



LABORATOIRE D'INFORMATIQUE FONDAMENTALE ET APPLIQUÉE DE TOURS



# BDTLN

## BASES DE DONNÉES ET TRAITEMENT DES LANGUES NATURELLES

**Bilan 2016-2021**  
**Projet 2024-2028**

# Bilan 2016-2021

## Structuration et positionnement scientifique

# Structuration de l'équipe

## □ Actuellement

- Responsable : B. Markhoff
- 4 PR ; 11 MCF
  - idem en 2016 (+1 HDR)
- 21 doctorants ; 6 postdocs ou ATER ou ingénieur.es
- **Parité femme-homme sur 2016-2021**
- 27% de doctorantes
- 26% de permanentes MCF+PR
- 0% de PR

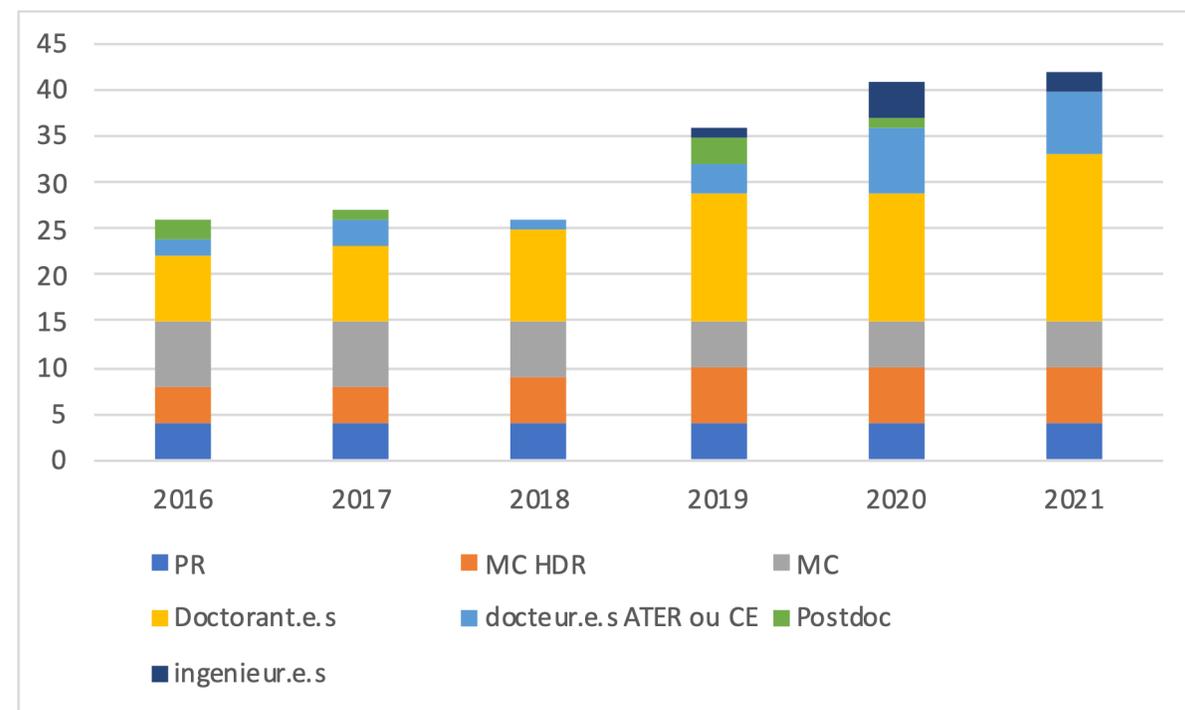
## □ Ratio PR/(PR+MCF)

- 0.26 au 1/1/23 (stable depuis 2016)

## □ Réunions mensuelles, séminaire annuel

## □ Nombreuses charges collectives

- Présidence de l'université, co-direction ICVL, co-direction LIFAT, responsable de département d'enseignement, etc.



# Positionnement scientifique : pour une IA intelligente

## □ Exploration et analyse interactive de données

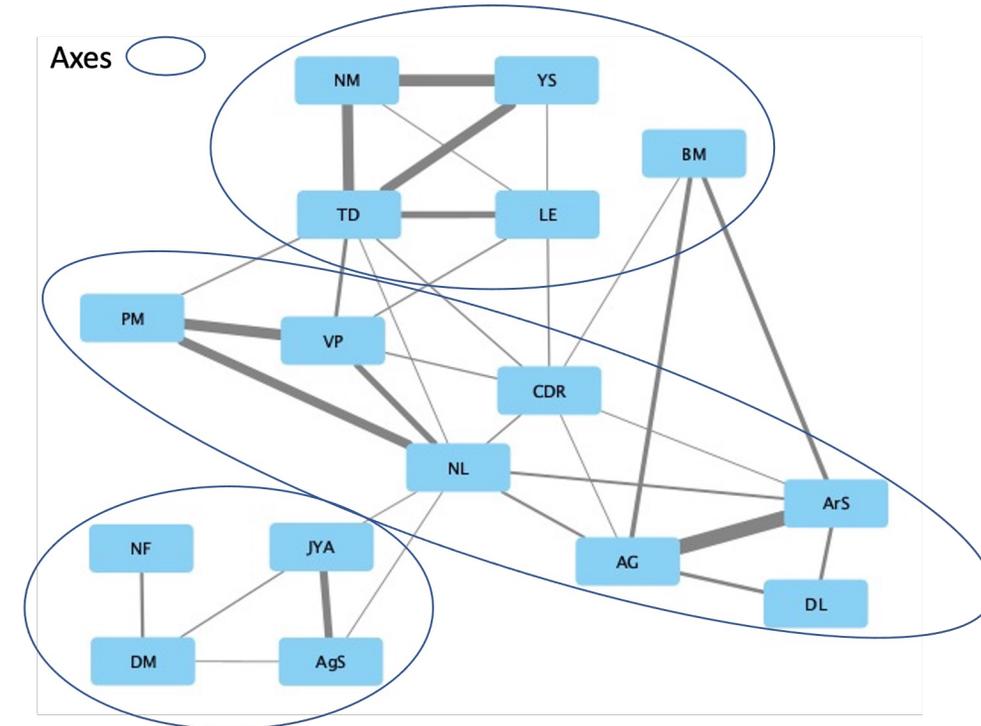
- Développement d'outils **centrés utilisateurs**
- Définition de **langages** et **algorithmes** pour l'extraction de **motifs** et l'**exploration** de données
- **Modélisation** et **découverte** des attentes et des **préférences** utilisateurs

## □ Intelligence des données et des services

- Point de vue **sémantique**
- Intégration de données et connaissances de **sources hétérogènes**, **composition de services** à partir de descriptions sémantiques
- Conception de systèmes pour exploiter des **données et services complexes**, en particulier des **données spatio-temporelles**

## □ Traitement automatique des langues et interaction

- **TA conscient du langage**
- **Ressources libres** (données systèmes)
- **Linguistique computationnelle** (comprendre la langue) ou applications en **situations réelles d'usage** : entités polylexicales, coréférence



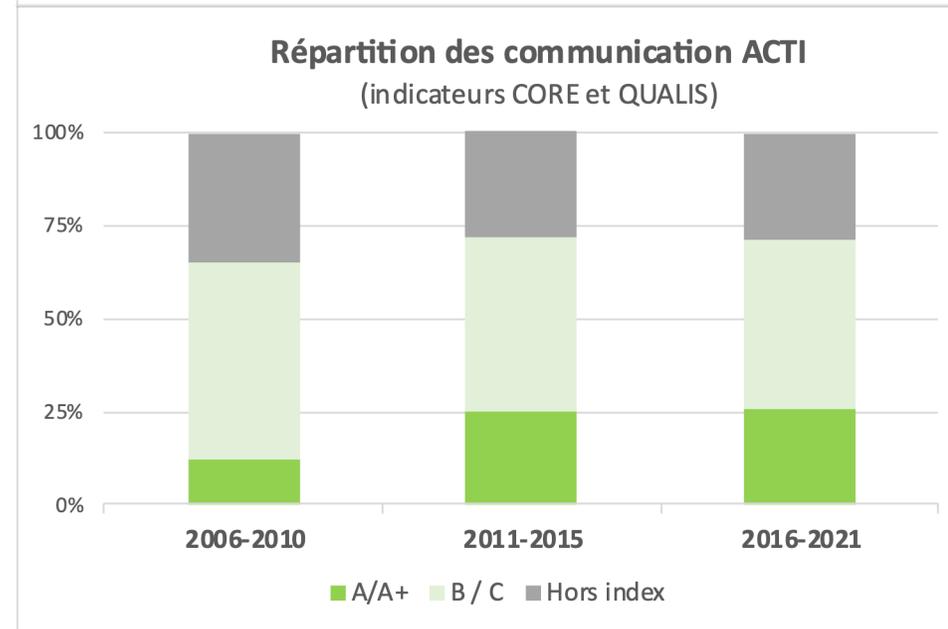
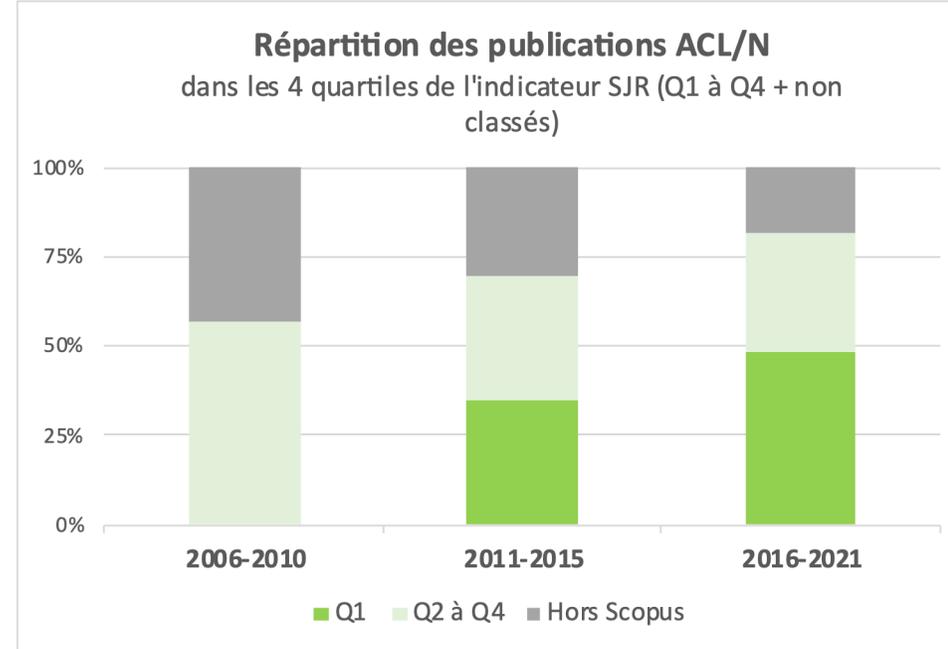
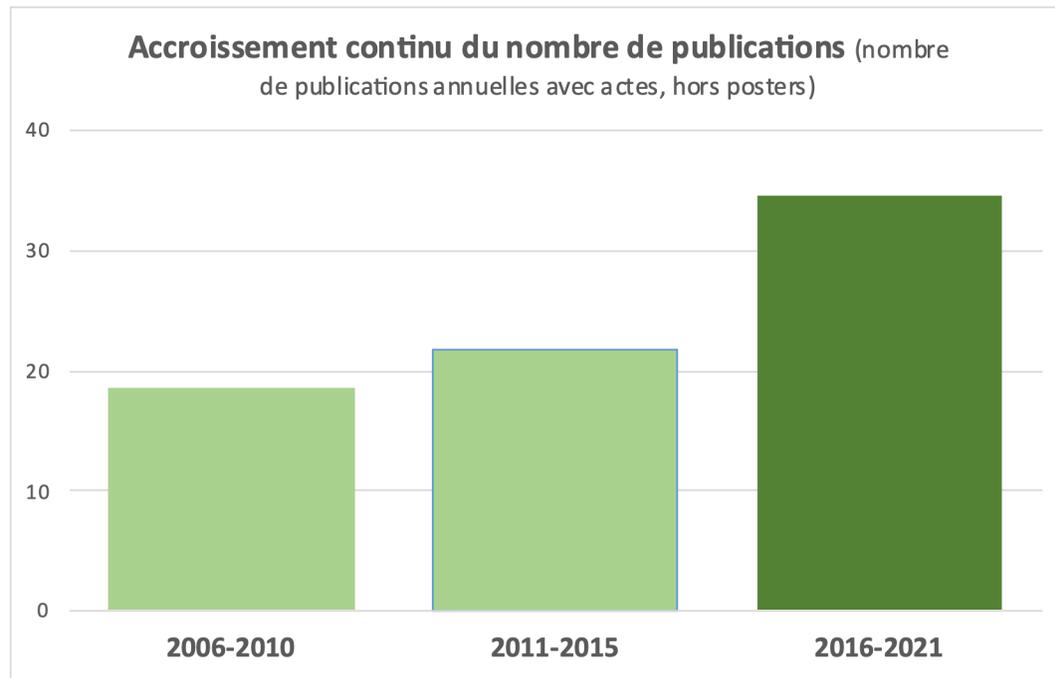
# Bilan 2016-2021

## Bilan scientifique

# Production scientifique

## Amélioration quantitative et qualitative

- à effectif permanent constant



# Rayonnement

Visibilité internationale affirmée

## □ Thèses

- 10 thèses en co-tutelle ou co-direction avec des universités étrangères
- 4 en co-direction avec Orléans

## □ Publications

- Co-auteur d'une université étrangère
- 47,6% (ACL) 32% (ACTI)

## □ Projets

- 1315 k€ sur la période (régionaux, nationaux, internationaux, CIFRE)
- Coordination du réseau COST PARSEME
- Participation à deux projets H2020 interdisciplinaires

# Encadrements doctoraux

- Visibilité auprès des acteurs économiques
- Structuration des activités du LIFAT
- Appui sur les formations locales

## □ 17 thèses soutenues

- 361 k€ obtenus entre 2016-2021 sur contrat CIFRE

## □ 21 thèses en cours

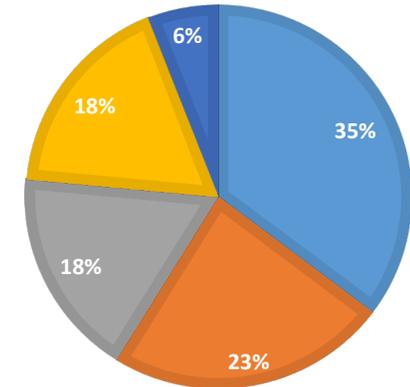
- 4 avec RFAI ou ROOT
- 8 co-tutelles ou co-directions avec d'autres universités/UR

## □ Master

- Erasmus Mundus (arrêté en 2018)
- 12 doctorants issus du Master BDMA (fac. S&T)
- Actions de sensibilisation locale : Apéros Info

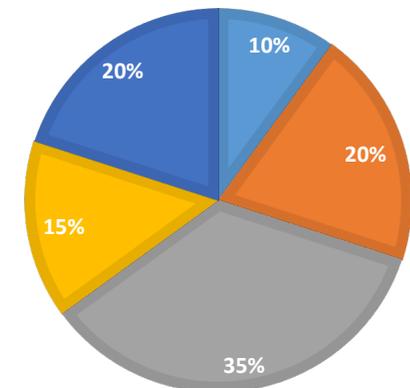
FINANCEMENTS THÈSE 2016-2021

■ Etranger ■ Ecole doctorale ■ CIFRE ■ Projet ■ Autre



FINANCEMENTS THÈSE 2022

■ Etranger ■ Ecole doctorale ■ CIFRE ■ Projet ■ Autre



# Contributions remarquables

## □ Fouille de données du Web sémantique

- Qualification des graphes de connaissances
  - estimation de la représentativité d'un graphe
  - découverte automatique de cardinalités maximales
- Analyse des graphes de connaissances
  - requêtes analytiques
  - génération de tableaux comparatifs

## □ Détection d'éléments textuels d'intérêt

- Détection des entités polylexicales
  - Travaux fondamentaux (diversité linguistique) et applicatifs : dimension plurilingue
  - Coordination d'un réseau européen COST
    - 29 langues, 6 dialectes

EGC 2021  ISWC 2018  ESWC 2021 	ANR SESAMES
Séminaire Dagstuhl Workshop MWE 2018 EACL 2017	COST Parseme ANR Parseme_FR ANR SELEXINI

# SWOT

# SWOT - interne

FORCES	FAIBLESSES
<p><b>Production scientifique</b> - Capacité à maintenir ou améliorer une production de qualité</p>	<p>Difficultés à lâcher des <b>responsabilités</b> (ces responsabilités garantissent plus d'aisance dans l'environnement d'exercice de nos fonctions)</p>
<p><b>Encadrement</b> - Capacité à encadrer de nombreuses thèses de qualité, (attesté par les publications et l'avenir des docteur.e.s)</p>	<p><b>Encadrement</b> - Plusieurs démarrages de thèses suivis de prise de poste dans le privé</p>
<p><b>Collaborations</b> - Multiplicité de projets financés, bon ancrage régional, nombreuses collaborations nationales, capacité à développer des recherches avec d'autres disciplines</p> <p><b>Visibilité internationale</b> - Capacité à développer et maintenir des réseaux internationaux</p>	
<p><b>Structuration</b> - Dynamisme, évolution des interactions intra et extra axes thématiques de l'équipe, augmentation des interactions avec les autres équipes du LIFAT</p>	<p><b>Structuration</b> par axe prévue remise en cause par les dynamiques d'interaction réelles au cours du contrat (sauf axe TAL)</p>
<p><b>Financement de la recherche</b> - Diversification de nos sources de financement, en particulier avec des entreprises via des financements de thèses CIFRE</p>	<p><b>Ratio PR/MCF très faible</b>, aggravé par le changement de fonction de A. Giacometti (président actuel de l'université) et la perte future d'un support de PU (retraite D. Maurel). <i>Risque de démobilisation des MC HDR par manque de perspective</i></p>
<p><b>Attractivité</b> - Capacité à attirer vers la recherche les membres du master local</p>	<p><b>Attractivité</b> - Difficulté à attirer de bonnes candidatures de l'extérieur sur les thèses à financements académiques</p>

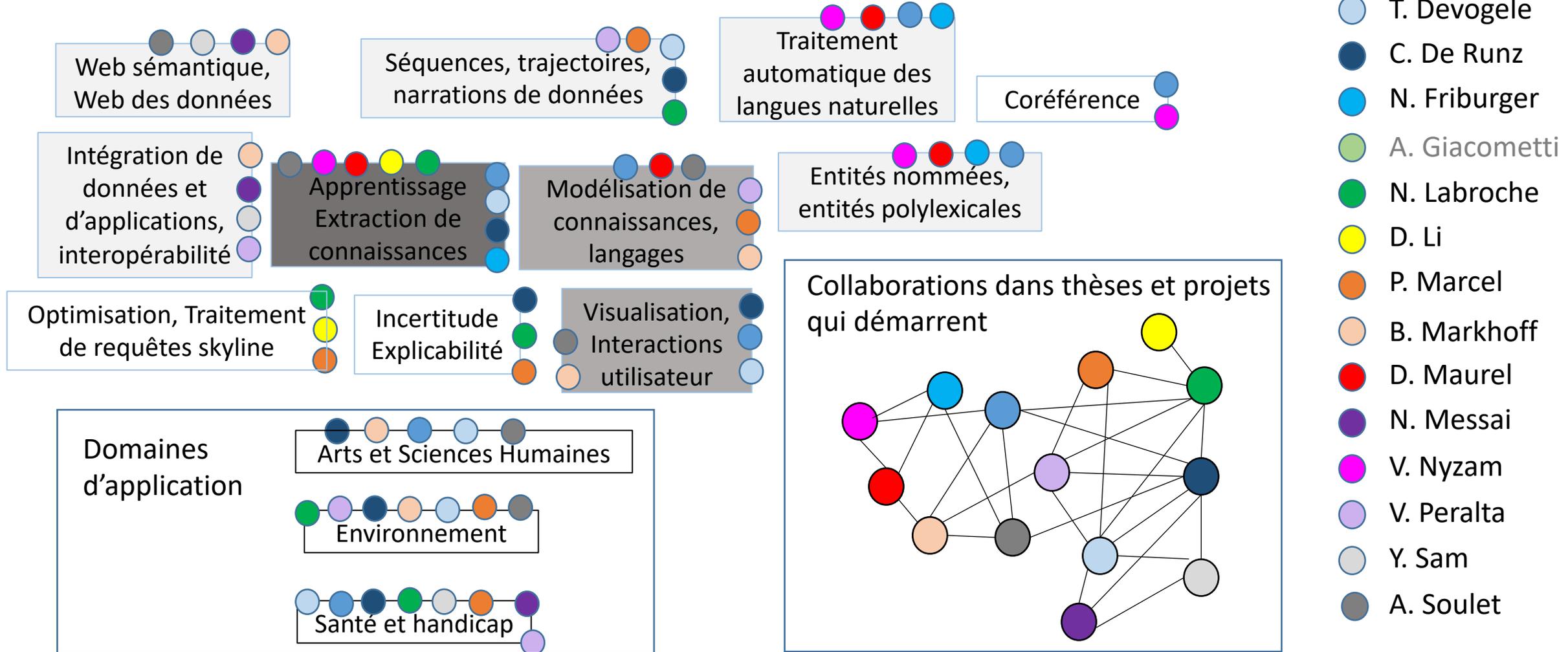
# SWOT - externe

OPPORTUNITES	MENACES
Demande importante des autorités européennes et nationales dans le domaine de la science ouverte et des principes FAIR	<b>Marché de l'emploi très concurrentiel</b> : difficulté à recruter des doctorants, postdocs et ingénieurs, des doctorants qui s'engagent puis décident de prendre un poste dans le privé
Nombreuses sollicitations des autres disciplines académiques pour collaborer sur leurs traitements de leurs données	<b>Concurrence entre disciplines</b> : il est nécessaire de prendre des responsabilités à tous les niveaux pour ne pas être "oubliés" dans la répartition des ressources (postes)
Synergies potentielles entre les thématiques de l'équipe	<b>Perte de deux chercheur.e.s clés</b> : Arnaud Giacometti (devenu président d'université) pour l'axe Exploration et Agata Savary (devenue PU à Saclay) pour l'axe TAL, sans perspective de remplacement dans le cas du premier

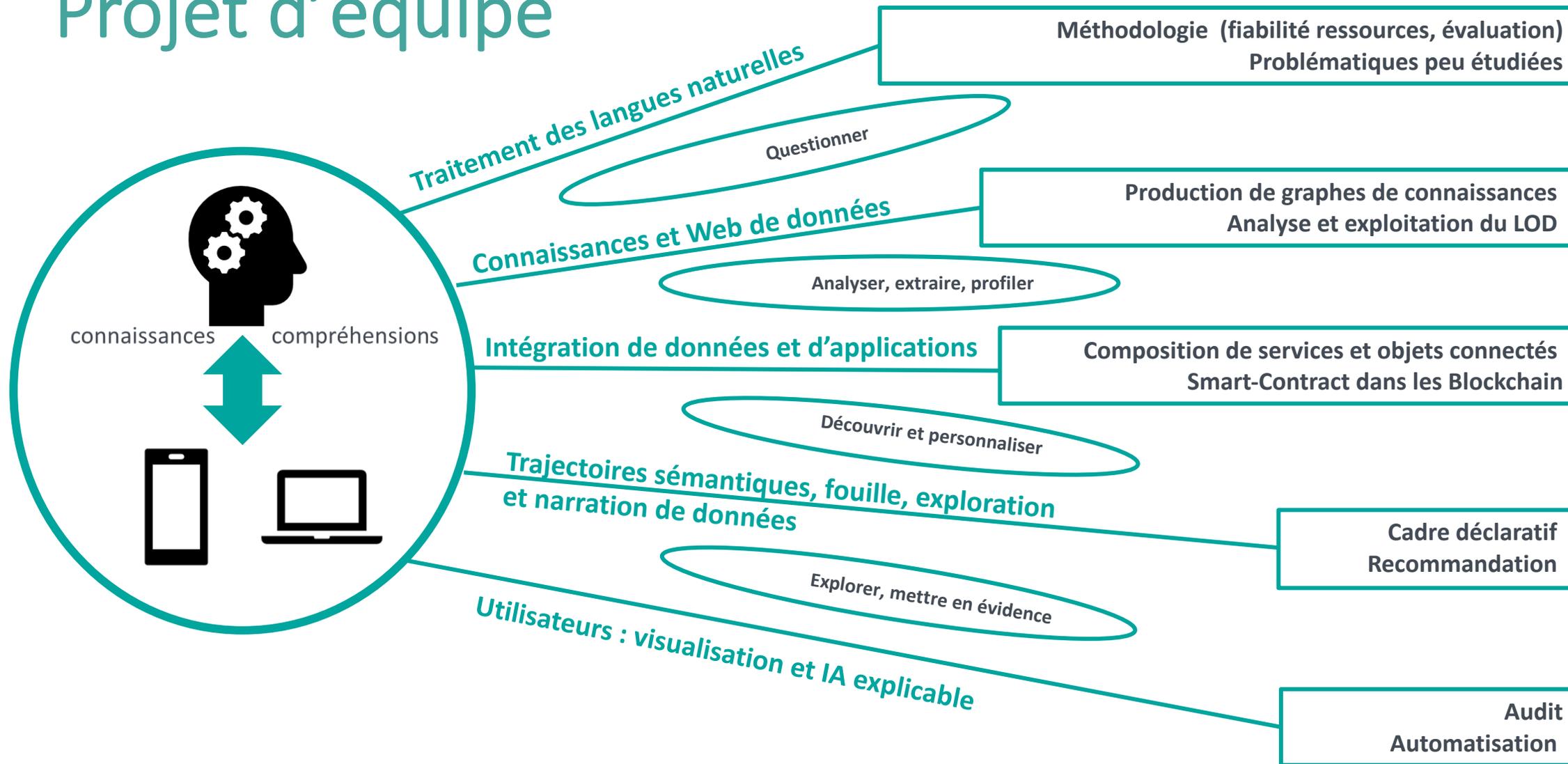
# Projet 2024-2028

## Projet d'équipe

# Ce qui nous rassemble : diversité et complémentarité



# Projet d'équipe



# Actions de stratégie scientifique

## □ Stratégie de publications

- Continuer à travailler à l'augmentation de la qualité des publications
  - Q1, A/A\*, A1/A2 dans Qualis
- Réduire l'empreinte carbone (Labo 1.5)
  - Privilégier les conférences de très bons niveaux en Europe
  - Prioriser les doctorants sur les déplacements longs

## □ Accompagnement vers l'HDR

- 3 HDR en cours de rédaction, 1 quatrième envisagée dans les 2/3 ans
- Poursuivre l'encouragement aux CRCT / délégations

## □ Nécessité d'une amélioration du différentiel PR/MCF

- Obtention du rehaussement d'un poste à l'IUT de Blois (2023)
- Demande du rehaussement de postes à la Faculté S&T
- Demandes de promotions internes (2023 et 2024)