900)
 - 2 thématiques (montagnes, ascensions)
 - 2 spatiaux (Pyrénées, Gavarnie)



Autres Verrous identifiés

Problèmes

- Comment homogénéiser les différentes facettes :
 - Représentations temporelles et spatiales de différentes échelles (exemple : communes et régions) vs thématique seulement termes
 - Calcul de similarité dédiée au temporel, au spatial (exemple : basé sur le recouvrement) vs approches statistiques pour les termes
- Comment combiner les résultats issus de SRI dédiés à chaque facette et personnaliser cette combinaison ?
- Comment évaluer un système de recherche d'information portant sur les 3 facettes ?

Objectif

Améliorer la RIG en combinant des systèmes de RI existants



Autres Verrous identifiés

Problèmes

- Comment homogénéiser les différentes facettes :
 - Représentations temporelles et spatiales de différentes échelles (exemple : communes et régions) vs thématique seulement termes
 - Calcul de similarité dédiée au temporel, au spatial (exemple : basé sur le recouvrement) vs approches statistiques pour les termes
- Comment combiner les résultats issus de SRI dédiés à chaque facette et personnaliser cette combinaison ?
- Comment évaluer un système de recherche d'information portant sur les 3 facettes ?

Objectif

Améliorer la RIG en combinant des systèmes de RI existants



Autres Verrous identifiés

Problèmes

- Comment homogénéiser les différentes facettes :
 - Représentations temporelles et spatiales de différentes échelles (exemple : communes et régions) vs thématique seulement termes
 - Calcul de similarité dédiée au temporel, au spatial (exemple : basé sur le recouvrement) vs approches statistiques pour les termes
- Comment combiner les résultats issus de SRI dédiés à chaque facette et personnaliser cette combinaison ?
- Comment évaluer un système de recherche d'information portant sur les 3 facettes ?

Objectif

Améliorer la RIG en combinant des systèmes de RI existants



Autres Verrous identifiés

Problèmes

- Comment homogénéiser les différentes facettes :
 - Représentations temporelles et spatiales de différentes échelles (exemple : communes et régions) vs thématique seulement termes
 - Calcul de similarité dédiée au temporel, au spatial (exemple : basé sur le recouvrement) vs approches statistiques pour les termes
- Comment combiner les résultats issus de SRI dédiés à chaque facette et personnaliser cette combinaison ?
- Comment évaluer un système de recherche d'information portant sur les 3 facettes ?

Objectif

Améliorer la RIG en combinant des systèmes de RI existants



Autres Verrous identifiés

Problèmes

- Comment homogénéiser les différentes facettes :
 - Représentations temporelles et spatiales de différentes échelles (exemple : communes et régions) vs thématique seulement termes
 - Calcul de similarité dédiée au temporel, au spatial (exemple : basé sur le recouvrement) vs approches statistiques pour les termes
- Comment combiner les résultats issus de SRI dédiés à chaque facette et personnaliser cette combinaison ?
- Comment évaluer un système de recherche d'information portant sur les 3 facettes ?

Objectif

Améliorer la RIG en combinant des systèmes de RI existants



Autres Verrous identifiés

Problèmes

- Comment homogénéiser les différentes facettes :
 - Représentations temporelles et spatiales de différentes échelles (exemple : communes et régions) vs thématique seulement termes
 - Calcul de similarité dédiée au temporel, au spatial (exemple : basé sur le recouvrement) vs approches statistiques pour les termes
- Comment combiner les résultats issus de SRI dédiés à chaque facette et personnaliser cette combinaison ?
- Comment évaluer un système de recherche d'information portant sur les 3 facettes ?

Objectif

Améliorer la RIG en combinant des systèmes de RI existants

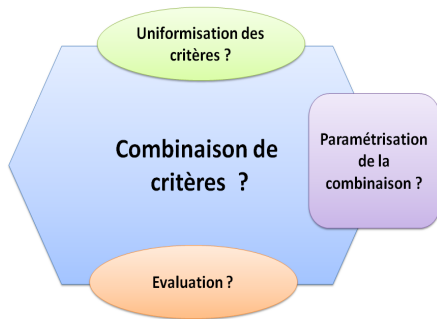


Plan

- 1 Introduction : La recherche d'information géographique dans des fonds documentaires textuels
 - Contexte : la Recherche d'Information Géographique (RIG)
 - Problématiques : Comment uniformiser, combiner, évaluer ?
 - Contributions : uniformisation générique, combinaison personnalisable et évaluations
- 2 Contribution : Vers la combinaison par contraintes de critères de recherche en RIG
- 3 Conclusion, discussions et perspectives



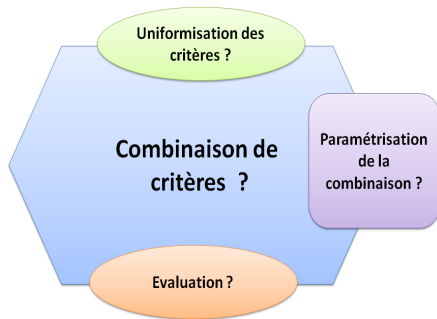
Contributions



- Approche générique d'uniformisation de données
- Approche de combinaison de critères paramétrable
- Un cadre expérimental pour évaluer un SRI géographique
- Evaluation et quantification combinaison facettes géographiques



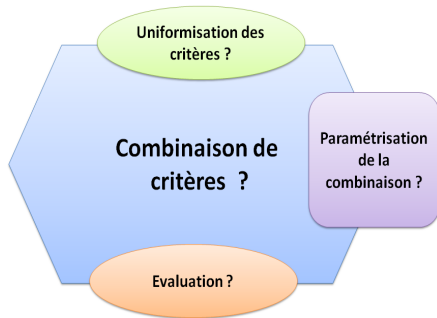
Contributions



- Approche générique d'uniformisation de données
- Approche de combinaison de critères paramétrable
- Un cadre expérimental pour évaluer un SRI géographique
- Evaluation et quantification combinaison facettes géographiques



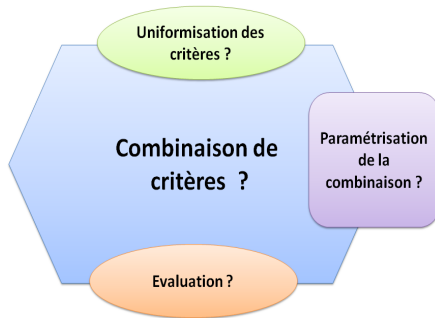
Contributions



- Approche générique d'uniformisation de données
- Approche de combinaison de critères paramétrable
- Un cadre expérimental pour évaluer un SRI géographique
- Evaluation et quantification combinaison facettes géographiques



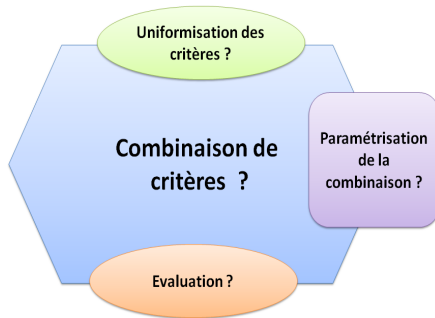
Contributions



- Approche générique d'uniformisation de données
- Approche de combinaison de critères paramétrable
- Un cadre expérimental pour évaluer un SRI géographique
- Evaluation et quantification combinaison facettes géographiques



Contributions



- Approche générique d'uniformisation de données
- Approche de combinaison de critères paramétrable
- Un cadre expérimental pour évaluer un SRI géographique
- Evaluation et quantification combinaison facettes géographiques



Plan

- 1 Introduction : La recherche d'information géographique dans des fonds documentaires textuels
- 2 Contribution : Vers la combinaison par contraintes de critères de recherche en RIG
 - Uniformisation générique
 - Recherche d'Information Géographique par combinaison de critères
 - Evaluation de systèmes de recherche d'information géographique
 - Expérimentations
- 3 Conclusion, discussions et perspectives



Approche classique en RI

- Représentations textuelles uniformisées [BYRN99, EGDC00] :
lemmatisation/troncature
- Utilisation de fréquences pour le calcul des scores (TF, IDF) [SYY75]



Approche classique en RI

- Représentations textuelles uniformisées [BYRN99, EGDC00] :
lemmatisation/troncature
- Utilisation de fréquences pour le calcul des scores (TF, IDF) [SYY75]

Termes	Lemmes
ai avoir avons	avoir



Approche classique en RI

- Représentations textuelles uniformisées [BYRN99, EGDC00] : lemmatisation/troncature
- Utilisation de fréquences pour le calcul des scores (TF, IDF) [SYY75]

Termes	Lemmes
ai avoir avons	avoir



Approche classique en RI

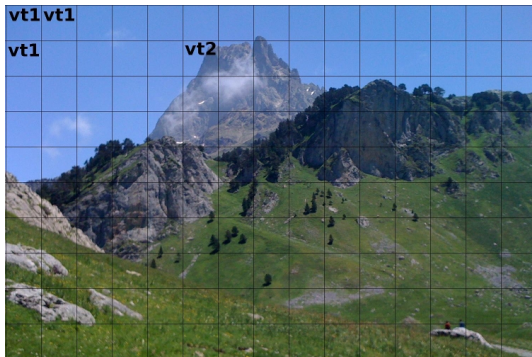
- Représentations textuelles uniformisées [BYRN99, EGDC00] : lemmatisation/troncature
- Utilisation de fréquences pour le calcul des scores (TF, IDF) [SYY75]

Termes	Lemmes (Fréquences)
ai avoir avons	avoir (3)



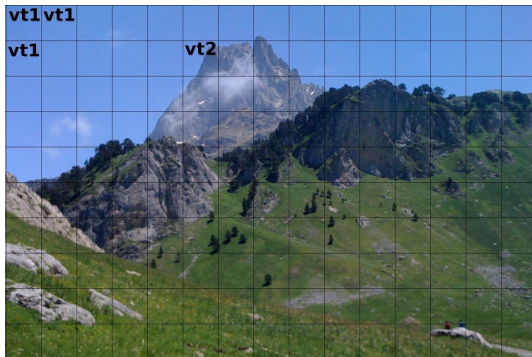
Application à la RI Multimedia

- Image représentée comme un sac de termes visuels (visterms) [PMLC07]
- Utilisation des fréquences pour calculer le poids de chaque visterms



Application à la RI Multimedia

- Image représentée comme un sac de termes visuels (visterms) [PMLC07]
- Utilisation des fréquences pour calculer le poids de chaque visterms



Indexation multidimensionnelle basée sur le « tuilage »

- **Indexation préalable existante**
- Objectif : uniformiser ces données
- « Tuile » : modèle de représentation uniforme
- « Tuilage » : segmentation en tuiles
- Chaque représentation de l'index initial rattachée à une ou plusieurs tuiles



Indexation multidimensionnelle basée sur le « tuilage »

- Indexation préalable existante
- Objectif : uniformiser ces données
- « Tuile » : modèle de représentation uniforme
- « Tuilage » : segmentation en tuiles
- Chaque représentation de l'index initial rattachée à une ou plusieurs tuiles

**Index
Existant
(données)**



Uniformisation

**Index
Uniformisé
(tuiles)**



Indexation multidimensionnelle basée sur le « tuilage »

- Indexation préalable existante
- Objectif : uniformiser ces données
- « Tuile » : modèle de représentation uniforme
- « Tuilage » : segmentation en tuiles
- Chaque représentation de l'index initial rattachée à une ou plusieurs tuiles

**Index
Existant
(données)**



Uniformisation

**Index
Uniformisé
(tuiles)**



Indexation multidimensionnelle basée sur le « tuilage »

- Indexation préalable existante
- Objectif : uniformiser ces données
- « Tuile » : modèle de représentation uniforme
- « Tuilage » : segmentation en tuiles
- Chaque représentation de l'index initial rattachée à une ou plusieurs tuiles

**Index
Existant
(données)**



Uniformisation

**Index
Uniformisé
(tuiles)**



Indexation multidimensionnelle basée sur le « tuilage »

- Indexation préalable existante
- Objectif : uniformiser ces données
- « Tuile » : modèle de représentation uniforme
- « Tuilage » : segmentation en tuiles
- Chaque représentation de l'index initial rattachée à une ou plusieurs tuiles

Index
Existant
(données)



Uniformisation

Index
Uniformisé
(tuiles)



Indexation multidimensionnelle basée sur le « tuilage »

- Au domaine O inclus dans l'espace R^n correspond un domaine T inclus dans l'espace R^n
- O est constitué d'un ensemble d'objets O_1, \dots, O_p
- T est constitué de l'union de m sous-espaces (les tuiles)
- Pour chaque sous-espace de T en intersection avec 1 ou plusieurs objets de O on retient le nombre d'intersections (N_{T_i})

$$O \subseteq \mathbb{R}^n \longrightarrow T \subseteq \mathbb{R}^n$$

$$O = \{O_1, O_2, O_3, \dots, O_p\}$$

$$T = \bigcup_{i=1}^m T_i \quad (1)$$

$$N_{T_i} = |T_i| \mid T_i \cap O_j \neq \emptyset \quad \forall j = 1, \dots, p \mid$$

avec $|x|$ la cardinalité de x



Indexation multidimensionnelle basée sur le « tuilage »

- Au domaine O inclus dans l'espace R^n correspond un domaine T inclus dans l'espace R^n
- O est constitué d'un ensemble d'objets O_1, \dots, O_p
- T est constitué de l'union de m sous-espaces (les tuiles)
- Pour chaque sous-espace de T en intersection avec 1 ou plusieurs objets de O on retient le nombre d'intersections (N_{T_i})

$$O \subseteq \mathbb{R}^n \longrightarrow T \subseteq \mathbb{R}^n$$

$$O = \{O_1, O_2, O_3, \dots, O_p\}$$

$$T = \bigcup_{i=1}^m T_i \quad (1)$$

$$N_{T_i} = |T_i| \mid T_i \cap O_j \neq \emptyset \quad \forall j = 1, \dots, p \mid$$

avec $|x|$ la cardinalité de x



Indexation multidimensionnelle basée sur le « tuilage »

- Au domaine O inclus dans l'espace R^n correspond un domaine T inclus dans l'espace R^n
- O est constitué d'un ensemble d'objets O_1, \dots, O_p
- T est constitué de l'union de m sous-espaces (les tuiles)
- Pour chaque sous-espace de T en intersection avec 1 ou plusieurs objets de O on retient le nombre d'intersections (N_{T_i})

$$O \subseteq \mathbb{R}^n \longrightarrow T \subseteq \mathbb{R}^n$$

$$O = \{O_1, O_2, O_3, \dots, O_p\}$$

$$T = \bigcup_{i=1}^m T_i \quad (1)$$

$$N_{T_i} = |T_i| \mid T_i \cap O_j \neq \emptyset \quad \forall j = 1, \dots, p \mid$$

avec $|x|$ la cardinalité de x



Indexation multidimensionnelle basée sur le « tuilage »

- Au domaine O inclus dans l'espace R^n correspond un domaine T inclus dans l'espace R^n
- O est constitué d'un ensemble d'objets O_1, \dots, O_p
- T est constitué de l'union de m sous-espaces (les tuiles)
- Pour chaque sous-espace de T en intersection avec 1 ou plusieurs objets de O on retient le nombre d'intersections (N_{T_i})

$$O \subseteq \mathbb{R}^n \longrightarrow T \subseteq \mathbb{R}^n$$

$$O = \{O_1, O_2, O_3, \dots, O_p\}$$

$$T = \bigcup_{i=1}^m T_i \quad (1)$$

$$N_{T_i} = |T_i| \mid T_i \cap O_j \neq \emptyset \quad \forall j = 1, \dots, p \mid$$

avec $|x|$ la cardinalité de x



Indexation multidimensionnelle basée sur le « tuilage »

- Au domaine O inclus dans l'espace R^n correspond un domaine T inclus dans l'espace R^n
- O est constitué d'un ensemble d'objets O_1, \dots, O_p
- T est constitué de l'union de m sous-espaces (les tuiles)
- Pour chaque sous-espace de T en intersection avec 1 ou plusieurs objets de O on retient le nombre d'intersections (N_{T_i})

$$O \subseteq \mathbb{R}^n \longrightarrow T \subseteq \mathbb{R}^n$$

$$O = \{O_1, O_2, O_3, \dots, O_p\}$$

$$T = \bigcup_{i=1}^m T_i \quad (1)$$

$$N_{T_i} = |T_i| \quad |T_i \cap O_j| \neq \emptyset \quad \forall j = 1, \dots, p$$

avec $|x|$ la cardinalité de x



Indexation multidimensionnelle basée sur le « tuilage »

Représentations



à 1
dimension

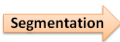
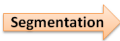
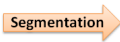


à 2
dimensions

...



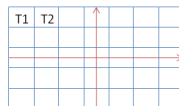
à n
dimensions



Tuilages



à 1 dimension



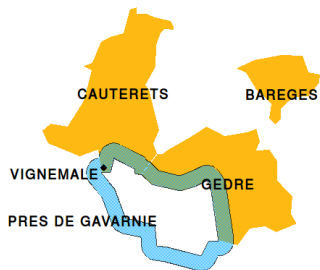
à 2 dimensions



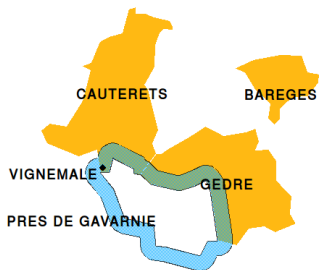
à n dimensions



Indexation multidimensionnelle basée sur le « tuilage » : exemple



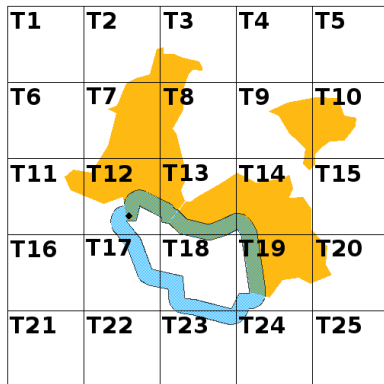
Indexation multidimensionnelle basée sur le « tuilage » : exemple



T1	T2	T3	T4	T5
T6	T7	T8	T9	T10
T11	T12	T13	T14	T15
T16	T17	T18	T19	T20
T21	T22	T23	T24	T25



Indexation multidimensionnelle basée sur le « tuilage » : exemple



Indexation multidimensionnelle basée sur le « tuilage » : exemple

T1	T2	T3	T4	T5
T6	T7	T8	T9	T10
T11	T12	T13	T14	T15
T16	T17	T18	T19	T20
T21	T22	T23	T24	T25

T1	T2	T3	T4	T5
T6	T7	T8	T9	T10
T11	T12	T13	T14	T15
T16	T17	T18	T19	T20
T21	T22	T23	T24	T25



Approche multi-échelles

- Informations contenues dans corpus de différentes échelles
 - Choix de l'index adapté à la granularité de la requête
- « documents sur Pau »
 - Pau : commune
 - Tuilage communal
- « documents sur l'Aquitaine »
 - Aquitaine : région
 - Tuilage régional



Approche multi-échelles

- Informations contenues dans corpus de différentes échelles
 - Choix de l'index adapté à la granularité de la requête

- « documents sur Pau »
 - Pau : commune
 - Tuilage communal

- « documents sur l'Aquitaine »
 - Aquitaine : région
 - Tuilage régional



Approche multi-échelles

- Informations contenues dans corpus de différentes échelles
 - Choix de l'index adapté à la granularité de la requête

- « documents sur Pau »

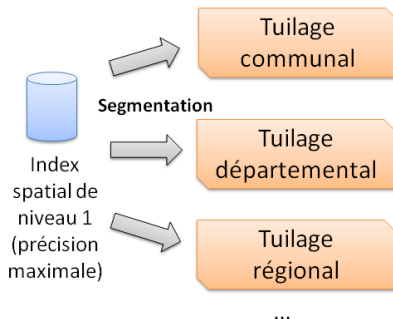
→ Pau : commune

→ Tuilage communal

- « documents sur l'Aquitaine »

→ Aquitaine : région

→ Tuilage régional



Approche multi-échelles

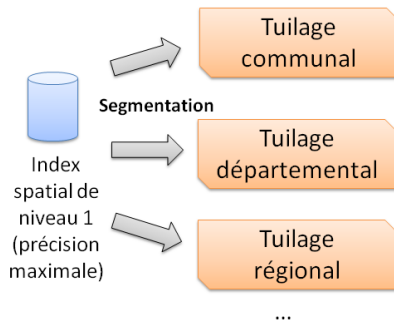
- Informations contenues dans corpus de différentes échelles
 - Choix de l'index adapté à la granularité de la requête

- « documents sur **Pau** »

- Pau : commune
- Tuilage communal

- « documents sur **l'Aquitaine** »

- Aquitaine : région
- Tuilage régional



Approche multi-échelles

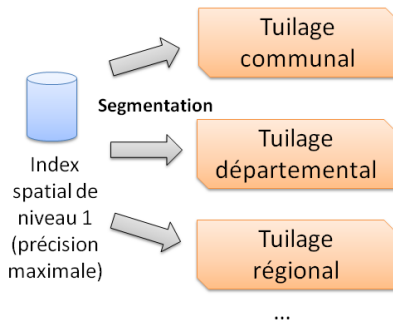
- Informations contenues dans corpus de différentes échelles
 - Choix de l'index adapté à la granularité de la requête

- « documents sur **Pau** »

- Pau : commune
- Tuilage communal

- « documents sur **l'Aquitaine** »

- Aquitaine : région
- Tuilage régional



Approche multi-échelles

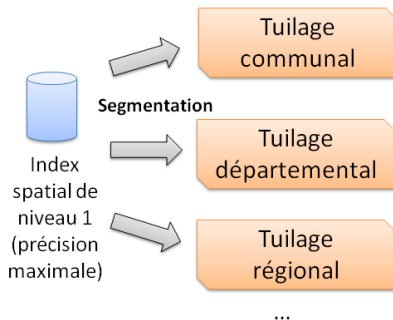
- Informations contenues dans corpus de différentes échelles
 - Choix de l'index adapté à la granularité de la requête

- « documents sur **Pau** »

- Pau : commune
- Tuilage communal

- « documents sur **l'Aquitaine** »

- Aquitaine : région
- Tuilage régional



Approche multi-échelles

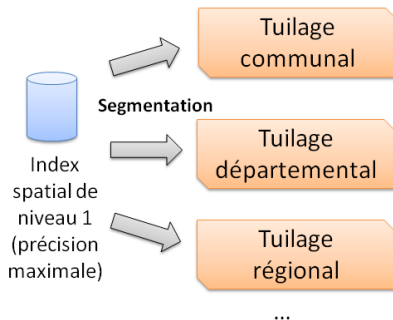
- Informations contenues dans corpus de différentes échelles
 - Choix de l'index adapté à la granularité de la requête

- « documents sur **Pau** »

- Pau : commune
- Tuilage communal

- « documents sur **l'Aquitaine** »

- Aquitaine : région
- Tuilage régional



Approche multi-échelles

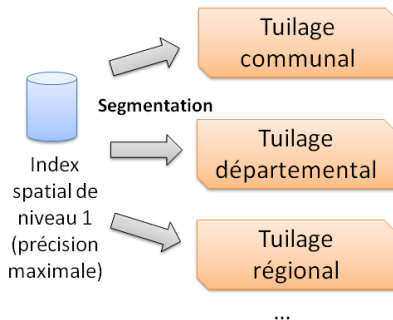
- Informations contenues dans corpus de différentes échelles
 - Choix de l'index adapté à la granularité de la requête

- « documents sur **Pau** »

- Pau : commune
- Tuilage communal

- « documents sur **l'Aquitaine** »

- Aquitaine : région
- Tuilage régional



Approche multi-échelles

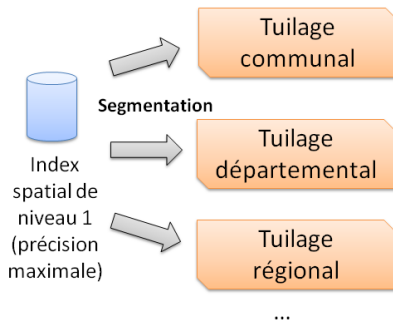
- Informations contenues dans corpus de différentes échelles
 - Choix de l'index adapté à la granularité de la requête

- « documents sur **Pau** »

- Pau : commune
- Tuilage communal

- « documents sur **l'Aquitaine** »

- Aquitaine : région
- Tuilage régional



Types de tuilages

- Tuilage « régulier »

- découpage géométrique basé sur une unité de référence (ex : distance pour spatial)
- tuiles de taille similaire
- utilisé par Pham et al [PMLC07] pour les *visterms*

- Avantages : taille et frontières des tuiles réglables
- Inconvénients : modification du corpus peut impliquer régénérer tuilage



Types de tuilages

- Tuilage « régulier »
 - découpage géométrique basé sur une unité de référence (ex : distance pour spatial)
 - tuiles de taille similaire
 - utilisé par Pham et al [PMLC07] pour les *visterms*
 - Avantages : taille et frontières des tuiles réglables
 - Inconvénients : modification du corpus peut impliquer régénérer tuilage



Types de tuilages

- Tuilage « régulier »
 - découpage géométrique basé sur une unité de référence (ex : distance pour spatial)
 - tuiles de taille similaire
 - utilisé par Pham et al [PMLC07] pour les *visterms*
 - Avantages : taille et frontières des tuiles réglables
 - Inconvénients : modification du corpus peut impliquer régénérer tuilage



Types de tuilages

- Tuilage « régulier »
 - découpage géométrique basé sur une unité de référence (ex : distance pour spatial)
 - tuiles de taille similaire
 - utilisé par Pham et al [PMLC07] pour les *visterms*
 - Avantages : taille et frontières des tuiles réglables
 - Inconvénients : modification du corpus peut impliquer régénérer tuilage



Types de tuilages

- Tuilage « régulier »
 - découpage géométrique basé sur une unité de référence (ex : distance pour spatial)
 - tuiles de taille similaire
 - utilisé par Pham et al [PMLC07] pour les *visterms*
 - Avantages : taille et frontières des tuiles réglables
 - Inconvénients : modification du corpus peut impliquer régénérer tuilage



Types de tuilages

- Tuilage « régulier »
 - découpage géométrique basé sur une unité de référence (ex : distance pour spatial)
 - tuiles de taille similaire
 - utilisé par Pham et al [PMLC07] pour les *visterms*
 - Avantages : taille et frontières des tuiles réglables
 - Inconvénients : modification du corpus peut impliquer régénérer tuilage



Types de tuilages

- Tuilage « explicite »
 - découpage « signifiant » bati sur des critères humains (sens commun)
 - tuiles de taille différentes
 - Avantages : tuiles figées et identifiées, peut être plus précis pour données à plus de 2 dimensions
 - Inconvénient : découpage ne peut être modifié



Types de tuilages

- Tuilage « explicite »
 - découpage « signifiant » bati sur des critères humains (sens commun)
 - tuiles de taille différentes
 - Avantages : tuiles figées et identifiées, peut être plus précis pour données à plus de 2 dimensions
 - Inconvénient : découpage ne peut être modifié



Types de tuilages

- Tuilage « explicite »
 - découpage « signifiant » bati sur des critères humains (sens commun)
 - tuiles de taille différentes
 - Avantages : tuiles figées et identifiées, peut être plus précis pour données à plus de 2 dimensions
 - Inconvénient : découpage ne peut être modifié



Types de tuilages

- Tuilage « explicite »
 - découpage « signifiant » bati sur des critères humains (sens commun)
 - tuiles de taille différentes
 - Avantages : tuiles figées et identifiées, peut être plus précis pour données à plus de 2 dimensions
 - Inconvénient : découpage ne peut être modifié



Types de tuilages

- Tuilage « explicite »
 - découpage « signifiant » bati sur des critères humains (sens commun)
 - tuiles de taille différentes
 - Avantages : tuiles figées et identifiées, peut être plus précis pour données à plus de 2 dimensions
 - Inconvénient : découpage ne peut être modifié



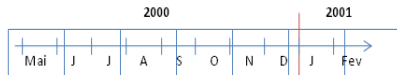
Application à la RIG : temporel

- Tuilage régulier et tuilage calendaire (découpage explicite : jours, mois, années, ...)



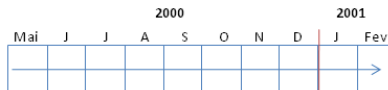
Application à la RIG : temporel

- Tuilage régulier et tuilage calendaire (découpage explicite : jours, mois, années, ...)



Application à la RIG : temporel

- Tuilage régulier et tuilage calendaire (découpage explicite : jours, mois, années, ...)



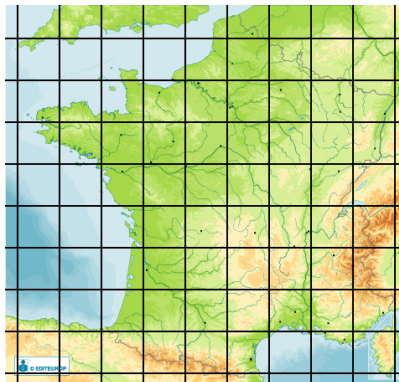
Application à la RIG : spatial

- Tuilage régulier et tuilage administratif (découpage explicite : communes, départements, régions, ...)



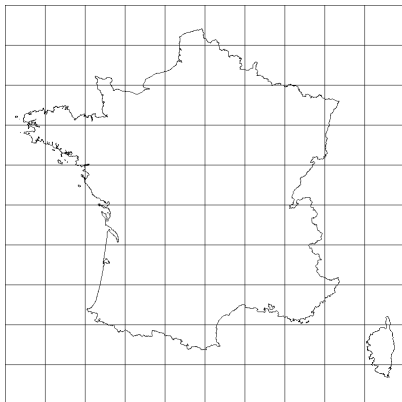
Application à la RIG : spatial

- Tuilage régulier et tuilage administratif (découpage explicite : communes, départements, régions, ...)



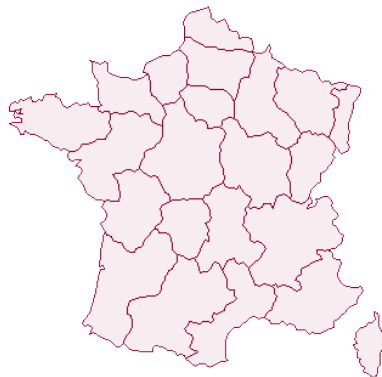
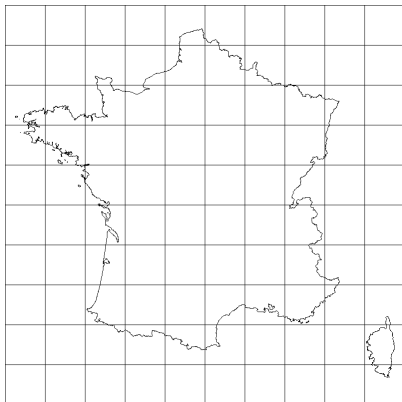
Application à la RIG : spatial

- Tuilage régulier et tuilage administratif (découpage explicite : communes, départements, régions, ...)



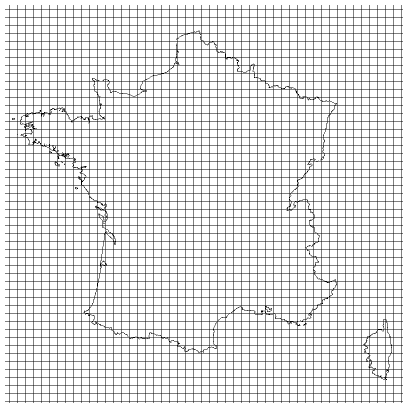
Application à la RIG : spatial

- Tuilage régulier et tuilage administratif (découpage explicite : communes, départements, régions, ...)



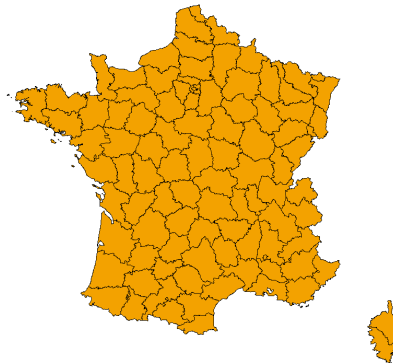
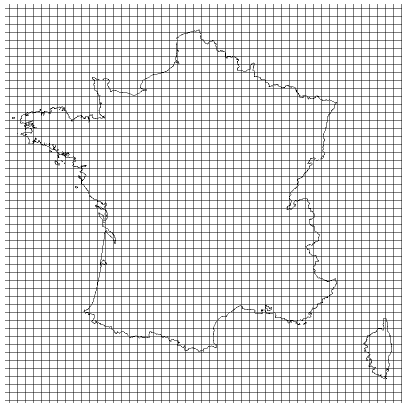
Application à la RIG : spatial

- Tuilage régulier et tuilage administratif (découpage explicite : communes, départements, régions, ...)



Application à la RIG : spatial

- Tuilage régulier et tuilage administratif (découpage explicite : communes, départements, régions, ...)



Pondération des tuiles

- Fréquence binaire (1 si intersection, 0 sinon)
 - Formules classiques de RI appliquées aux tuiles : TF, TF-IDF, Okapi (fréquence binaire normalisée)
- Fréquence proportionnelle : ratio recouvrement tuile/objet (valeur entre 0 et 1)
 - Application du TF sur fréquence proportionnelle normalisée (TFp)



Pondération des tuiles

- Fréquence binaire (1 si intersection, 0 sinon)
 - Formules classiques de RI appliquées aux tuiles : TF, TF·IDF, Okapi (fréquence binaire normalisée)
- Fréquence proportionnelle : ratio recouvrement tuile/objet (valeur entre 0 et 1)
 - Application du TF sur fréquence proportionnelle normalisée (TFp)



Pondération des tuiles

- Fréquence binaire (1 si intersection, 0 sinon)
 - Formules classiques de RI appliquées aux tuiles : TF, TF-IDF, Okapi (fréquence binaire normalisée)
- Fréquence proportionnelle : ratio recouvrement tuile/objet (valeur entre 0 et 1)
 - Application du TF sur fréquence proportionnelle normalisée (TFp)



Pondération des tuiles

- Fréquence binaire (1 si intersection, 0 sinon)
 - Formules classiques de RI appliquées aux tuiles : TF, TF-IDF, Okapi (fréquence binaire normalisée)
- Fréquence proportionnelle : ratio recouvrement tuile/objet (valeur entre 0 et 1)
 - Application du TF sur fréquence proportionnelle normalisée (TFp)

Pondération des tuiles : Exemple



id_t	liste id_{O_i}	fréquence binaire	fréquence proportionnelle
T1	[]	0	0
T2	[O ₁]	1	0,10
...			
T12	[O ₁ , O ₄ , O ₅]	3	0,75
...			

Pondération des tuiles : Exemple



id_t	liste id_{O_i}	fréquence binaire	fréquence proportionnelle
T1	[]	0	0
T2	[O_1]	1	0,10
...			
T12	[O_1, O_4, O_5]	3	0,75
...			

Après l'uniformisation ?

Bilan

- Approche d'uniformisation appliquée aux informations spatiales et temporelles
- Thématique : utilisation des approches classiques pour les termes

Ensuite

- Comment combiner différents critères et comment paramétrer cette combinaison ?



Après l'uniformisation ?

Bilan

- Approche d'uniformisation appliquée aux informations spatiales et temporelles
- Thématique : utilisation des approches classiques pour les termes

Ensuite

- Comment combiner différents critères et comment paramétrer cette combinaison ?



Après l'uniformisation ?

Bilan

- Approche d'uniformisation appliquée aux informations spatiales et temporelles
- Thématique : utilisation des approches classiques pour les termes

Ensuite

- Comment combiner différents critères et comment paramétrer cette combinaison ?

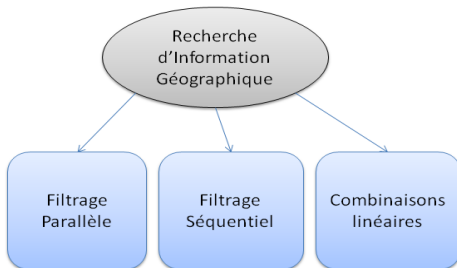


Plan

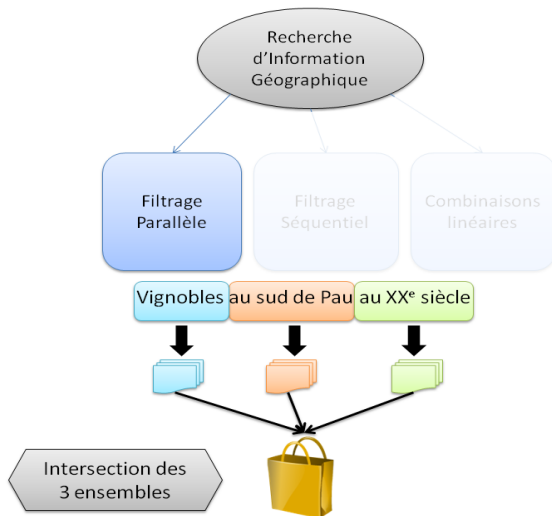
- 1 Introduction : La recherche d'information géographique dans des fonds documentaires textuels
- 2 Contribution : Vers la combinaison par contraintes de critères de recherche en RIG
 - Uniformisation générique
 - Recherche d'Information Géographique par combinaison de critères
 - Evaluation de systèmes de recherche d'information géographique
 - Expérimentations
- 3 Conclusion, discussions et perspectives



Approches de Combinaison en RIG

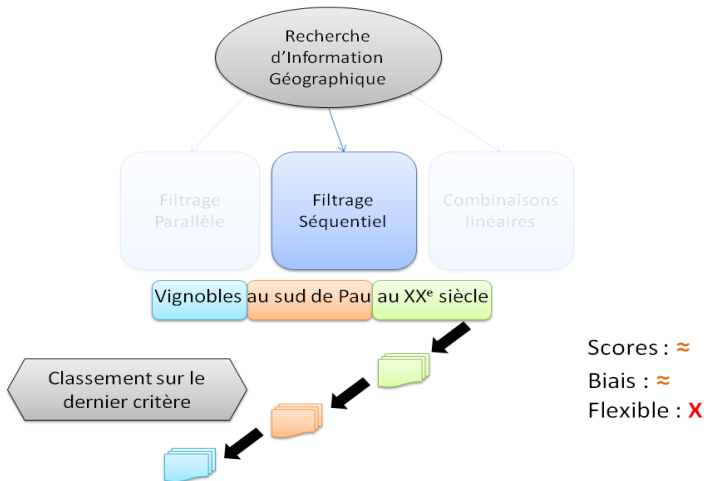


Approches de Combinaison en RIG

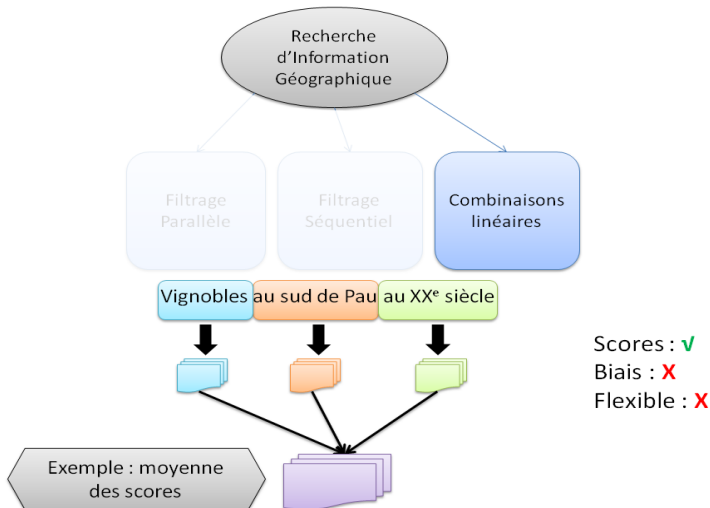


Scores : **X**
Biais : **✓**
Flexible : **X**

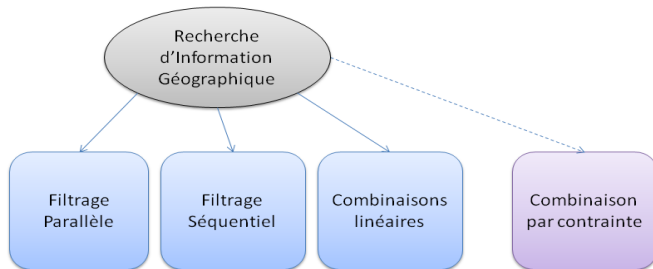
Approches de Combinaison en RIG



Approches de Combinaison en RIG



Approches de Combinaison en RIG



Scores :	X	≈	✓	!
Biais :	✓	≈	X	!
Flexible :	X	X	X	!



Approche par contraintes - Descriptif général

Le système de combinaison par contraintes modélisé comme un quadruplet (C, O, P, E) :

- les choix $C = (c_1, \dots, c_n)$: un choix est un critère de recherche
- les opérateurs $O = (o_1, \dots, o_n)$ comparant un document D et un choix C. Opérateurs : intersection, égalité, inclusion, et proximité
- les exigences $E = (e_1, \dots, e_n)$: obligatoire, facultatif, exclus
- les préférences $P = (p_1, \dots, p_n)$ pondèrent chaque critère de recherche en fonction de son importance.



Approche par contraintes - Descriptif général

Le système de combinaison par contraintes modélisé comme un quadruplet (C, O, P, E) :

- les choix $C = (c_1, \dots, c_n)$: un choix est un critère de recherche
- les opérateurs $O = (o_1, \dots, o_n)$ comparant un document D et un choix C. Opérateurs : intersection, égalité, inclusion, et proximité
- les exigences $E = (e_1, \dots, e_n)$: obligatoire, facultatif, exclus
- les préférences $P = (p_1, \dots, p_n)$ pondèrent chaque critère de recherche en fonction de son importance.



Approche par contraintes - Descriptif général

Le système de combinaison par contraintes modélisé comme un quadruplet (C, O, P, E) :

- les choix $C = (c_1, \dots, c_n)$: un choix est un critère de recherche
- les opérateurs $O = (o_1, \dots, o_n)$ comparant un document D et un choix C. Opérateurs : intersection, égalité, inclusion, et proximité
- les exigences $E = (e_1, \dots, e_n)$: obligatoire, facultatif, exclus
- les préférences $P = (p_1, \dots, p_n)$ pondèrent chaque critère de recherche en fonction de son importance.



Approche par contraintes - Descriptif général

Le système de combinaison par contraintes modélisé comme un quadruplet (C, O, P, E) :

- les choix $C = (c_1, \dots, c_n)$: un choix est un critère de recherche
- les opérateurs $O = (o_1, \dots, o_n)$ comparant un document D et un choix C. Opérateurs : intersection, égalité, inclusion, et proximité
- les exigences $E = (e_1, \dots, e_n)$: obligatoire, facultatif, exclus
- les préférences $P = (p_1, \dots, p_n)$ pondèrent chaque critère de recherche en fonction de son importance.



Approche par contraintes - Descriptif général

Le système de combinaison par contraintes modélisé comme un quadruplet (C, O, P, E) :

- les choix $C = (c_1, \dots, c_n)$: un choix est un critère de recherche
- les opérateurs $O = (o_1, \dots, o_n)$ comparant un document D et un choix C. Opérateurs : intersection, égalité, inclusion, et proximité
- les exigences $E = (e_1, \dots, e_n)$: obligatoire, facultatif, exclus
- les préférences $P = (p_1, \dots, p_n)$ pondèrent chaque critère de recherche en fonction de son importance.



Approche par contraintes - Scénarios

- A - Exclusion (filtrage négatif) : exclure un critère
- B - Dévalorisation : défavoriser un critère
- C - Valorisation : favoriser un critère
- D - Cible (Focus) : objet de la recherche
- E - Prérequis (filtrage positif) : obligatoire

	Exigences			Préférences			
	-	N	+	-	N	+	
A	✓				✓		Exclusion (Filtrage négatif)



Approche par contraintes - Scénarios

- A - Exclusion (filtrage négatif) : exclure un critère
- B - Dévalorisation : défavoriser un critère
- C - Valorisation : favoriser un critère
- D - Cible (Focus) : objet de la recherche
- E - Prérequis (filtrage positif) : obligatoire

	Exigences			Préférences			
	-	N	+	-	N	+	
A	✓				✓		Exclusion (Filtrage négatif)
B		✓		✓			Dévalorisation



Approche par contraintes - Scénarios

- A - Exclusion (filtrage négatif) : exclure un critère
- B - Dévalorisation : défavoriser un critère
- C - Valorisation : favoriser un critère
- D - Cible (Focus) : objet de la recherche
- E - Prérequis (filtrage positif) : obligatoire

	Exigences			Préférences			
	-	N	+	-	N	+	
A	✓				✓		Exclusion (Filtrage négatif)
B		✓		✓			Dévalorisation
C		✓				✓	Valorisation



Approche par contraintes - Scénarios

- A - Exclusion (filtrage négatif) : exclure un critère
- B - Dévalorisation : défavoriser un critère
- C - Valorisation : favoriser un critère
- D - Cible (Focus) : objet de la recherche
- E - Prérequis (filtrage positif) : obligatoire

	Exigences			Préférences			
	-	N	+	-	N	+	
A	✓				✓		Exclusion (Filtrage négatif)
B		✓		✓			Dévalorisation
C		✓				✓	Valorisation
D			✓			✓	Cible (Focus)



Approche par contraintes - Scénarios

- A - Exclusion (filtrage négatif) : exclure un critère
- B - Dévalorisation : défavoriser un critère
- C - Valorisation : favoriser un critère
- D - Cible (Focus) : objet de la recherche
- E - Prérequis (filtrage positif) : obligatoire

	Exigences			Préférences			
	-	N	+	-	N	+	
A	✓				✓		Exclusion (Filtrage négatif)
B		✓		✓			Dévalorisation
C		✓				✓	Valorisation
D			✓			✓	Cible (Focus)
E			✓		✓		Prérequis (Filtrage positif)



Approche par contraintes - Scores

- Calcul du score d'un document si exigences satisfaites
- Somme pondérée avec préférences

$$\text{score}(d_j) = \begin{cases} 0 & \text{si } \text{exigences_non_satisfaites}(d_j, E) \\ \frac{1}{\sum_{i=1}^n p_i > 0} * \sum_{i=1}^n (p_i * o_i(d_j, c_i)) & \text{sinon} \end{cases} \quad (2)$$



Approche par contraintes - Scores

- Calcul du score d'un document si exigences satisfaites
- Somme pondérée avec préférences

$$\text{score}(d_j) = \begin{cases} 0 & \text{si } \text{exigences_non_satisfaites}(d_j, E) \\ \frac{1}{\sum_{i=1}^n p_i > 0} * \sum_{i=1}^n (p_i * o_i(d_j, c_i)) & \text{sinon} \end{cases} \quad (2)$$



Approche par contraintes - Scores

- Calcul du score d'un document si exigences satisfaites
- Somme pondérée avec préférences

$$\text{score}(d_j) = \begin{cases} 0 & \text{si } \text{exigences_non_satisfaites}(d_j, E) \\ \frac{1}{\sum_{i=1}^n p_i > 0} * \sum_{i=1}^n (p_i * o_i(d_j, c_i)) & \text{sinon} \end{cases} \quad (2)$$



Approche par contraintes - Exemple d'application

- « les risques accidentels en montagne si possible liés à des balades ou randonnées »
- 3 critères : a) « balade »(valorisation), (b) « montagne »(prérequis), (c) « accident »(cible)

Choix	balade	montagne	accident
Opérateur	égalité	égalité	égalité
Exigence	N	+	+
Préférences	+	N	+

- $d8 : \frac{1}{1+0+1} \times (1 \times 0,9 + 0 \times 0,7 + 1 \times 0,1) = 0,5$
- $d3 : \frac{1}{1+0+1} \times (1 \times 0 + 0 \times 0,7 + 1 \times 0,9) = 0,45$
- $d5 : \frac{1}{1+0+1} \times (1 \times 0,4 + 0 \times 0,3 + 1 \times 0,2) = 0,3$



Approche par contraintes - Exemple d'application

- « les risques accidentels en montagne si possible liés à des balades ou randonnées »
- 3 critères : a) « balade »(valorisation), (b) « montagne »(prérequis), (c) « accident »(cible)

Choix	balade	montagne	accident
Opérateur	égalité	égalité	égalité
Exigence	N	+	+
Préférences	+	N	+

- $d8 : \frac{1}{1+0+1} \times (1 \times 0,9 + 0 \times 0,7 + 1 \times 0,1) = 0,5$
- $d3 : \frac{1}{1+0+1} \times (1 \times 0 + 0 \times 0,7 + 1 \times 0,9) = 0,45$
- $d5 : \frac{1}{1+0+1} \times (1 \times 0,4 + 0 \times 0,3 + 1 \times 0,2) = 0,3$



Approche par contraintes - Exemple d'application

- « les risques accidentels en montagne si possible liés à des balades ou randonnées »
- 3 critères : a) « balade »(valorisation), (b) « montagne »(prérequis), (c) « accident »(cible)

Choix	balade	montagne	accident
Opérateur	égalité	égalité	égalité
Exigence	N	+	+
Préférences	+	N	+

- $d8 : \frac{1}{1+0+1} \times (1 \times 0,9 + 0 \times 0,7 + 1 \times 0,1) = 0,5$
- $d3 : \frac{1}{1+0+1} \times (1 \times 0 + 0 \times 0,7 + 1 \times 0,9) = 0,45$
- $d5 : \frac{1}{1+0+1} \times (1 \times 0,4 + 0 \times 0,3 + 1 \times 0,2) = 0,3$



Approche par contraintes - Exemple d'application

- « les risques accidentels en montagne si possible liés à des balades ou randonnées »
- 3 critères : a) « balade »(valorisation), (b) « montagne »(prérequis), (c) « accident »(cible)

Choix	balade	montagne	accident
Opérateur	égalité	égalité	égalité
Exigence	N	+	+
Préférences	+	N	+

- $d8 : \frac{1}{1+0+1} \times (1 \times 0,9 + 0 \times 0,7 + 1 \times 0,1) = 0,5$
- $d3 : \frac{1}{1+0+1} \times (1 \times 0 + 0 \times 0,7 + 1 \times 0,9) = 0,45$
- $d5 : \frac{1}{1+0+1} \times (1 \times 0,4 + 0 \times 0,3 + 1 \times 0,2) = 0,3$



Approche par contraintes - Exemple d'application

- « les risques accidentels en montagne si possible liés à des balades ou randonnées »
- 3 critères : a) « balade »(valorisation), (b) « montagne »(prérequis), (c) « accident »(cible)

Choix	balade	montagne	accident
Opérateur	égalité	égalité	égalité
Exigence	N	+	+
Préférences	+	N	+

- $d8 : \frac{1}{1+0+1} \times (1 \times 0,9 + 0 \times 0,7 + 1 \times 0,1) = 0,5$
- $d3 : \frac{1}{1+0+1} \times (1 \times 0 + 0 \times 0,7 + 1 \times 0,9) = 0,45$
- $d5 : \frac{1}{1+0+1} \times (1 \times 0,4 + 0 \times 0,3 + 1 \times 0,2) = 0,3$

Approche par contraintes - Exemple d'application

- « les risques accidentels en montagne si possible liés à des balades ou randonnées »
- 3 critères : a) « balade »(valorisation), (b) « montagne »(prérequis), (c) « accident »(cible)

Choix	balade	montagne	accident
Opérateur	égalité	égalité	égalité
Exigence	N	+	+
Préférences	+	N	+

- $d8 : \frac{1}{1+0+1} \times (1 \times 0,9 + 0 \times 0,7 + 1 \times 0,1) = 0,5$
- $d3 : \frac{1}{1+0+1} \times (1 \times 0 + 0 \times 0,7 + 1 \times 0,9) = 0,45$
- $d5 : \frac{1}{1+0+1} \times (1 \times 0,4 + 0 \times 0,3 + 1 \times 0,2) = 0,3$

Approche par contraintes - Exemple d'application

- Préférences proportionnelles améliorent les résultats

Choix	balade	montagne	accident
Opérateur	égalité	égalité	égalité
Exigence	N	+	+
Préférences	0,1	0	0,9

- $d3 : \frac{1}{0,1+0+0,9} \times (0,1 \times 0 + 0 \times 0,7 + 0,9 \times 0,9) = 0,81$
- $d5 : \frac{1}{0,1+0+0,9} \times (0,1 \times 0,4 + 0 \times 0,3 + 0,9 \times 0,2) = 0,22$
- $d8 : \frac{1}{0,1+0+0,9} \times (0,1 \times 0,9 + 0 \times 0,7 + 0,9 \times 0,1) = 0,18$



Approche par contraintes - Exemple d'application

- Préférences proportionnelles améliorent les résultats

Choix	balade	montagne	accident
Opérateur	égalité	égalité	égalité
Exigence	N	+	+
Préférences	0,1	0	0,9

- $d3 : \frac{1}{0,1+0+0,9} \times (0,1 \times 0 + 0 \times 0,7 + 0,9 \times 0,9) = 0,81$

- $d5 : \frac{1}{0,1+0+0,9} \times (0,1 \times 0,4 + 0 \times 0,3 + 0,9 \times 0,2) = 0,22$

- $d8 : \frac{1}{0,1+0+0,9} \times (0,1 \times 0,9 + 0 \times 0,7 + 0,9 \times 0,1) = 0,18$



Approche par contraintes - Exemple d'application

- Préférences proportionnelles améliorent les résultats

Choix	balade	montagne	accident
Opérateur	égalité	égalité	égalité
Exigence	N	+	+
Préférences	0,1	0	0,9

- $d3 : \frac{1}{0,1+0+0,9} \times (0,1 \times 0 + 0 \times 0,7 + 0,9 \times 0,9) = 0,81$
- $d5 : \frac{1}{0,1+0+0,9} \times (0,1 \times 0,4 + 0 \times 0,3 + 0,9 \times 0,2) = 0,22$
- $d8 : \frac{1}{0,1+0+0,9} \times (0,1 \times 0,9 + 0 \times 0,7 + 0,9 \times 0,1) = 0,18$



Après la combinaison ?

Bilan

- Combinaison des différentes facettes via des opérateurs de combinaisons linéaires
- Paramétrisation via l'approche proposée par contraintes

Ensuite

- Comment évaluer l'apport de la combinaison des différentes facettes de l'information géographique ?



Après la combinaison ?

Bilan

- Combinaison des différentes facettes via des opérateurs de combinaisons linéaires
- Paramétrisation via l'approche proposée par contraintes

Ensuite

- Comment évaluer l'apport de la combinaison des différentes facettes de l'information géographique ?



Après la combinaison ?

Bilan

- Combinaison des différentes facettes via des opérateurs de combinaisons linéaires
- Paramétrisation via l'approche proposée par contraintes

Ensuite

- Comment évaluer l'apport de la combinaison des différentes facettes de l'information géographique ?



Plan

- 1 Introduction : La recherche d'information géographique dans des fonds documentaires textuels
- 2 Contribution : Vers la combinaison par contraintes de critères de recherche en RIG
 - Uniformisation générique
 - Recherche d'Information Géographique par combinaison de critères
 - Evaluation de systèmes de recherche d'information géographique
 - Expérimentations
- 3 Conclusion, discussions et perspectives



Evaluations existantes

- RI : traitements performants
 - efficacité en terme de qualité
- RIG : traitements encore lourds
 - efficacité en terme de stockage ou de temps de calcul



Evaluations existantes

- RI : traitements performants
 - efficacité en terme de qualité
- RIG : traitements encore lourds
 - efficacité en terme de stockage ou de temps de calcul



Evaluations existantes

- RI : traitements performants
 - efficacité en terme de qualité
- RIG : traitements encore lourds
 - efficacité en terme de stockage ou de temps de calcul



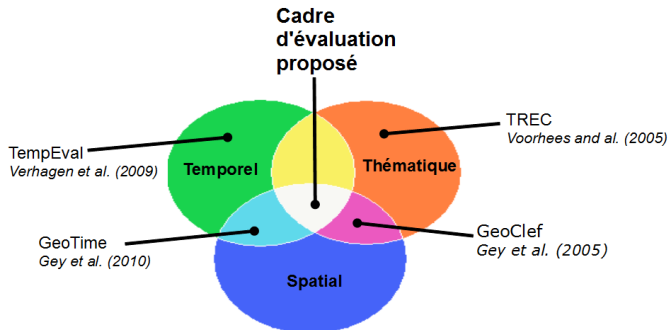
Evaluations existantes

- RI : traitements performants
 - efficacité en terme de qualité
- RIG : traitements encore lourds
 - efficacité en terme de stockage ou de temps de calcul



Evaluations existantes

- RI : traitements performants
 - efficacité en terme de qualité
- RIG : traitements encore lourds
 - efficacité en terme de stockage ou de temps de calcul



Constitution d'une collection de test

Adaptation de la méthodologie de TREC [Har05, VH05] : topic (requêtes), corpus et qrels. Particularités :

- *qrels* par facette : assesseur évalue l'adéquation entre le document et la requête (topic). Pertinence $\in \{0; 1; 2; 3; 4\}$ (topic : « Pyrénées entre 1800 et 1900 », doc : balades dans les Pyrénées + Henri est décédé en 1800)
- *ressources géographiques* nécessaires (précises difficiles à obtenir)



Constitution d'une collection de test

Adaptation de la méthodologie de TREC [Har05, VH05] : topic (requêtes), corpus et qrels. Particularités :

- *qrels* par facette : assesseur évalue l'adéquation entre le document et la requête (topic). Pertinence $\in \{0; 1; 2; 3; 4\}$ (topic : « Pyrénées entre 1800 et 1900 », doc : balades dans les Pyrénées + Henri est décédé en 1800)
- *ressources géographiques* nécessaires (précises difficiles à obtenir)



Constitution d'une collection de test

Adaptation de la méthodologie de TREC [Har05, VH05] : topic (requêtes), corpus et qrels. Particularités :

- *qrels* par facette : assesseur évalue l'adéquation entre le document et la requête (topic). Pertinence $\in \{0; 1; 2; 3; 4\}$ (topic : « Pyrénées entre 1800 et 1900 », doc : balades dans les Pyrénées + Henri est décédé en 1800)
- *ressources géographiques* nécessaires (précises difficiles à obtenir)



Protocole d'analyse comparative de SRI géographiques

- Evaluation vise à mesurer l'efficacité relative des SRI suivants :
 - SRI monofacette : thématique (Th), spatial (S) et temporel (Te) ;
 - SRI bifacette : Th+S, Th+Te et S+Te permettant de mesurer l'apport de chacune des facettes dans l'efficacité du SRI ;
 - SRI géographique combinant les trois facettes : Th+S+Te.
- Pour chaque topic, mesure de pertinence graduelle *Normalized Discounted Cumulative Gain (NDCG)* proposée par Järvelin et al. [JK02]
- Pour mesurer SRI, moyenne des NDCG



Protocole d'analyse comparative de SRI géographiques

- Evaluation vise à mesurer l'efficacité relative des SRI suivants :
 - SRI monofacette : thématique (Th), spatial (S) et temporel (Te) ;
 - SRI bifacette : Th+S, Th+Te et S+Te permettant de mesurer l'apport de chacune des facettes dans l'efficacité du SRI ;
 - SRI géographique combinant les trois facettes : Th+S+Te.
- Pour chaque topic, mesure de pertinence graduelle *Normalized Discounted Cumulative Gain (NDCG)* proposée par Järvelin et al. [JK02]
- Pour mesurer SRI, moyenne des NDCG



Protocole d'analyse comparative de SRI géographiques

- Evaluation vise à mesurer l'efficacité relative des SRI suivants :
 - SRI monofacette : thématique (Th), spatial (S) et temporel (Te) ;
 - SRI bifacette : Th+S, Th+Te et S+Te permettant de mesurer l'apport de chacune des facettes dans l'efficacité du SRI ;
 - SRI géographique combinant les trois facettes : Th+S+Te.
- Pour chaque topic, mesure de pertinence graduelle *Normalized Discounted Cumulative Gain (NDCG)* proposée par Järvelin et al. [JK02]
- Pour mesurer SRI, moyenne des NDCG



Protocole d'analyse comparative de SRI géographiques

- Evaluation vise à mesurer l'efficacité relative des SRI suivants :
 - SRI monofacette : thématique (Th), spatial (S) et temporel (Te) ;
 - SRI bifacette : Th+S, Th+Te et S+Te permettant de mesurer l'apport de chacune des facettes dans l'efficacité du SRI ;
 - SRI géographique combinant les trois facettes : Th+S+Te.
- Pour chaque topic, mesure de pertinence graduelle *Normalized Discounted Cumulative Gain (NDCG)* proposée par Järvelin et al. [JK02]
- Pour mesurer SRI, moyenne des NDCG



Protocole d'analyse comparative de SRI géographiques

- Evaluation vise à mesurer l'efficacité relative des SRI suivants :
 - SRI monofacette : thématique (Th), spatial (S) et temporel (Te) ;
 - SRI bifacette : Th+S, Th+Te et S+Te permettant de mesurer l'apport de chacune des facettes dans l'efficacité du SRI ;
 - SRI géographique combinant les trois facettes : Th+S+Te.
- Pour chaque topic, mesure de pertinence graduelle *Normalized Discounted Cumulative Gain (NDCG)* proposée par Järvelin et al. [JK02]
- Pour mesurer SRI, moyenne des NDCG



Protocole d'analyse comparative de SRI géographiques

- Evaluation vise à mesurer l'efficacité relative des SRI suivants :
 - SRI monofacette : thématique (Th), spatial (S) et temporel (Te) ;
 - SRI bifacette : Th+S, Th+Te et S+Te permettant de mesurer l'apport de chacune des facettes dans l'efficacité du SRI ;
 - SRI géographique combinant les trois facettes : Th+S+Te.
- Pour chaque topic, mesure de pertinence graduelle *Normalized Discounted Cumulative Gain (NDCG)* proposée par Järvelin et al. [JK02]
- Pour mesurer SRI, moyenne des NDCG



Mise en oeuvre avec MIDR2010

- *Corpus* de 5645 paragraphes issus de 11 ouvrages numérisés, 8983 informations spatiales et 1702 informations temporelles
- 34 *topics* couvrant tout ou partie des 3 facettes de l'information géographique
- *Qrels* obtenus en interrogeant trois SRI : thématique, spatial et temporel (12 assesseurs bénévoles ont participé, merci encore !)
- Ressources géographiques liées au corpus issues de l'Institut Géographique National (BD NYME[®]) et d'un *gazetteer* (un dictionnaire géographique) contributif local



Mise en oeuvre avec MIDR2010

- *Corpus* de 5645 paragraphes issus de 11 ouvrages numérisés, 8983 informations spatiales et 1702 informations temporelles
- 34 *topics* couvrant tout ou partie des 3 facettes de l'information géographique
- *Qrels* obtenus en interrogeant trois SRI : thématique, spatial et temporel (12 assesseurs bénévoles ont participé, merci encore !)
- Ressources géographiques liées au corpus issues de l'Institut Géographique National (BD NYME[®]) et d'un *gazetteer* (un dictionnaire géographique) contributif local



Mise en oeuvre avec MIDR2010

- *Corpus* de 5645 paragraphes issus de 11 ouvrages numérisés, 8983 informations spatiales et 1702 informations temporelles
- 34 *topics* couvrant tout ou partie des 3 facettes de l'information géographique
- *Qrels* obtenus en interrogeant trois SRI : thématique, spatial et temporel (12 assesseurs bénévoles ont participé, merci encore !)
- Ressources géographiques liées au corpus issues de l'Institut Géographique National (BD NYME[®]) et d'un *gazetteer* (un dictionnaire géographique) contributif local



Mise en oeuvre avec MIDR2010

- *Corpus* de 5645 paragraphes issus de 11 ouvrages numérisés, 8983 informations spatiales et 1702 informations temporelles
- 34 *topics* couvrant tout ou partie des 3 facettes de l'information géographique
- *Qrels* obtenus en interrogeant trois SRI : thématique, spatial et temporel (12 assesseurs bénévoles ont participé, merci encore !)
- Ressources géographiques liées au corpus issues de l'Institut Géographique National (BD NYME[®]) et d'un *gazetteer* (un dictionnaire géographique) contributif local



Après le cadre d'évaluation ?

Bilan

- Cadre permettant d'évaluer des SRI traitant une ou plusieurs facettes de l'information géographique
- Mise en oeuvre du cadre d'évaluation avec la collection MIDR_2010

Ensuite

- Vérifier expérimentalement la viabilité de l'uniformisation et l'apport de la combinaison



Après le cadre d'évaluation ?

Bilan

- Cadre permettant d'évaluer des SRI traitant une ou plusieurs facettes de l'information géographique
- Mise en oeuvre du cadre d'évaluation avec la collection MIDR_2010

Ensuite

- Vérifier expérimentalement la viabilité de l'uniformisation et l'apport de la combinaison



Après le cadre d'évaluation ?

Bilan

- Cadre permettant d'évaluer des SRI traitant une ou plusieurs facettes de l'information géographique
- Mise en oeuvre du cadre d'évaluation avec la collection MIDR_2010

Ensuite

- Vérifier expérimentalement la viabilité de l'uniformisation et l'apport de la combinaison



Plan

- 1 Introduction : La recherche d'information géographique dans des fonds documentaires textuels
- 2 Contribution : Vers la combinaison par contraintes de critères de recherche en RIG
 - Uniformisation générique
 - Recherche d'Information Géographique par combinaison de critères
 - Evaluation de systèmes de recherche d'information géographique
 - Expérimentations
- 3 Conclusion, discussions et perspectives



Uniformisation spatiale

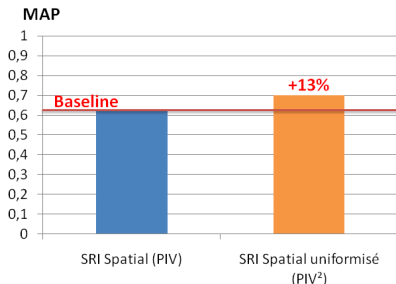
- Approche d'uniformisation a-t-elle un intérêt ? Perte de précision trop importante par rapport à un SRI spatial classique (en général basé sur l'intersection ou la distance [JP08]) tel que notre système PIV[Les07]

- Fréquences des tuiles :
compense pertes dues au
tuilage
- Amélioration significative
des résultats (T-test $< 0,05$)



Uniformisation spatiale

- Approche d'uniformisation a-t-elle un intérêt ? Perte de précision trop importante par rapport à un SRI spatial classique (en général basé sur l'intersection ou la distance [JP08]) tel que notre système PIV[Les07]

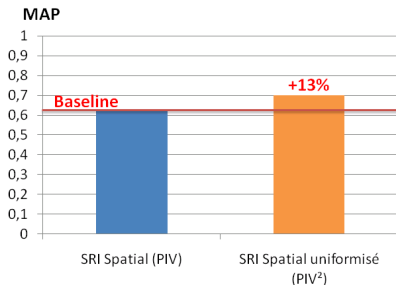


- Fréquences des tuiles :
compense pertes dues au
tuilage
- Amélioration significative
des résultats (T-test < 0,05)



Uniformisation spatiale

- Approche d'uniformisation a-t-elle un intérêt ? Perte de précision trop importante par rapport à un SRI spatial classique (en général basé sur l'intersection ou la distance [JP08]) tel que notre système PIV[Les07]

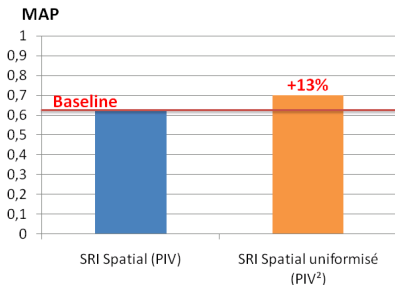


- Fréquences des tuiles :
compense pertes dues au
tuilage
- Amélioration significative
des résultats (T-test < 0,05)



Uniformisation spatiale

- Approche d'uniformisation a-t-elle un intérêt ? Perte de précision trop importante par rapport à un SRI spatial classique (en général basé sur l'intersection ou la distance [JP08]) tel que notre système PIV[Les07]



- Fréquences des tuiles :
compense pertes dues au
tuilage
- Amélioration significative
des résultats (T-test < 0,05)



Uniformisation spatiale

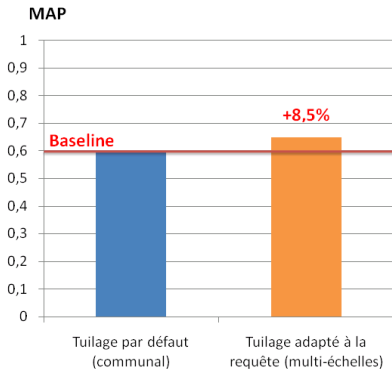
- Utiliser un index adapté à la granularité de la requête améliore-t-il les résultats ?

- Amélioration qualité non significative (T-test $> 0,05$)
- Amélioration rapidité d'accès



Uniformisation spatiale

- Utiliser un index adapté à la granularité de la requête améliore-t-il les résultats ?

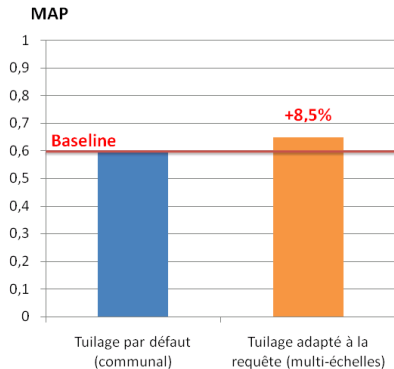


- Amélioration qualité non significative (T-test $> 0,05$)
- Amélioration rapidité d'accès



Uniformisation spatiale

- Utiliser un index adapté à la granularité de la requête améliore-t-il les résultats ?

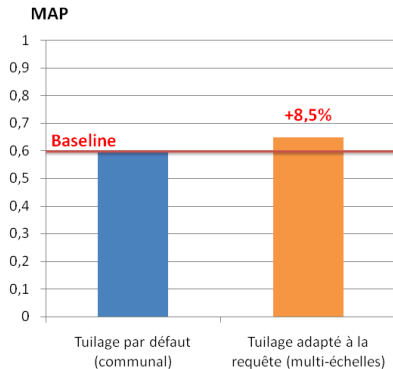


- Amélioration qualité non significative (T-test $> 0,05$)
- Amélioration rapidité d'accès



Uniformisation spatiale

- Utiliser un index adapté à la granularité de la requête améliore-t-il les résultats ?



- Amélioration qualité non significative (T-test $> 0,05$)
- Amélioration rapidité d'accès



Uniformisation temporelle

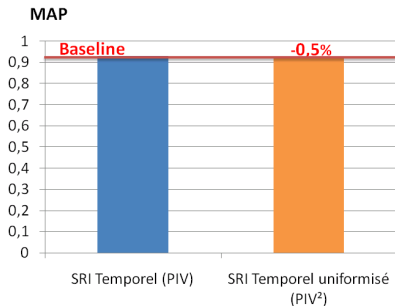
- Approche d'uniformisation appliquée à l'information temporelle donne-t-elle de bons résultats comme pour l'information spatiale ?

- Fréquences ne compensent pas (peu d'infos tempo)
- Différence non significative (T-test $> 0,05$)



Uniformisation temporelle

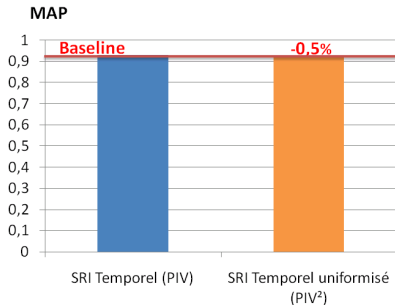
- Approche d'uniformisation appliquée à l'information temporelle donne-t-elle de bons résultats comme pour l'information spatiale ?



- Fréquences ne compensent pas (peu d'infos tempo)
- Différence non significative (T-test > 0,05)

Uniformisation temporelle

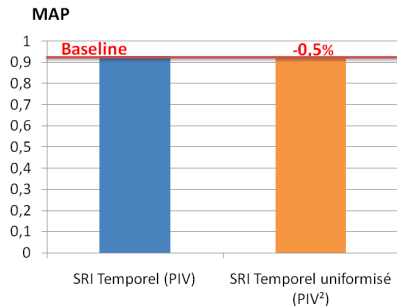
- Approche d'uniformisation appliquée à l'information temporelle donne-t-elle de bons résultats comme pour l'information spatiale ?



- Fréquences ne compensent pas (peu d'infos tempo)
- Différence non significative (T-test > 0,05)

Uniformisation temporelle

- Approche d'uniformisation appliquée à l'information temporelle donne-t-elle de bons résultats comme pour l'information spatiale ?



- Fréquences ne compensent pas (peu d'infos tempo)
- Différence non significative (T-test > 0,05)

Combinaisons linéaires

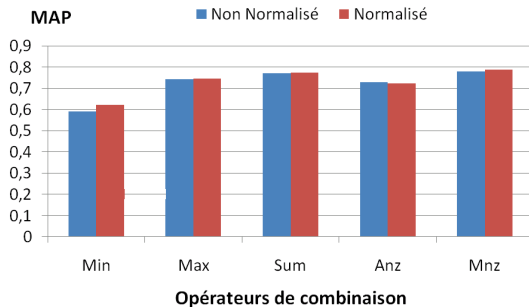
- Meilleur opérateur de combinaison linéaire ?

- CombMNZ
normalisé comme
en RI
[FS93, Lee97]



Combinaisons linéaires

• Meilleur opérateur de combinaison linéaire ?

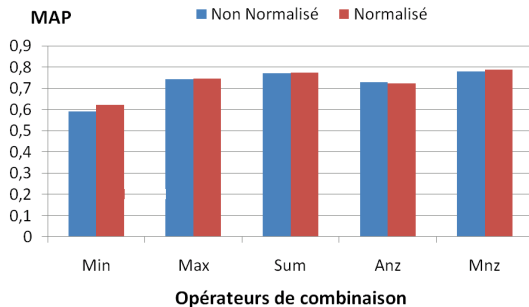


- CombMNZ normalisé comme en RI [FS93, Lee97]



Combinaisons linéaires

• Meilleur opérateur de combinaison linéaire ?



- CombMNZ normalisé comme en RI [FS93, Lee97]



Combinaisons linéaires

- Comparaison des différentes combinaisons possibles
- Apport de la combinaison des différentes facettes de l'information géographique ?



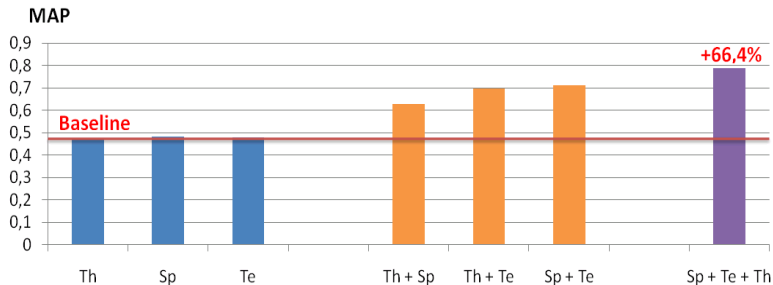
Combinaisons linéaires

- Comparaison des différentes combinaisons possibles
- Apport de la combinaison des différentes facettes de l'information géographique ?



Combinaisons linéaires

- Comparaison des différentes combinaisons possibles
- Apport de la combinaison des différentes facettes de l'information géographique ?



Thématique (Th)										
Spatial (Sp)										
Temporel (Te)										

Combinaisons par contraintes

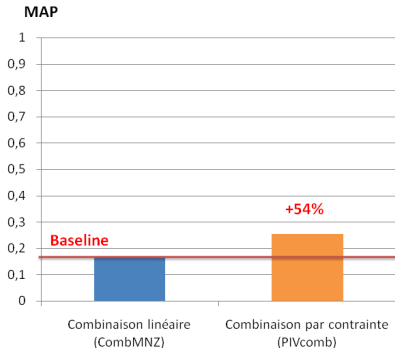
- Comparaison PIVComb avec CombMNZ

- Amélioration de la qualité des résultats
- Test avec 10 requêtes contraintes



Combinaisons par contraintes

● Comparaison PIVComb avec CombMNZ

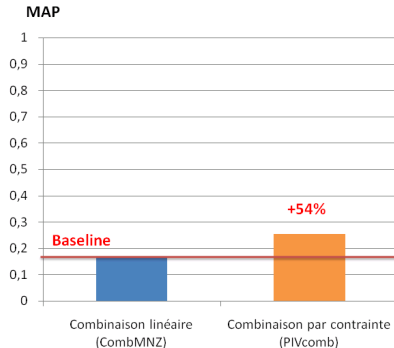


- Amélioration de la qualité des résultats
- Test avec 10 requêtes contraintes



Combinaisons par contraintes

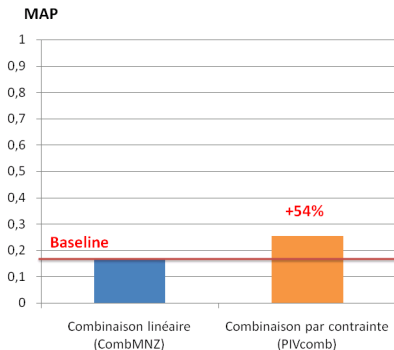
● Comparaison PIVComb avec CombMNZ



- Amélioration de la qualité des résultats
- Test avec 10 requêtes contraintes

Combinaisons par contraintes

● Comparaison PIVComb avec CombMNZ



- Amélioration de la qualité des résultats
- Test avec 10 requêtes contraintes



Bilan expérimentations

- Uniformisation appliquée aux informations spatiales et temporelles viables
- Combinaison des différentes facettes de l'information géographique améliore de manière significative la pertinence des résultats



Bilan expérimentations

- Uniformisation appliquée aux informations spatiales et temporelles viables
- Combinaison des différentes facettes de l'information géographique améliore de manière significative la pertinence des résultats



Plan

- 1 Introduction : La recherche d'information géographique dans des fonds documentaires textuels
- 2 Contribution : Vers la combinaison par contraintes de critères de recherche en RIG
- 3 Conclusion, discussions et perspectives
 - Bilan
 - Perspectives



Uniformisations

Avantages

- Approche générique
- Approche adaptée de la RI classique → compatible avec les SRI textuels
- Uniformisation spatiale améliore les résultats, uniformisation temporelle donne résultats équivalents

Limites

- Expérimenté uniquement sur 1 et 2 dimensions (temporel et spatial)



Uniformisations

Avantages

- Approche générique
- Approche adaptée de la RI classique → compatible avec les SRI textuels
- Uniformisation spatiale améliore les résultats, uniformisation temporelle donne résultats équivalents

Limites

- Expérimenté uniquement sur 1 et 2 dimensions (temporel et spatial)



Uniformisations

Avantages

- Approche générique
- Approche adaptée de la RI classique → compatible avec les SRI textuels
- Uniformisation spatiale améliore les résultats, uniformisation temporelle donne résultats équivalents

Limites

- Expérimenté uniquement sur 1 et 2 dimensions (temporel et spatial)



Uniformisations

Avantages

- Approche générique
- Approche adaptée de la RI classique → compatible avec les SRI textuels
- Uniformisation spatiale améliore les résultats, uniformisation temporelle donne résultats équivalents

Limites

- Expérimenté uniquement sur 1 et 2 dimensions (temporel et spatial)



Combinaisons linéaires

Avantages

- Faciles à utiliser et rapides
- Amélioration de la pertinence des résultats pour la RIG (+66,4 %)

Inconvénients

- Pas de paramétrisation de l'utilisateur
- Requêtes contraintes non possibles



Combinaisons linéaires

Avantages

- Faciles à utiliser et rapides
- Amélioration de la pertinence des résultats pour la RIG (+66,4 %)

Inconvénients

- Pas de paramétrisation de l'utilisateur
- Requêtes contraintes non possibles

Combinaisons linéaires

Avantages

- Faciles à utiliser et rapides
- Amélioration de la pertinence des résultats pour la RIG (+66,4 %)

Inconvénients

- Pas de paramétrisation de l'utilisateur
- Requêtes contraintes non possibles



Combinaisons linéaires

Avantages

- Faciles à utiliser et rapides
- Amélioration de la pertinence des résultats pour la RIG (+66,4 %)

Inconvénients

- Pas de paramétrisation de l'utilisateur
- Requêtes contraintes non possibles



Combinaison par contraintes

Avantages

- Amélioration de la pertinence des résultats
- Application possible à d'autres critères
- Possibilité d'étendre des SRI classiques

Limites

- Expérimentation sur 10 requêtes types
- Non prise en compte de groupes de critères :
montagnes ET (Pyrénées OU Alpes)

Combinaison par contraintes

Avantages

- Amélioration de la pertinence des résultats
- Application possible à d'autres critères
- Possibilité d'étendre des SRI classiques

Limites

- Expérimentation sur 10 requêtes types
- Non prise en compte de groupes de critères :
montagnes ET (Pyrénées OU Alpes)



Combinaison par contraintes

Avantages

- Amélioration de la pertinence des résultats
- Application possible à d'autres critères
- Possibilité d'étendre des SRI classiques

Limites

- Expérimentation sur 10 requêtes types
- Non prise en compte de groupes de critères :
montagnes ET (Pyrénées OU Alpes)

Combinaison par contraintes

Avantages

- Amélioration de la pertinence des résultats
- Application possible à d'autres critères
- Possibilité d'étendre des SRI classiques

Limites

- Expérimentation sur 10 requêtes types
- Non prise en compte de groupes de critères :
montagnes ET (Pyrénées OU Alpes)



Combinaison par contraintes

Avantages

- Amélioration de la pertinence des résultats
- Application possible à d'autres critères
- Possibilité d'étendre des SRI classiques

Limites

- Expérimentation sur 10 requêtes types
- Non prise en compte de groupes de critères :
montagnes ET (Pyrénées OU Alpes)



Cadre d'évaluation

Avantages

- Permet d'évaluer les 3 facettes de l'information géographique
- Extensible à d'autres critères
- Mis en oeuvre avec la collection de test MIDR

Limites

- Collection MIDR en français
- Collection MIDR de 5645 paragraphes



Cadre d'évaluation

Avantages

- Permet d'évaluer les 3 facettes de l'information géographique
- Extensible à d'autres critères
- Mis en oeuvre avec la collection de test MIDR

Limites

- Collection MIDR en français
- Collection MIDR de 5645 paragraphes



Cadre d'évaluation

Avantages

- Permet d'évaluer les 3 facettes de l'information géographique
- Extensible à d'autres critères
- Mis en oeuvre avec la collection de test MIDR

Limites

- Collection MIDR en français
- Collection MIDR de 5645 paragraphes



Cadre d'évaluation

Avantages

- Permet d'évaluer les 3 facettes de l'information géographique
- Extensible à d'autres critères
- Mis en oeuvre avec la collection de test MIDR

Limites

- Collection MIDR en français
- Collection MIDR de 5645 paragraphes



Cadre d'évaluation

Avantages

- Permet d'évaluer les 3 facettes de l'information géographique
- Extensible à d'autres critères
- Mis en oeuvre avec la collection de test MIDR

Limites

- Collection MIDR en français
- Collection MIDR de 5645 paragraphes

Plan

- 1 Introduction : La recherche d'information géographique dans des fonds documentaires textuels
- 2 Contribution : Vers la combinaison par contraintes de critères de recherche en RIG
- 3 Conclusion, discussions et perspectives**
 - Bilan
 - **Perspectives**



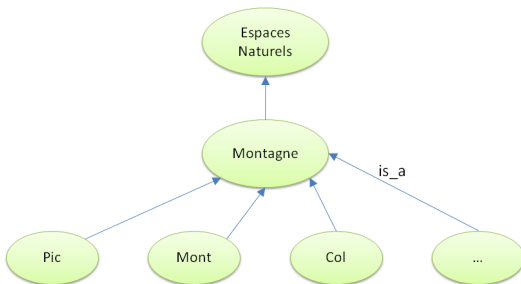
Perspective : uniformisation thématique

- Souvent, plusieurs termes portent sur le même thème (exemple : automobile et voiture)
- Ressources supplémentaires (ontologies)
- Appliquer l'approche d'uniformisation au thématique ?
- Sur combien de dimensions représenter ce type d'information ?



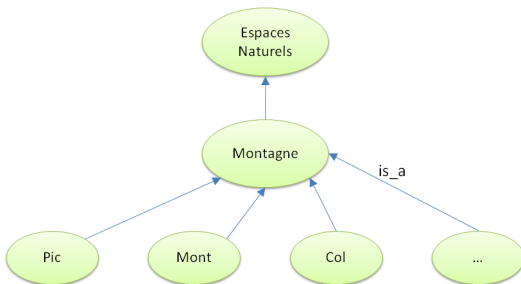
Perspective : uniformisation thématique

- Souvent, plusieurs termes portent sur le même thème (exemple : automobile et voiture)
- Ressources supplémentaires (ontologies)
- Appliquer l'approche d'uniformisation au thématique ?
- Sur combien de dimensions représenter ce type d'information ?



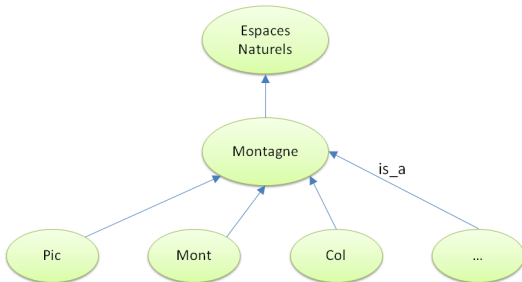
Perspective : uniformisation thématique

- Souvent, plusieurs termes portent sur le même thème (exemple : automobile et voiture)
- Ressources supplémentaires (ontologies)
- Appliquer l'approche d'uniformisation au thématique ?
- Sur combien de dimensions représenter ce type d'information ?



Perspective : uniformisation thématique

- Souvent, plusieurs termes portent sur le même thème (exemple : automobile et voiture)
- Ressources supplémentaires (ontologies)
- Appliquer l'approche d'uniformisation au thématique ?
- Sur combien de dimensions représenter ce type d'information ?



Perspective : combinaison par contraintes

- Continuer les assessements de requêtes
- Expérimentation sur corpus plus conséquent (TREC) ?



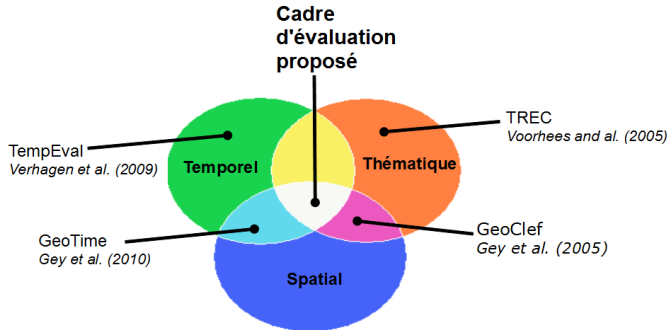
Perspective : combinaison par contraintes

- Continuer les assessements de requêtes
- Expérimentation sur corpus plus conséquent (TREC) ?



Perspective : cadre d'évaluation

- Nouveau track sur campagne existante ?



Interfaces adaptées : spatial

- Requête : « autour de Laruns »

RECHERCHE

autour de Laruns

Interprétation graphique

Ajouter une ligne Ajouter une zone Modifier une ligne / zone Déplacer une ligne / zone Supprimer une ligne / zone

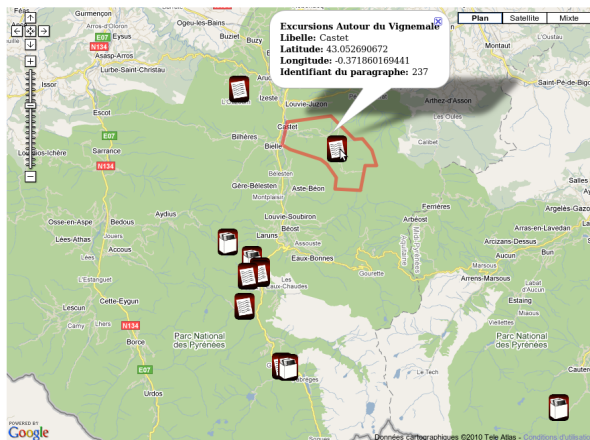
Modification d'une ligne ou zone
Pour modifier une ligne ou une zone, sélectionner une des punaises et la déplacer.

Plan Satellite Mixte

Données cartographiques ©2010 Tele Atlas

Lancer la recherche

Interfaces adaptées : spatial

Excursions Autour du Vigne**Paragraphe numéro 237**

[...] celle des Isards jusqu' à la croix de la Mousquère de l' Homme ; et de celle-ci à la croix de la Rapardère ; et de celle-ci à la croix de ' la Linde , et de celle de la Linde au Turoun de l' Eu (lac) d' ilhéu ; et de ce Turoun aux croix du Col déra Héugade ; et de celles-ci aux deux croix de Castèl-Abarque d' Arras et Sireix ; et de celles-ci à la croix du sommet du Pic déra Cardinquère , Et descendant dudit Pic au Plaia de l' Aubassole , et au Castèl dudit Plaia . Près d' un ruisseau , il y a une croix au Cuyèu déra Labasse ; et de celle du Cuyèu déra Labasse à une autre croix du Bout de Terre déra Labasse qui est la dernière borne faisant la fin des montagnes d' Arras et Sireix , qui confrontent avec celles dudit abbé et des religieux depuis la croix des Isards qui est près du Pic de Mountré et de Brum . [...]

Image Texte Vidéo Son Documents évoquant le même endroit



Interfaces adaptées : temporel

- Requête : « vers 1903 »

Recherche temporelle



Formulation de la recherche

Formulez votre recherche : vers 1903

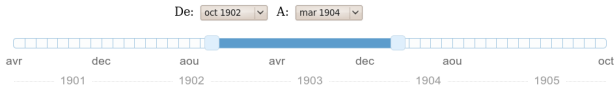
Ok

Vous avez saisi: vers 1903



Votre requete a été interprétée comme une période allant du **01 Octobre 1902** au **31 Mars 1904**

Vous pouvez modifier cette période soit en déplaçant les curseurs soit en choisissant directement une date dans les listes déroulantes



Chercher

Interfaces adaptées : temporel

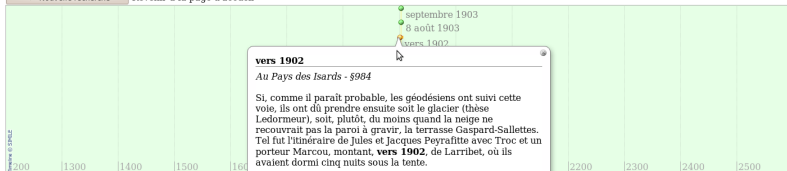
Recherche temporelle



Votre recherche est : **vers 1903**
Cette page affiche les résultats de la recherche de documents pour des dates allant du **01 Octobre 1902** au **31 Mars 1904**
N'hésitez pas à consulter [l'aide](#) pour connaître les fonctionnalités de cette page



<- Modifier cette recherche Modifier la période de recherche
<- Nouvelle recherche Revenir à la page d'accueil

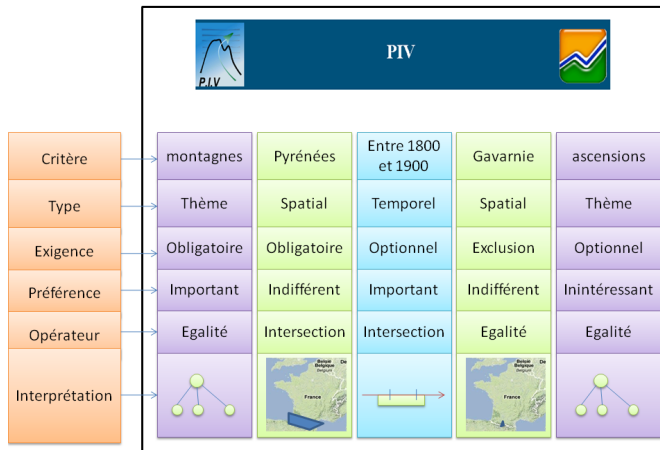


résultats 1 à 10 sur 33

- septembre 1903
 - [Au Pays des Isards - \\$692](#)
- 8 août 1903
 - [Excursions Autour du Vignemale - \\$469](#)
- vers 1902
 - [Au Pays des Isards - \\$984](#)



Interfaces adaptées : combinaison



Merci pour votre attention !





Ricardo Baeza-Yates and Berthier Ribeiro-Neto.

Modern Information Retrieval.

ACM Press, New York, 1999.



Paul Clough, Hideo Joho, and Ross Purves.

Judging the Spatial Relevance of Documents for GIR.

In *ECIR'06 : Proceedings of the 28th European Conference on IR Research*, volume 3936 of *Lecture Notes in Computer Science*, pages 548–552. Springer, 2006.



Gaussier Eric, GREFENSTETTE Gregory, HULL David, and ROUX Claude.

Recherche d'information en français et traitement automatique des langues.

Traitement automatique des langues, 41(2) :473–493, 2000.



Edward A. Fox and Joseph A. Shaw.

Combinaison de Multiple Searches.



In Donna K. Harman, editor, *TREC-1 : Proceedings of the First Text REtrieval Conference*, pages 243–252, Gaithersburg, MD, USA, February 1993. NIST.



Mauro Gaio.

Traitements de l'Information Géographique : Représentations et Structures.

In Mémoire d'HDR, Université de Caen, 2001.



Donna K. Harman.

The TREC Test Collections.

In Ellen M. Voorhees and Donna K. Harman, editors, *TREC : Experiment and Evaluation in Information Retrieval*, chapter 2, pages 21–53. MIT Press, 2005.



Kalervo Järvelin and Jaana Kekäläinen.

Cumulated gain-based evaluation of ir techniques.

ACM Trans. Inf. Syst., 20(4) :422–446, 2002.





Christopher B. Jones and Ross S. Purves.

Geographical information retrieval.

International Journal of Geographical Information Science,
22(3) :219–228, 2008.



Joon Ho Lee.

Analyses of Multiple Evidence Combination.

In *SIGIR'97 : Proceedings of the 20th annual international ACM SIGIR conference*, pages 267–276, New York, NY, USA, 1997.
ACM Press.



J. Lesbegueries.

Plate-forme pour l'indexation spatiale multi-niveaux d'un corpus territorialisé.

PhD thesis, Université de Pau et des Pays de l'Adour, 2007.



Bruno Martins, Mário J. Silva, and Leonardo Andrade.

Indexing and ranking in Geo-IR systems.



In *GIR '05 : Proceedings of the 2005 workshop on Geographic information retrieval*, pages 31–34, New York, NY, USA, 2005. ACM.



Trong-Ton Pham, Nicolas Maillot, Joo-Hwee Lim, and Jean-Pierre Chevallet.

Latent semantic fusion model for image retrieval and annotation. In Mário J. Silva and Alberto H. F. Laender and Ricardo A. Baeza-Yates and Deborah L. McGuinness and Bjørn Olstad and Øystein Haug Olsen and André O. Falcão, editor, *CIKM*, pages 439–444. ACM, 2007.



G. Salton, C. S. Yang, and C. T. Yu.

A theory of term importance in automatic text analysis. *Journal of the American Society for Information Science*, 26(1) :33–44, 1975.



E. Lynn Usery.

A feature-based geographic information system model.



Photogrammetric Engineering & Remote Sensing,
62(7) :833–838, 1996.



Ellen M. Voorhees and Donna K. Harman.

TREC : Experiment and Evaluation in Information Retrieval.
MIT Press, Cambridge, MA, USA, 2005.

