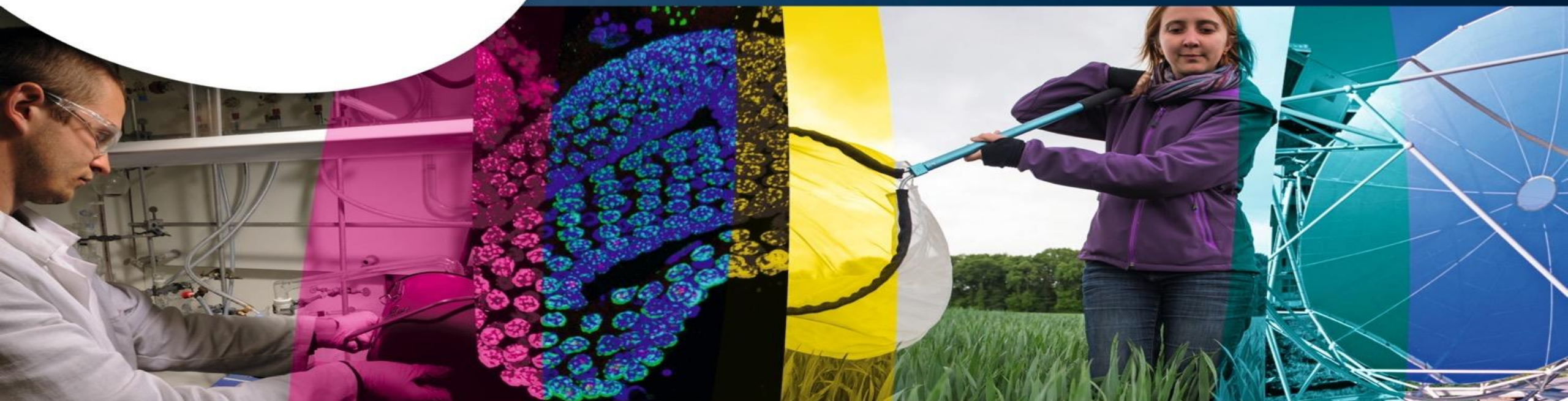




www.cnrs.fr

Vers des infrastructures de services avancés d'analyse de textes et de données

Défis et verrous



Enjeu majeur - trésor caché

Big data, small data, **NO DATA**, C Borgman

Deuxième axe : structurer et ouvrir les données de la recherche



© Camelia.boban - Own work, CC BY-SA 3.0,

ACCEDER - FAIR



Fouille des contenus

- Indexation, enrichissement de base
- Moteur de recherche (facettes)
- Extraction d'un ou des corpus contenant l'information recherchée
- Homogénéisation si nécessaire => **CORPUS**



80 % du temps du projet consacrés à la constitution du CORPUS !



www.cnrs.fr



Fouille des contenus

- Indexation, enrichissement de base
- Moteur de recherche (facettes)
- Extraction d'un ou des corpus contenant l'information recherchée
- Homogénéisation si nécessaire => **CORPUS**



Extraction de connaissance

- Enrichissements sémantiques, algorithmes apprenants
- Développements informatiques et développement d'ontologies, de thésaurus



www.cnrs.fr



Fouille des contenus

- Indexation, enrichissement de base
- Moteur de recherche (facettes)
- Extraction d'un ou des corpus contenant l'information recherchée
- Homogénéisation si nécessaire => **CORPUS**



Extraction de connaissance

- Enrichissements sémantiques, algorithmes apprenants
- Développements informatiques et développement d'ontologies, de thésaurus



Capitalisation

- Connaissance
 - Corpus enrichis mis à disposition
 - Algorithmes réutilisables
- => **Reproductibilité, mutualisation, innovations sociétales**



Comment ?



www.cnrs.fr



INFRASTRUCTURE PERENNE

SERVICES INCLUS

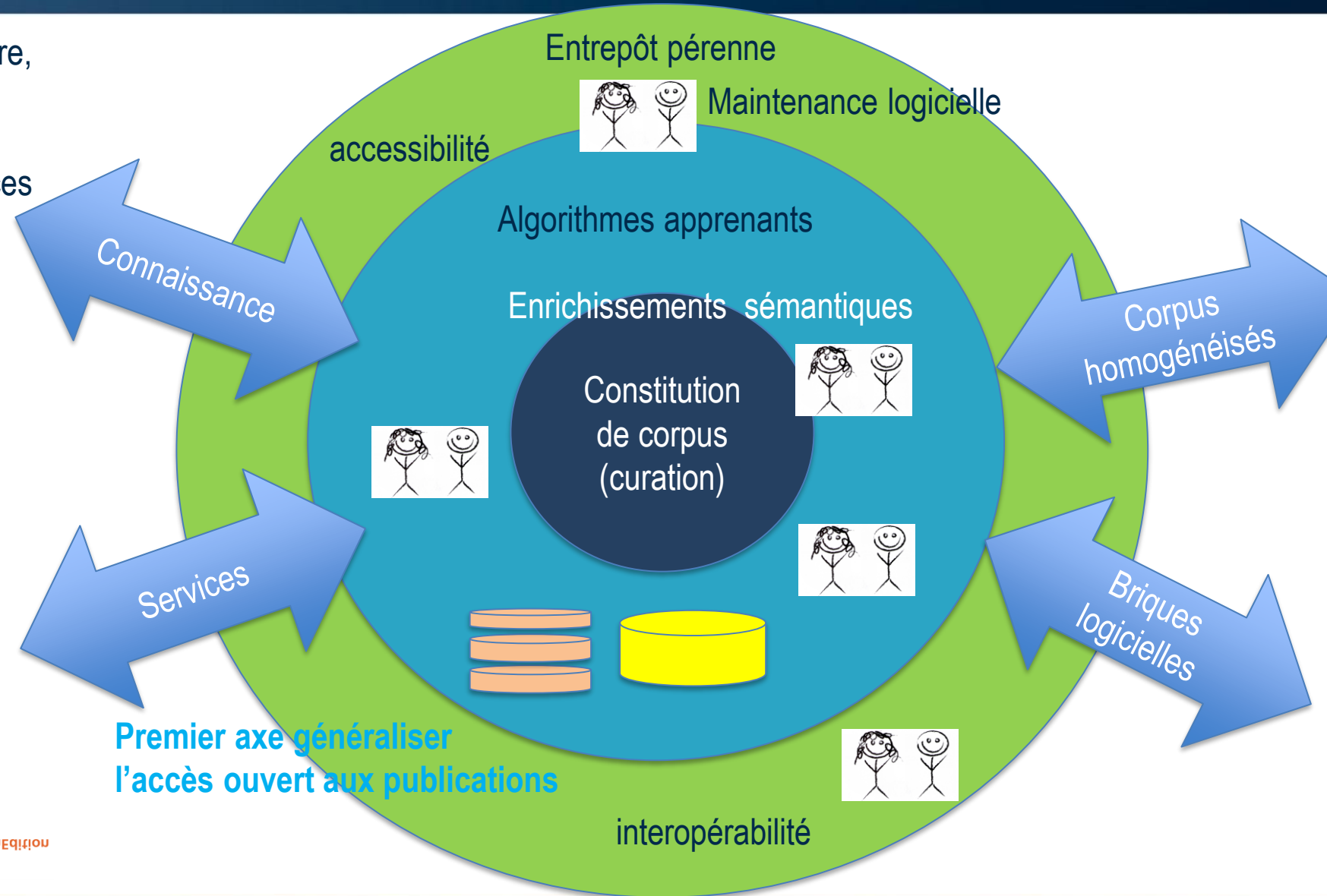
COMPETENCES transversales

INFRASTRUCTURE HUMAINE

GASTRONOMIE

INFRASTRUCTURE PERENNE

Environnement multi-disciplinaire,
Multi-métiers et compétences
Pour un innovation ouverte
Données-connaissances-services



DEFIS et VEROUS

If you want to go fast go alone,
if you want to go far **go together**



Transversalité : travailler ensemble

Informaticiens, data scientists, documentalistes
Ingénieurs systèmes, curateurs, scientifiques

Compétences :

Formations continues, Master, écoles doctorales

Troisième axe : s'inscrire dans une dynamique durable,
Maintenance et durabilité des infrastructures et des données

Accessibilité : E-infrastructure de services

Mise à disposition sous licences (à choisir)

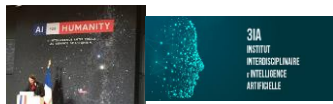


Collaborative reuse and new data creation

Normes Interopérabilité Partage Souveraineté



TEXTCROWD



ANR

Evaluation : Récompenses et motivations