



Utilisation de données numériques expérientielles en discours spécialisés : *deux exemples dans le domaine du tourisme et de l'oénologie*

TEDonnées 2017

« *Techniques Émergentes des Données : veille et e-réputation* »

24 Novembre 2017 – CLA - Besançon

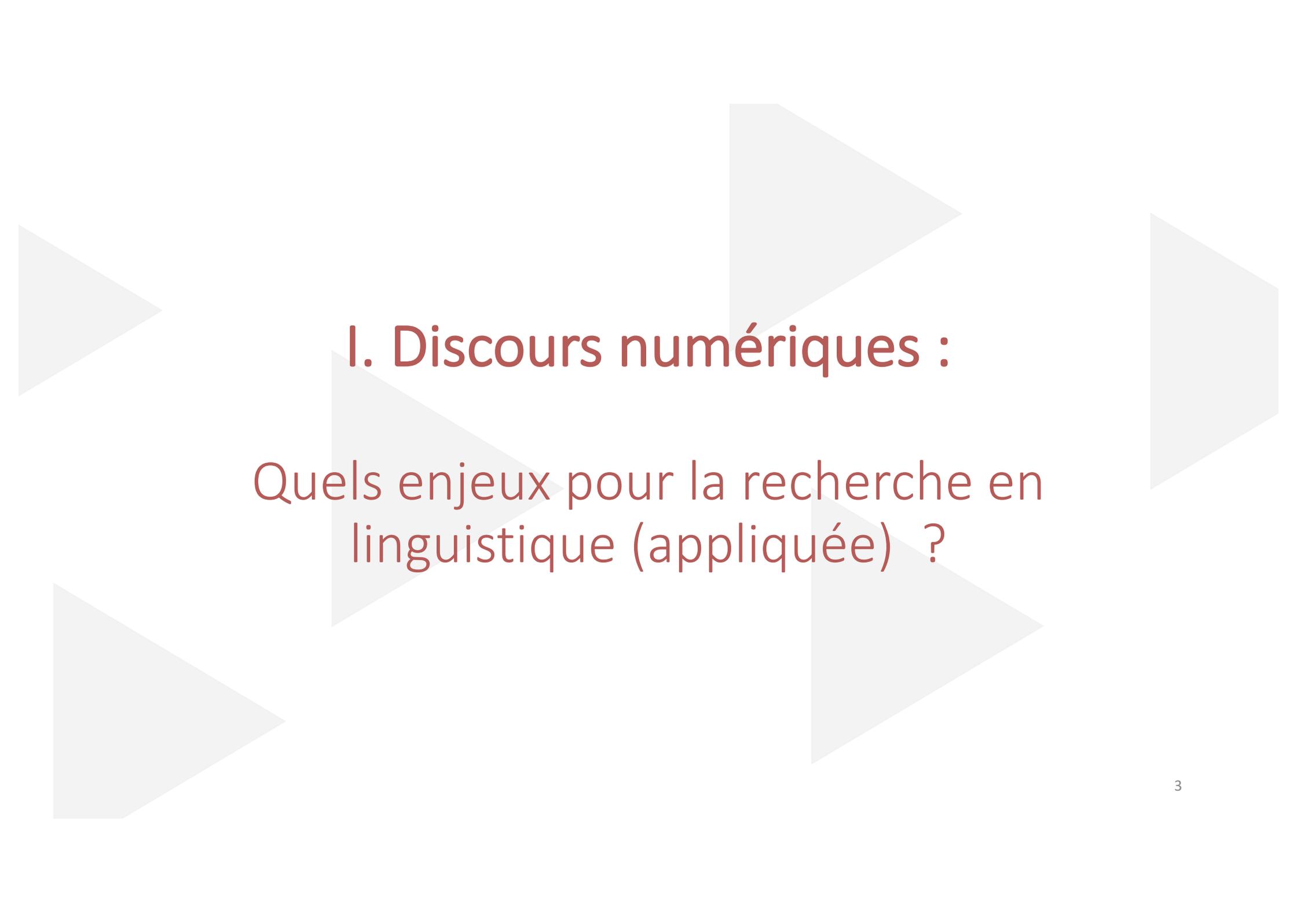
Laurent GAUTIER, PR, Centre Interlangues Texte Image Langage (UBFC, EA 4182)

Cyril NGUYEN VAN, IE CNRS, Maison des Sciences de l'Homme Dijon (USR CNRS-uB 3516)

Matthieu BACH, Doctorant, Centre Interlangues Texte Image Langage (UBFC, EA 4182)

Structure

1. **Discours numériques** : Quels enjeux pour la recherche en linguistique (appliquée) ?
2. **TouristYou** : un « E-Greeters » dans votre poche
3. **AdWine** : Système de recommandation intelligent pour le vin et la gastronomie
4. **Perspectives**



I. Discours numériques :

Quels enjeux pour la recherche en linguistique (appliquée) ?

Discours numériques : Quels enjeux pour la recherche en linguistique (appliquée) ?

Éléments de contexte

Tournant / révolution / transition numérique

- Révolution des espaces de stockage et des temps de traitement des données
 - => explosion des masses de données disponibles
- Appropriation du web 2.0 (web social)
 - ⇒ contribution à l'échange des informations et des interactions
 - ⇒ nouvelles formes de communication avec apparition de nouveaux genres discursifs
 - ⇒ Nécessité de saisie par la recherche linguistique (avis et classement de consos, blogs/forums d'amateurs, etc.)
 - ⇒ Nécessité de nouvelles formes de traitement comme « corpus intelligent »
- Sous couvert de multimodalité
 - => omniprésence des données langagières

Discours numériques : Quels enjeux pour la recherche en linguistique (appliquée) ?
Éléments de contexte

***Big et linked data* : un 'terrain de jeu' pour linguistes**

Enjeux reposant sur une maîtrise et le développement d'outils linguistiques :

- **Extraction de l'information** : traitement de la syntaxe (y compris des formes émergentes, par exemple liées au #), connaissances sémantiques, modélisation des savoirs encyclopédiques
- **Circulation de l'information** : traitement de marqueurs de polyphonie, d'implicite, d'ironie, etc.
- **Représentation / visualisation de l'information** : création de banques de données intelligentes, techniques d'annotation, d'apprentissage machine

Discours numériques : Quels enjeux pour la recherche en linguistique (appliquée) ?
Un défi pour la linguistique (appliquée) : quelques exemples

L'analyse des discours évaluatifs pour l'extraction d'opinions



tripelle69
Lyon, France
👤 145 🍷 51



Reviewed 6 days ago

bon hotel a paris

[Google Translation](#)

Enfin un hotel ou l'on se sent bien, belle chambre, propre, bien décorée, calme...seul bémol l'accueil qui laisse à désirer, manque de chaleur..; petit déjeuner fort agréable, peut être victime de leur succès les prix sont élevés et l'hotel souvent complet

- Essor des sites d'avis de consommateurs
- Traitement du lexique valencé (+, neutre, -) : création de dictionnaires appropriés + traitement de la 'tonalité' des textes
- Multiplication des domaines : tourisme, achats, finances...

Discours numériques : Quels enjeux pour la recherche en linguistique (appliquée) ?
Un défi pour la linguistique (appliquée) : quelques exemples

***L'analyse de blogs / forums pour l'extraction d'informations
« terminologiques » en vue de la recommandation***

Re: [Domaine Morey-Coffinet](#)

□ par [tpinault](#) » Mar 9 Fév 2010 13:18

Voilà donc le vin :

Chassagne-Montrachet - Les Blanchots Dessus 2007

Lez nez est boisé, trop, un peu de citron en arrière plan mais guère plus.

En bouche, c'est la banane et la vanille bourbon qui dominant. On sent tout de même de la tension dans ce vin et une finale intéressante.

Mais il n'est pas dans sa meilleure phase.

Bien -, à attendre impérativement ...

- Liens entre ontologie experte d'un domaine et « terminologie consos » pour recommander à partir des mots du consommateur
- Extraction et catégorisation du lexique de la dégustation avec triple dimension :
 - Sensoriel
 - Evaluative
 - Hédonique

II. TouristYou un « E-Greeters » dans votre poche





Au commencement...

Contexte :

- **Projet européen sur les discours des guides conférenciers :**
KA2 « Tell Me A Story : Multilingual Interactive Learning for Tourist Guides »
- **Thèse d'Olivier Méric (sous la direction de Laurent Gautier) :**
Organisation discursive de la visite médiée de sites touristiques : théorisation contributionnelle et valorisation d'une praxis professionnelle, Dijon, 2016
- **Contrat de collaboration avec la communauté de communes de Gevrey-Chambertin sur les mémoires du vin : enregistrement oraux des habitants**



Au commencement était l'anecdote ... et la rupture avec le discours touristique institutionnel...

TouristYou, c'est une application **touristique** d'un **nouveau genre** basée sur un **corpus** oral d'**anecdotes** authentiques et originales visant à mettre en valeur le **patrimoine** d'une **zone géographique** ou lié à un **thème**



Entre **expérientiel** et **historicité**, les anecdotes révèlent les **secrets** de l'histoire, des paysages, des traditions : les **partager** est la mission de **TouristYou**



Approche interdisciplinaire

- TouristYou** où l'histoire d'un agrégat de compétences interdisciplinaires :
- les enregistrements audio/vidéo sont **collectés** et **préparés** pour révéler l'aspect **historique** de l'anecdote
 - une analyse **linguistique** et un traitement **sémantique** viennent affiner ce **découpage** pour extraire, par annotations fines, le meilleur de l'anecdote et alimenter
 - une **application** mobile associant **cartographie**, **base de données** et **intelligence artificielle**
 - pour que l'expérience touristique du **greeter** devienne l'expérience de l'**utilisateur** : « *tourisme expérientiel* »



Indexation et annotation sémantique

Deux principes :

- l'information venant à l'utilisateur, il n'a pas besoin d'aller la chercher
 - => « **sérendipité contrôlée** » par notifications push (notamment après l'échec des circuits par QR codes)
 - l'information est **profilée** suivant ses centres d'intérêts
- ⇒ Nécessité d'une **annotation sémantique fine** des anecdotes par **balisage** (« *tags* ») des segments transcrits associés au flux audio / vidéo et au lieu
- Double couche sémantique :
 - Sémantique « naturelle », Sémantique « informatique »

Taxonomie (extraits) des descripteurs de Gevrey-Chambertin

LIEUX / SITES DIVERS

- > Abattoirs (Dijon)
- > Carrière de Brochon
- > Creux Tombain (le) (Ternant)
- > Falaises de Brochon
- > Falaises de Fixin
- > Ferme de Charmoy (ruines - Clémencey)
- > Ferme de Collonges (Fleurey-sur-Ouche)
- > Ferme du Leuzeu (Fleurey-sur-Ouche)
- > Fromagerie Gaugry (Brochon)
- > Goulotte (la) (réservoir d'eau urbain - Gevrey-Chambertin)
- > Légende de la Dame Blanche (Messigny-et-Vantoux)
- > Légende de la Pierre qui vire (Fleurey-sur-Ouche)

MONUMENTS / LIEUX ET SITES REMARQUABLES

- > Abbaye de Cluny (département Saône-et-Loire)
- > Bareuzai (le) (Dijon - monument)
- > Château de Brochon (Château de Stéphen Liégeard)
- > Château de Gevrey-Chambertin
- > Château de Montculot (Urcy)
- > Clos de Vougeot
- > Creux Tombain (le) (Ternant)
- > Falaises de Brochon
- > Falaises de Fixin
- > Fontaine de Jouvence (la) (Messigny-et-Vantoux)
- > Maison de Noisot (Fixin)

Démonstration mode déambulation

Lunetterie Henri Jullien

Lunetterie Henri Jullien. Vous trouverez plus d'informations sur l'histoire de la lunetterie à Morez au Musée de la Lunette.

Lunetterie, Industrie locale, Lunettes, Savoir-faire, Artisanat-Industriel, Fabrication artisanale, Fabrication industrielle, Qualité, Morez

Flux vidéos

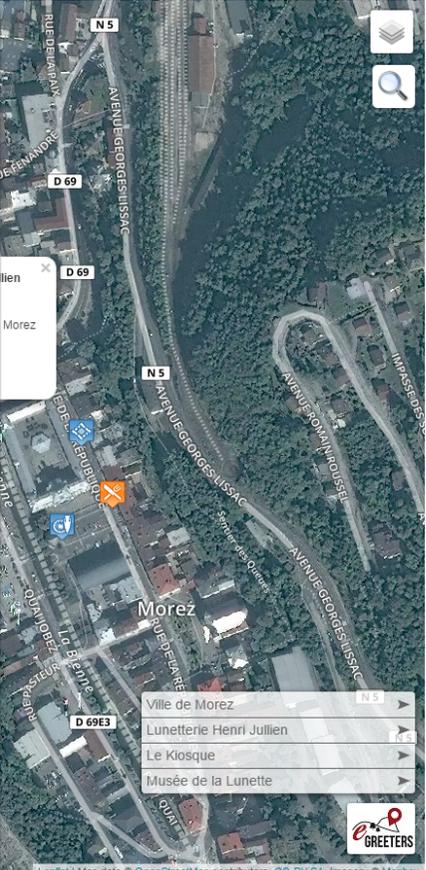
- [Henri Jullien : les Lunettiers artisans-industriels](#)

Henri Jullien : les lunettiers artisans-industriels



On est ici dans une usine, donc un bâtiment industriel, ça veut dire qu'on travaille des séries industrielles. Même si ce sont de petites séries, ce sont quand même des séries industrielles. En revanche, ces séries industrielles sont souvent travaillées de façon artisanale, c'est ce qui fait aussi la qualité qui est reconnue à nos produits c'est que on est des artisans-industriels ou des industriels-artisans. On a l'impression qu'une lunette quand on voit c'est fait plat. En réalité, c'est pas plat une lunette, c'est galbé dans tous les sens. C'est galbé dans le sens, dans ce sens-là, c'est galbé dans le sens-là. Chaque verre est galbé également dans les deux sens donc c'est vraiment quelque chose d'excessivement compliqué. Je fais le soudage de face, donc je positionne mes cercles

Film documentaire de Laurent Lutaud "Les lunettes font de la résistance" - Copyright : FTV/NomadesTV



Morez

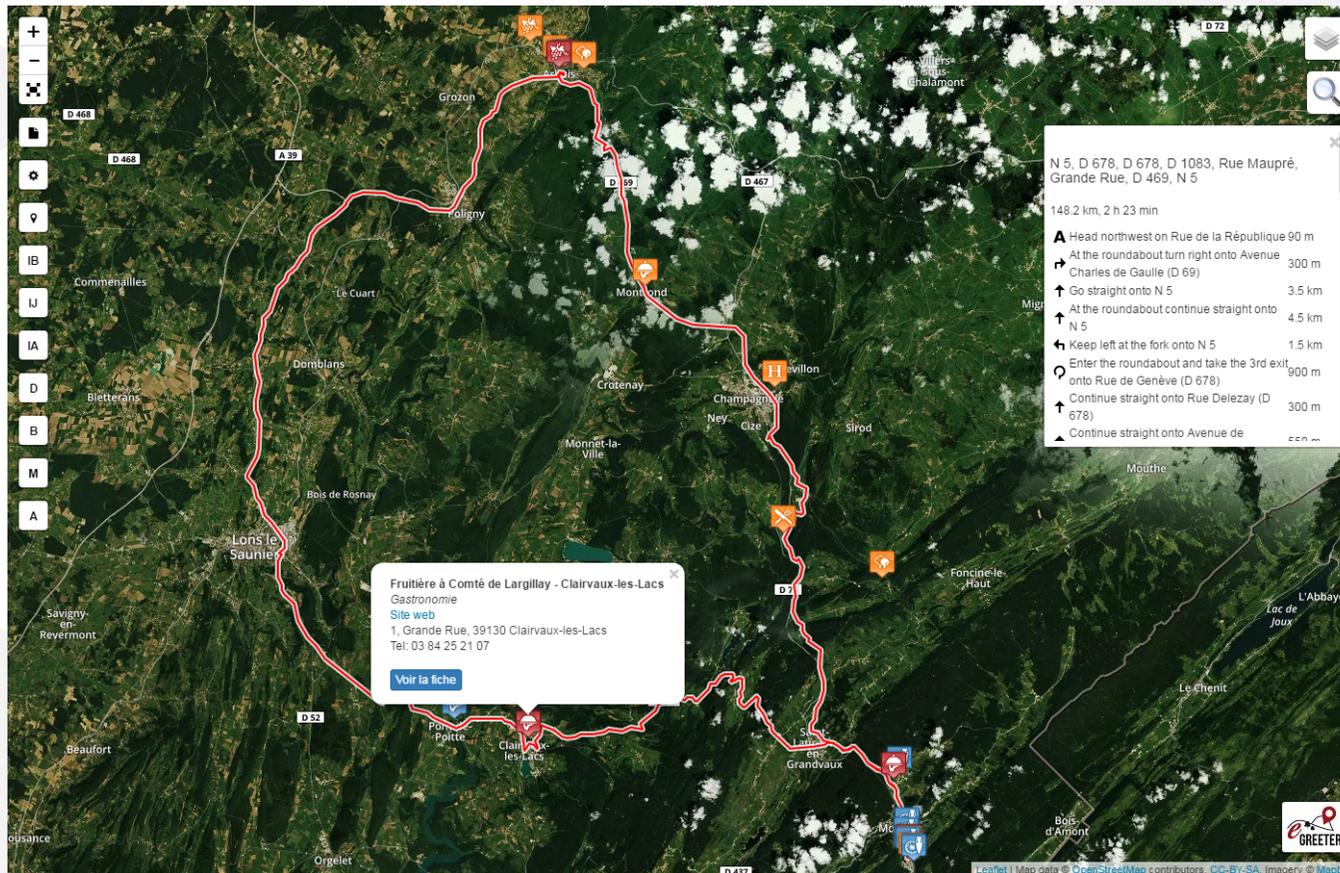
- Ville de Morez
- Lunetterie Henri Jullien
- Le Kiosque
- Musée de la Lunette

Les anecdotes arrivent au touriste en mode push en fonction de ses déplacements, sous forme de propositions de visites personnalisées.

Présentation du nom du lieu, de sa description, de ses « tags », des flux multimédias associés.

Pour chaque flux :
visualisation et transcription

Démonstration mode itinéraire



Le touriste paramètre le type d'anecdotes qu'il souhaite obtenir pour les intégrer à son plan de route, sa randonnée etc. ;

Les découvertes de points d'intérêts sont adaptées au profil utilisateur : affichage, accroche, anecdote.

Exemples :

- Passionné de gastronomie
- En famille
- Passionné par le patrimoine



Plus-values de TouristYou

Acteurs du tourisme

- Mise à disposition d'une offre **sur-mesure** : **déclinaisons zonale** et **thématique**
- Une offre **authentique** et **originale**, loin des habituels discours institutionnels stéréotypés/formatés
- Mise en lumière de tous les patrimoines à travers l'**expérience utilisateur**
- Partenariats avec les commerces/services environnant pour favoriser l'**économie de proximité** => **veille possible avec les retour, détection des tendances, ...**

Visiteurs

- Une application qui s'adapte à son **utilisateur** (déambulation/itinéraire + sélection des points d'intérêts)
- Effacement des contraintes et **liberté** de circulation dans la zone visitée
- Ressenti expérientiel « comme les habitants »
- **Sérendipité profilée et contrôlée !**

III. AdWine

Systeme de recommandation intelligent pour le vin
et la gastronomie





Présentation

« Au service d'une communication ciblée non seulement **pour** les consommateurs, mais **partant** des consommateurs et utilisant les canaux numériques. »

Lauréat AAP CVT Athéna

- **Adapter** le langage et la terminologie des experts et des professionnels à celles des consommateurs et non l'inverse.
- Sous forme d'**ontologies** (au sens informatique): ensemble structuré des termes et concepts représentant le sens d'un champ d'informations (ici l'œnologie).

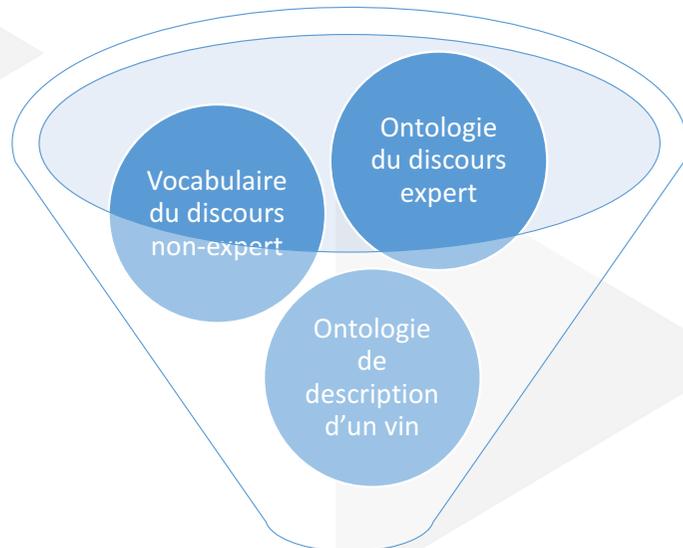
Outil de recommandation intelligent

- **croisement** ontologies professionnelles et vocabulaire **produit** par les consommateurs par vin
- profilage classique + lexical
- extension vers **gastronomie** et accord mets-vins

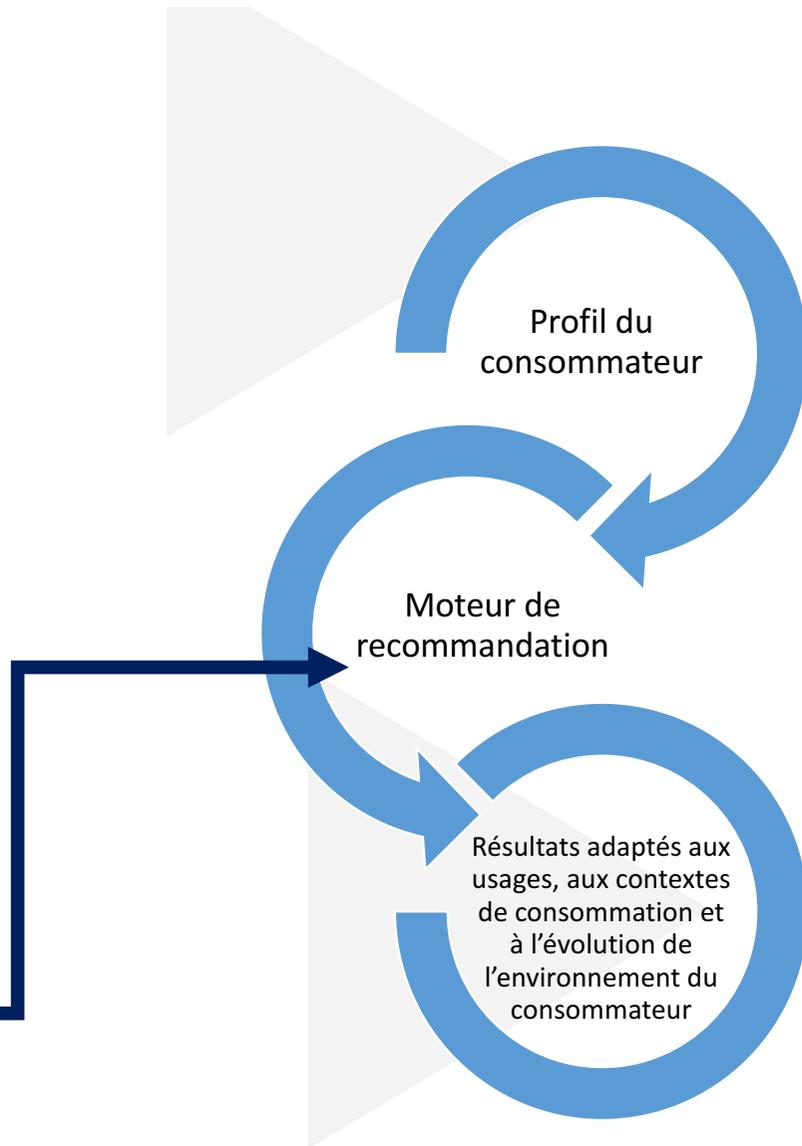
=> **Recommandation d'un vin à partir du profil spécifique de l'utilisateur consommateur**



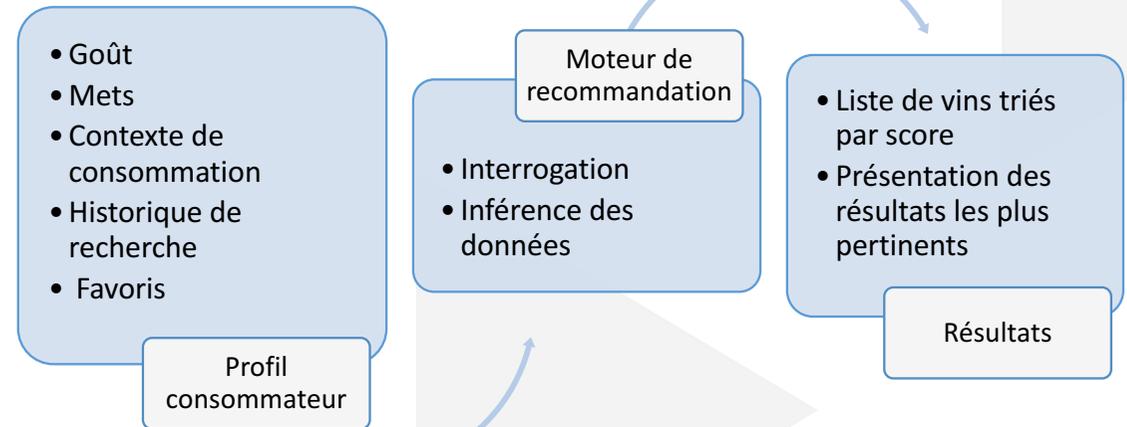
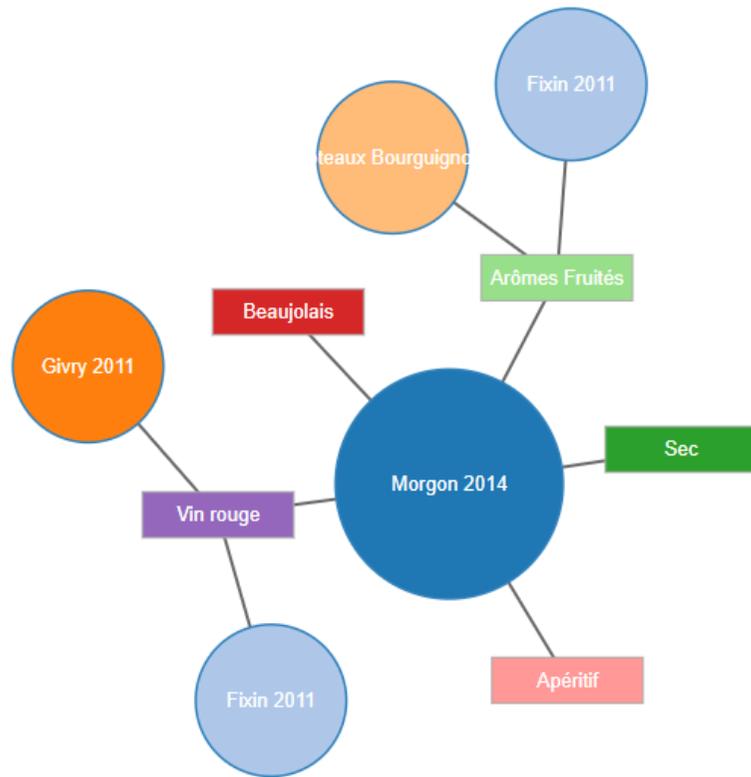
Fonctionnement



Fiche d'un vin



Processus de recommandation





Démonstration du prototype

1 Profil du vin 2 Mets 3 Résultat

Région viticole: Champagne

Type du vin: blanc effervescent, rosé effervescent

Gamme de prix: Toutes gammes, Entrée de gamme, Milieu de gamme, Haut de gamme, Prestige

Circonstances de repas: Apéritif, En couple, Entre amis, Célébrations

→

Démonstration du prototype

Que souhaitez-vous déguster avec ce vin ?

Catégorie de repas Choisir les mets

Langoustine

-  **Langoustine**
Fruits de mer - Crustacé
Entrée de poisson - Entrée de crustacé
-  **Langoustine**
Apéro de fruits de mer - Apéro de crustacé
Apéritif de poisson
-  **Langoustine**
Plateau de fruits de mer
Plat unique
-  **Langoustine**
Plateau de fruits de mer
Plat de poisson
-  **Langoustines au four**
Coquillage et crustacé grillés - Coquillage et crustacé poêlés - Mollusque grillé
Plat de poisson

 **Rechercher** 

 **Langoustine**
Fruits de mer - Crustacé
Entrée de poisson - Entrée de crustacé

Démonstration du prototype

Recommandation

Graphique



Champagne grand cru Blanc de Blancs



Domaine A.R. Lenoble

- Région :** Champagne
- Appellation :** Champagne
- Dénomination :** Champagne grand cru
- Cépage :** Chardonnay blanc
- Circonstances :** Apéritif
- Gamme de prix :** Haut de gamme

Arômes

- Arômes floraux
- Arômes fruités
- Arôme de beurre frais
- Arôme de brioche

Accords mets/vins

- Soufflé au fromage
- Queues d'écrevisses
- Crustacé en sauce
- Poisson grillé
- Langoustine

Vins similaires

Arômes Fruités

Givry 2011



Bourgogne

Givry



Morgon 2014



Beaujolais

Morgon



Fixin premier cru Hervelets 2011



Bourgogne



Plus-values d'AdWine

Filière viti-vinicole, grande distribution, restauration

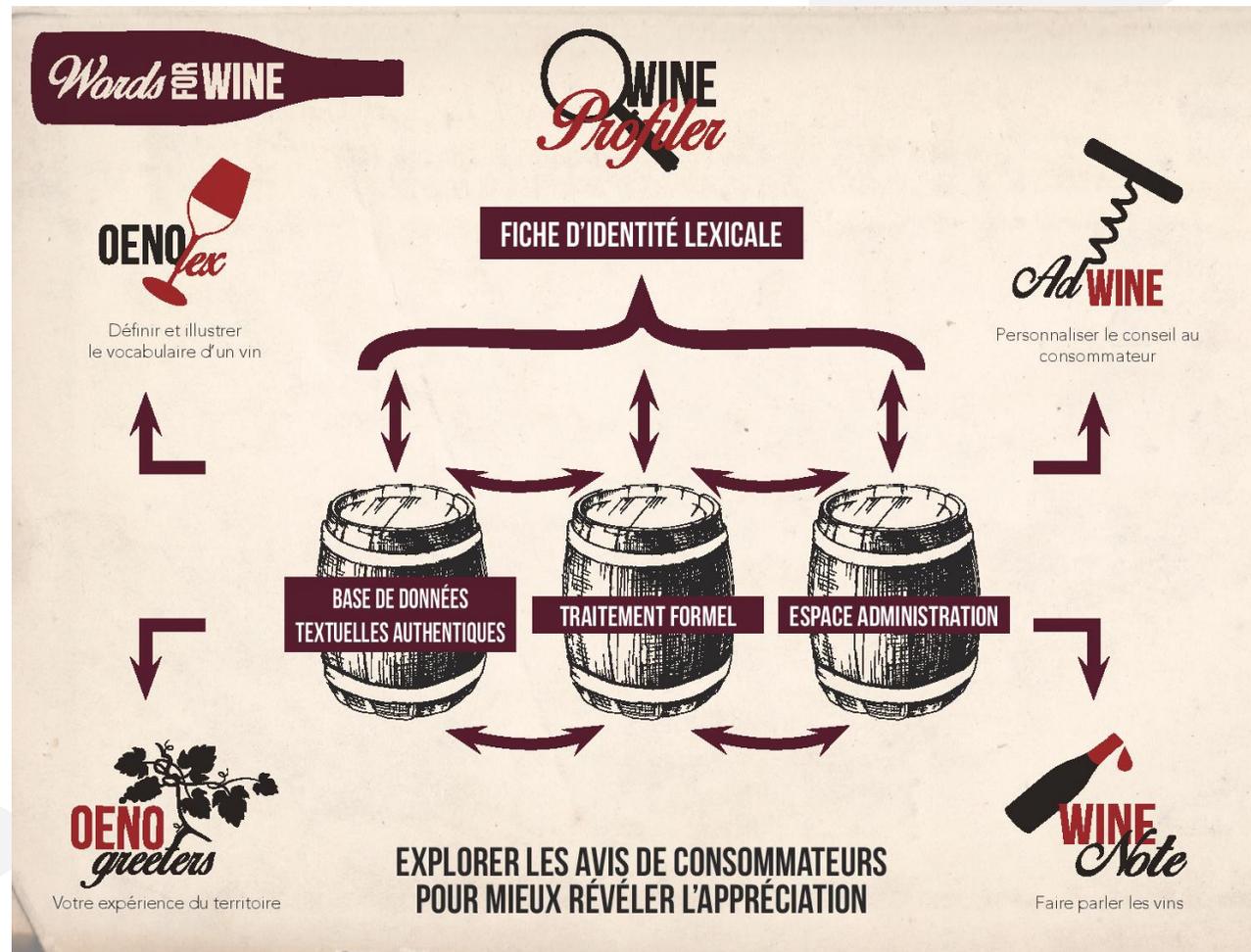
- Constitution d'un **dispositif numérique** associé aux études en sciences humaines sur la vigne, le vin et l'oenotourisme ;
- Inscription dans la tendance actuelle d'**associer** les nouvelles technologies à la communication autour des vins ;
- Répondre aux **besoins de communication pédagogique** sur le vin et d'éducation au goût des consommateurs.

Words4Wine

- Ensemble de fonctionnalités multi-agents indexés sur une base de données communes ;
- Double visée : **professionnels** de la filière / **utilisateurs finaux**.

Words4Wine

Quand le numérique révolutionne la communication autour du vin



4. Perspectives

« Le fait d'être appliquée donne à la linguistique une dimension pluridisciplinaire. » (Williams 2009 : 205-206)

- Interactions entre linguistique, informatique, sciences cognitives, (intelligence artificielle) ;
- Compilation de nouveaux types de données pouvant servir directement des problématiques commerciales et marketing orientées « client » ;
- Développement d'une recherche appliquée aux besoins industriels avec, en retour, des développements méthodologiques et de la capitalisation de connaissances en recherche fondamentale.

=> Perspective gagnant-gagnant !

Merci !

Laurent GAUTIER, PR, Centre Interlangues Texte Image Langage (UBFC, EA 4182)

laurent.gautier@ubfc.fr

Cyril NGUYEN VAN, IE CNRS, Maison des Sciences de l'Homme Dijon (USR CNRS-uB 3516)

cyril.nguyen-van@ubfc.fr

Matthieu BACH, Doctorant, Centre Interlangues Texte Image Langage (UBFC, EA 4182)

matthieu_bach@etu.u-bourgogne.fr

