

IA NEURO-SYMBOLIQUE POUR L'INTERPRÉTATION GRANULAIRE DES DONNÉES DES BASES STUPS© ET OTARIES© : DÉFIS APPLICATIFS ET PERSPECTIVES

S. Guillemin¹, L. Dujourdy², L. Journaux³, A. Roxin¹

¹Université de Bourgogne, Laboratoire d'Informatique de Bourgogne (LIB) EA 7534

²Institut Agro Dijon, Cellule d'Appui à la recherche en science des données

³Institut Agro Dijon, Laboratoire d'Informatique de Bourgogne (LIB) EA 7534

Contexte

Problématiques

Rappels
théoriques

IA Neuro-
symbolique

Travaux
réalisés

Conclusion et
travaux futurs

Thèse 100% financée par le Ministère de l'Enseignement Supérieur, de la Recherche et de l'Innovation (MESRI)

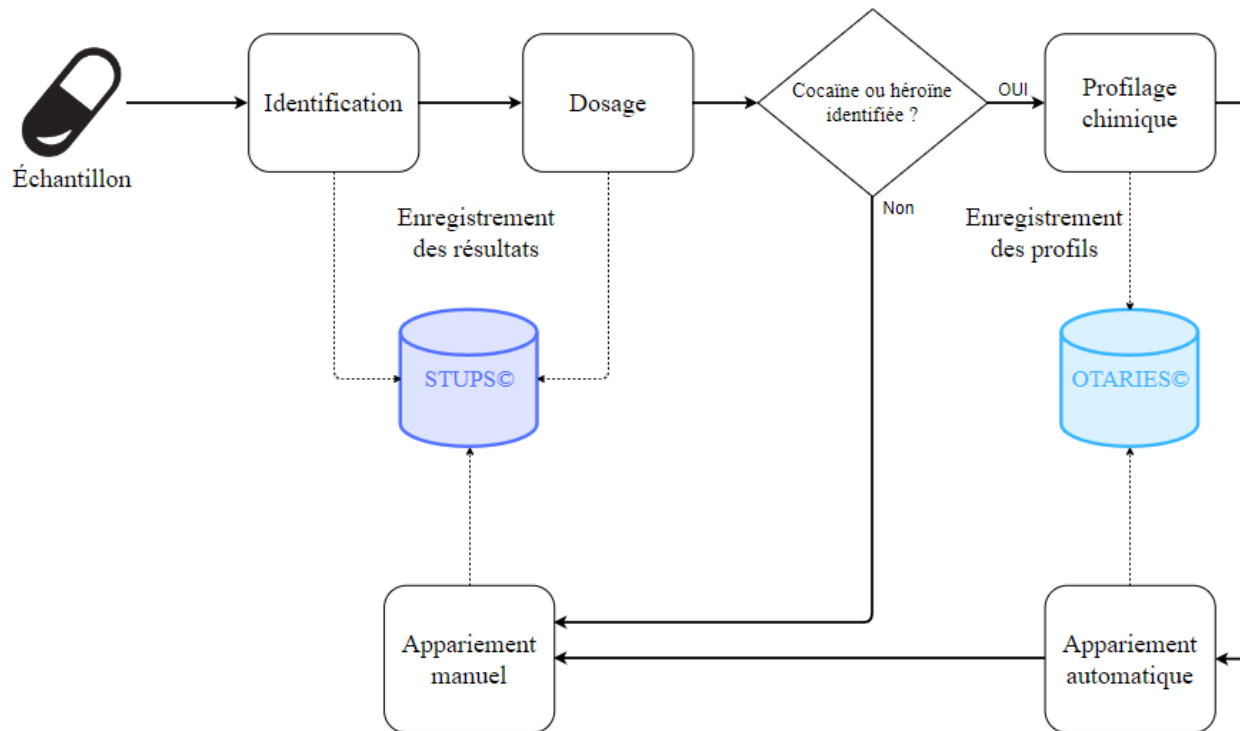
- Débutée en Octobre 2022

Collaboration avec le Ministère de l'Intérieur (MI)

- Plan national STUP français (2019) - lutte contre la drogue
- Identification des méthodes des trafiquants et anticiper leurs évolutions

Au travers du projet "Artificial Intelligence for Narcotic Profiling" (AI4NP)

- Regroupant: le LIB, l'Institut Agro Dijon et le Service National de Police Scientifique (SNPS)
- Support pour accès aux données confidentielles



Processus d'analyse d'un échantillon de substance supposée illicite

Limites du processus

- **Hétérogénéités** – base de données alimentées par 5 laboratoires de la Police Scientifique et par l'Institut de Recherche Criminelle
- **Inconsistantes dans le temps**
- **Peu d'informations** utilisées
- **Pas de connaissances métiers**
- **Analyse complexe** des données/informations contenues
- **Pas de travail proactif**

Problématiques métier

- **(PM1) Rapprocher automatiquement des profils** selon les caractéristiques des échantillons
- **(PM2) Établir des modèles prédictifs** pour caractériser les produits stupéfiants circulant en France
- **(PM3) Comparer des compositions de produits** par années, niveaux de trafic, territoires etc.

Problématiques métier

- **(PM1) Rapprocher automatiquement des profils** selon les caractéristiques des échantillons
- **(PM2) Établir des modèles prédictifs** pour caractériser les produits stupéfiants circulant en France
- **(PM3) Comparer des compositions de produits** par années, niveaux de trafic, territoires etc.

Problématiques scientifiques

- **(PS1) Spécifier des connaissances et règles métier par des approches d'IA symbolique** pour des résultats explicables
- **(PS2) Coupler des approches d'IA symbolique et d'IA connexionniste** pour la prédiction et l'aide à la décision
- **(PS3) Spécifier et gérer efficacement différentes échelles d'analyse - informatique granulaire dans les modèles d'IA connexionniste et symbolique**

IA Symbolique

- Raisonnement cognitif humain via un système de représentation des connaissances (des symboles) Flasiński et al., 2016 [18]
- Une ontologie : spécification explicite et formelle d'une conceptualisation partagée d'un domaine de connaissances Studer et al., 1998 [42]

IA Connexionniste

- Utilisation de neurones formels organisés en réseau Dastres et al., 2021 [7]
- Entraînement du réseau en modifiant la valeur de ses poids synaptiques

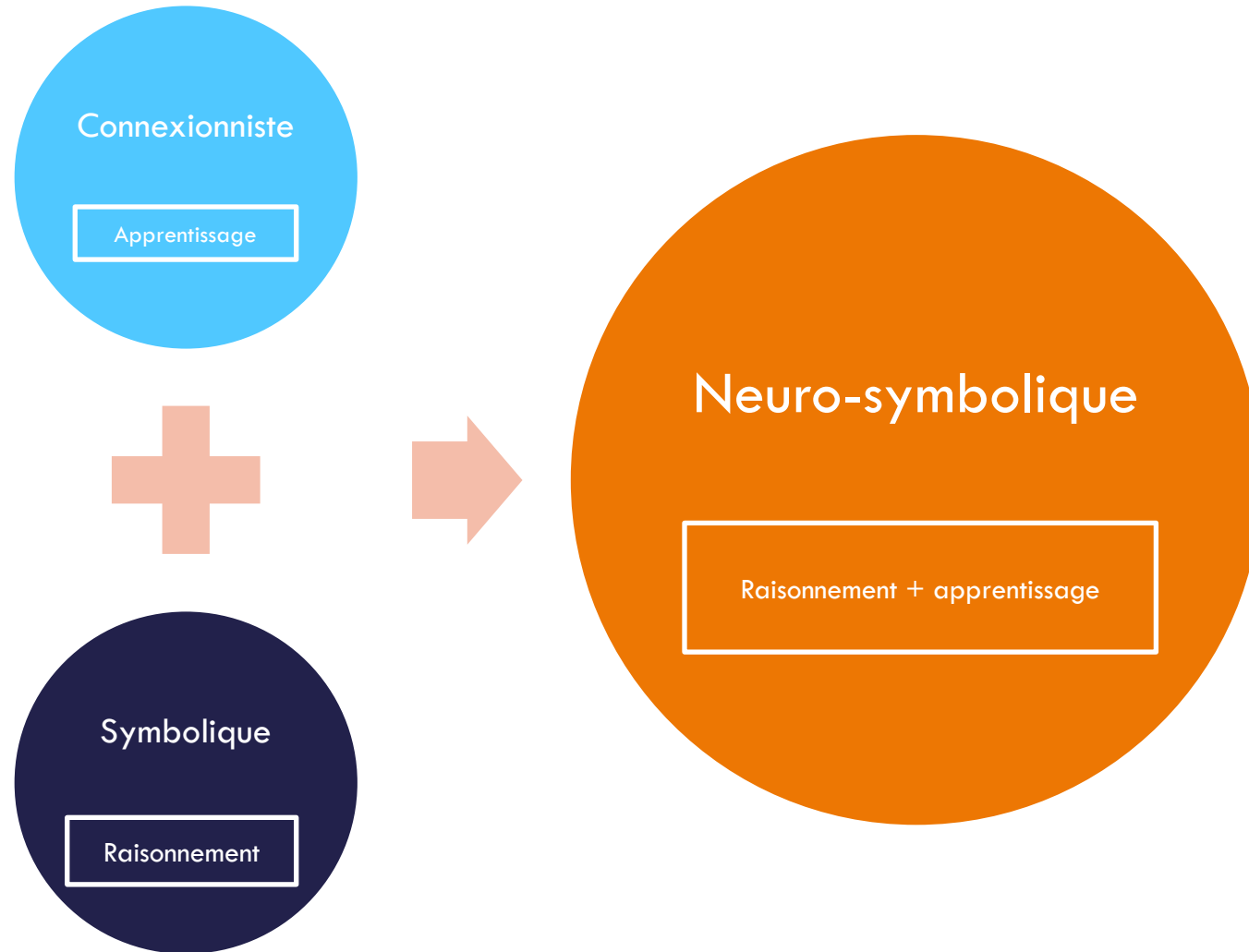
IA Symbolique

- ✓ Modélisation de la connaissances métier
- ✓ Explicabilité des résultats
- × Pas d'apprentissage
- × Pas de manipulation de données non structurées
- × Peu tolérant au bruit

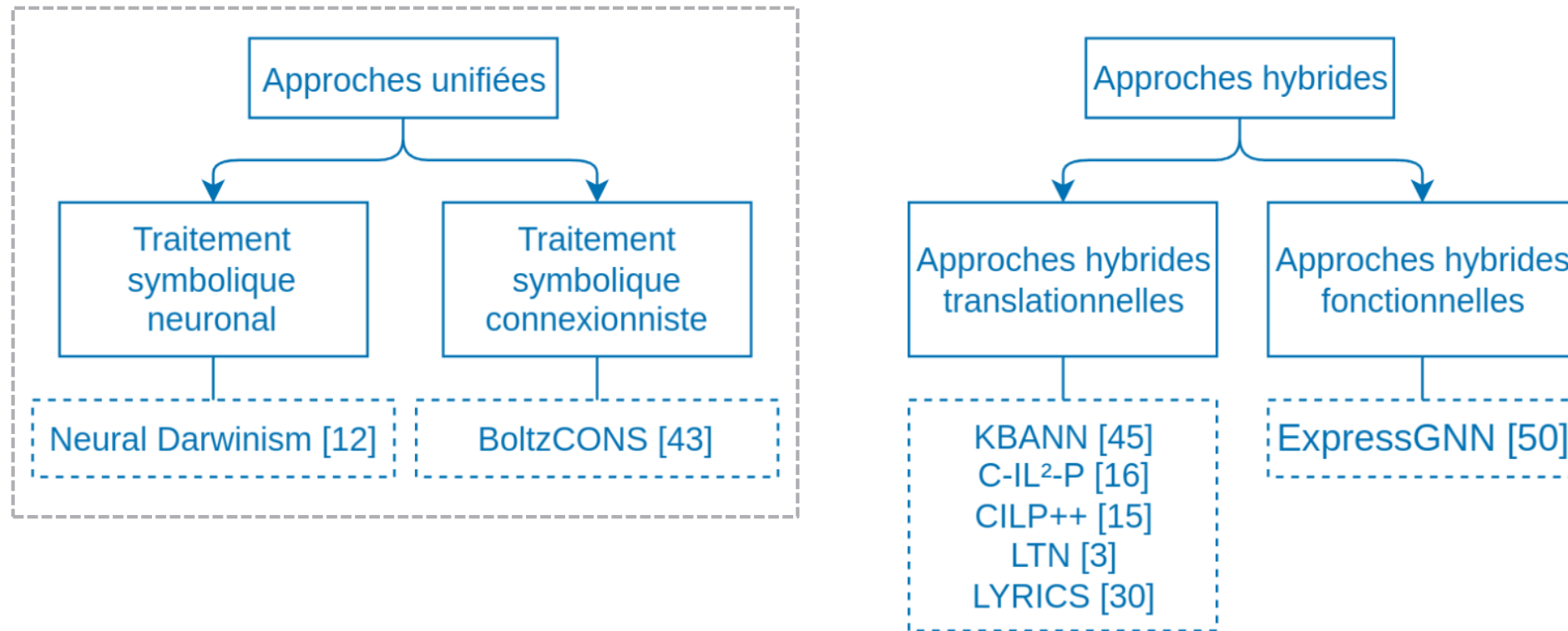
IA Connexionniste

- ✓ Capacité d'apprentissage et de généralisation
- × Résultats difficilement explicables (modèles en boîte noire)
- × Non prise en compte de la connaissance métier

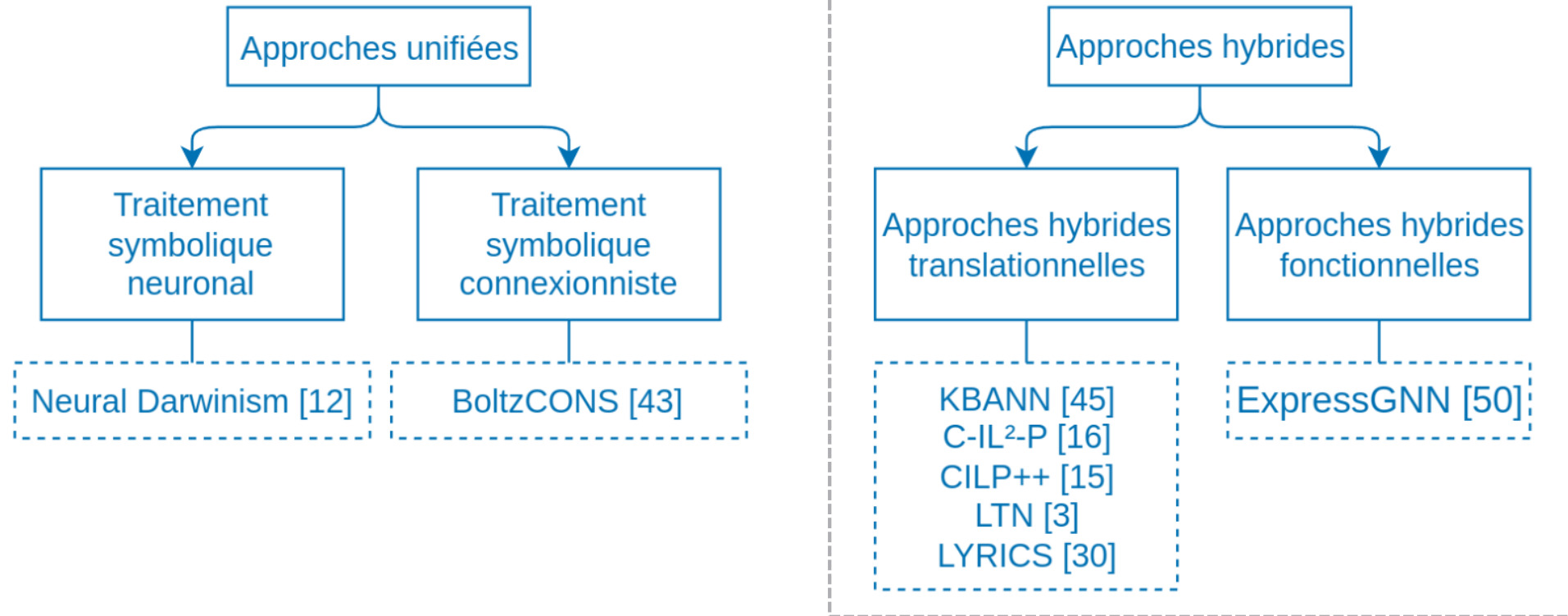
Calegari et al., 2020 [4]
Hitzler et al., 2022 [21]
Yu et al., 2021 [47]



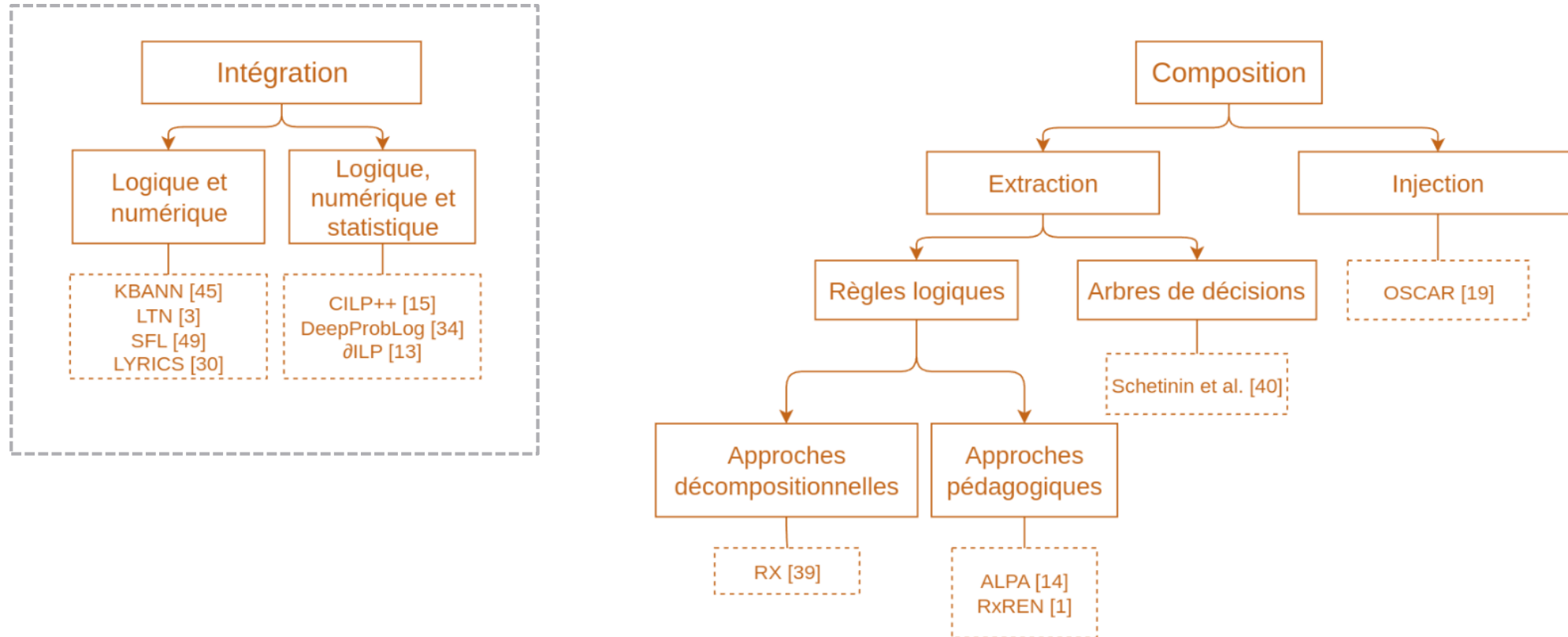
Hilario, 1997 [20]



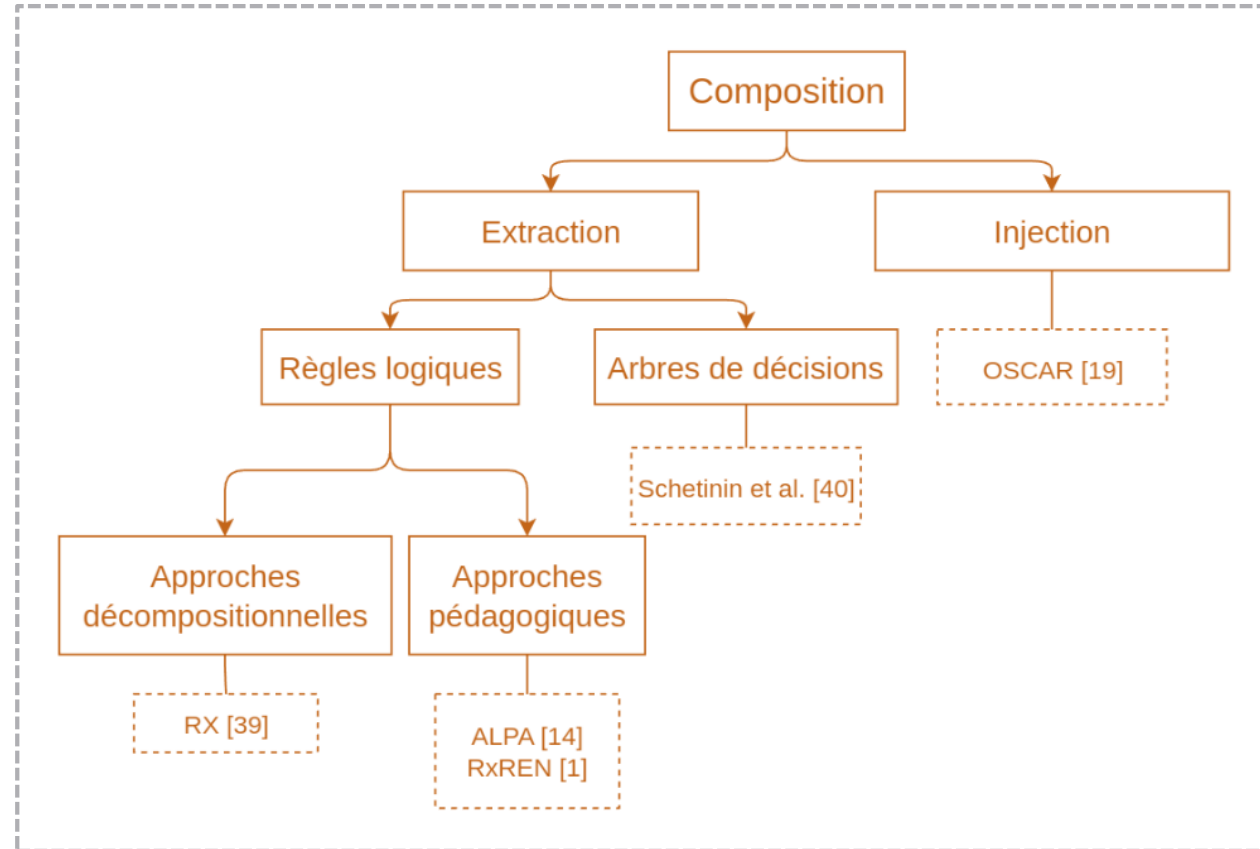
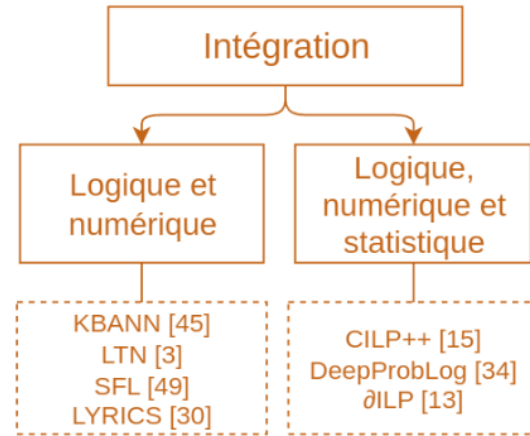
Hilario, 1997 [20]



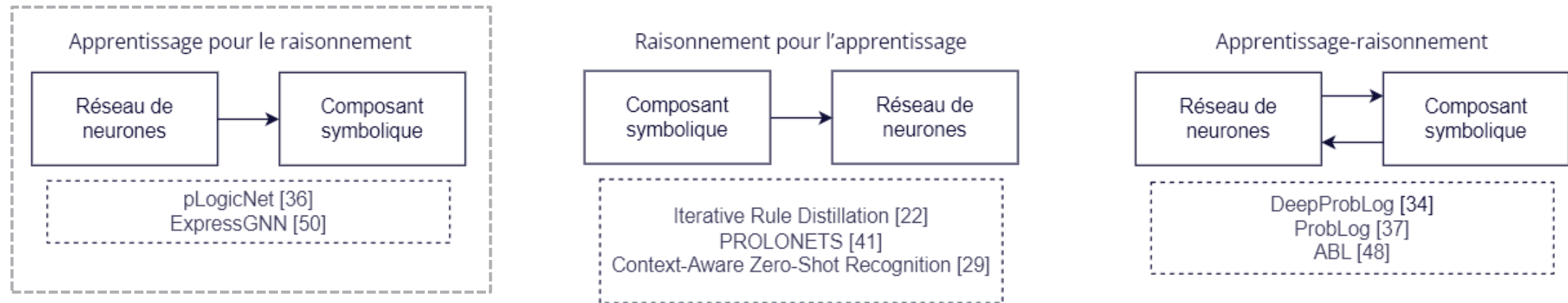
Calegari et al., 2020 [4]



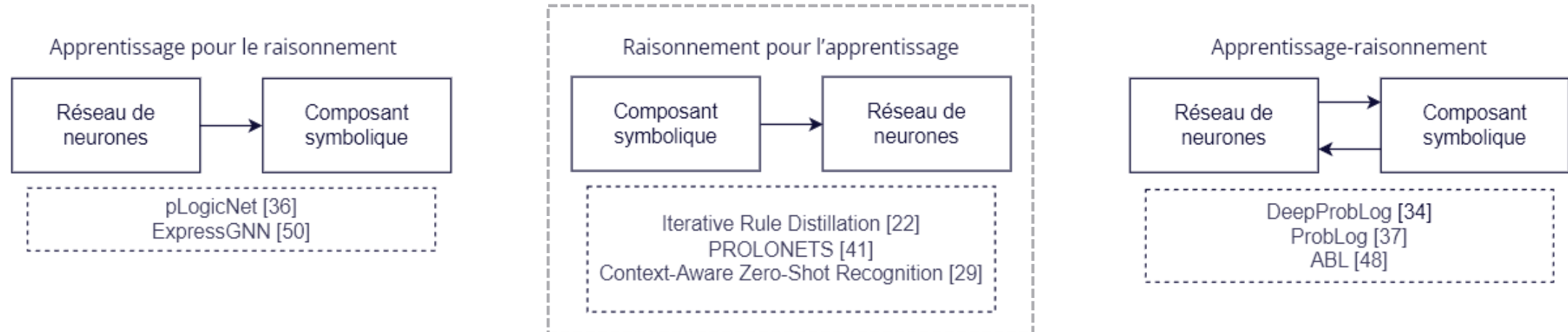
Calegari et al., 2020 [4]



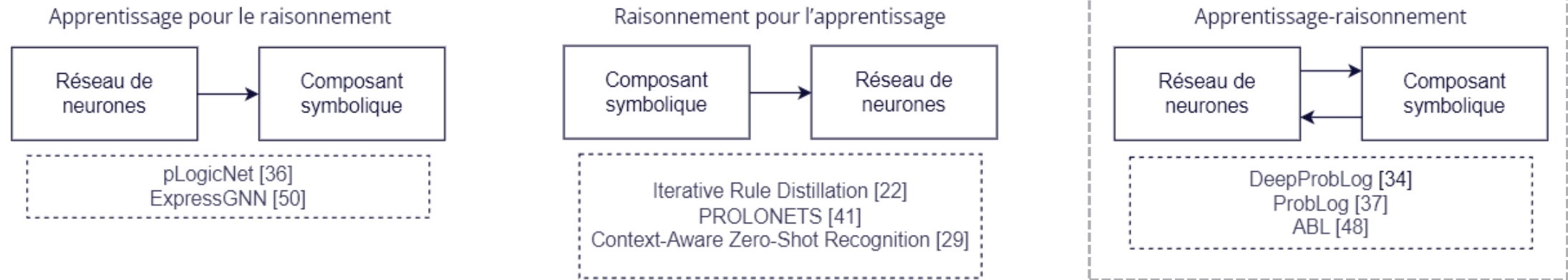
Yu et al., 2021 [47]



Yu et al., 2021 [47]



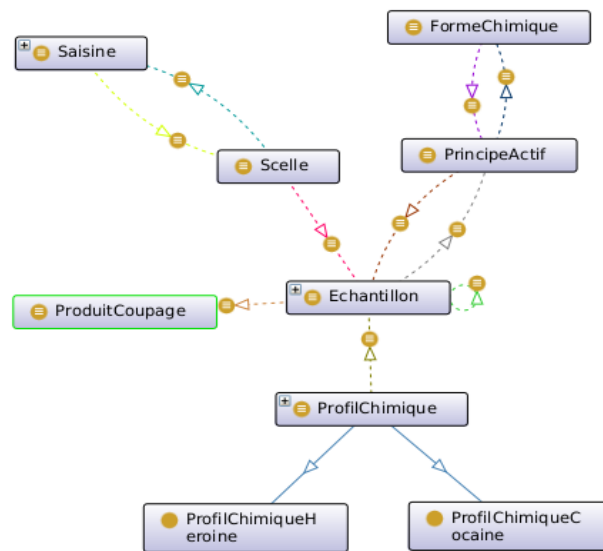
Yu et al., 2021 [47]



(PS1) Spécifier des **connaissances et règles métier par des approches d'IA symbolique** pour des résultats explicables

Représentation la connaissance du domaine métier

Conception d'une ontologie (20ne de concepts, 40ne de propriétés objets et une 40ne de propriété de données).



Extrait de la Tbox de l'ontologie

- owl:topObjectProperty
- aAspectExterne
- aAspectInterne
- aDepartement
- aEchantillon
- aFormeChimique
- aLotInterne
- aPaysDestination
- aPic
- aPrincipeActif
- aProduitCoupage
- aProfilChimique
- aRegion
- aScelle
- aServiceCapteur
- aServiceRequerant
- aSubstance
- aUnite
- aVille
- estAspectExterneDe
- estAspectInterneDe
- estDansDepartement
- estDansPays
- estDansRegion
- estDansSaisine
- estEloigneDe
- estFaitA
- estFormeChimiqueDe
- estLieA
- estLieuSaisineDe
- estLotInterneDe
- estPaysDestinationDe
- estPicDe
- estPrincipeActifDe
- estProcheDe
- estProduitCoupageDe
- estProfilChimiqueDe
- estRepresenteePar
- estServiceCapteurDe

Extrait des propriétés objet

- owl:topDataProperty
- airPic
- caracteristiqueMacroscopique
- abime
- couleur
- couleur1
- couleur2
- couleurInterieure
- couleurInterieure1
- diametre
- epaisseur
- forme
- hauteur
- largeur
- longueur
- masse
- ovule
- presentation
- secabiliteRecto
- secabiliteVerso
- taux
- trace
- unite
- visqueux
- volume
- commentaire
- date
- description
- dosage
- id
- libelleAspect
- libelleFormeChimique
- libelleService
- logo
- nomLogo
- nomService

Extrait des propriétés de données

(PS1) Spécifier des **connaissances et règles métier par des approches d'IA symbolique** pour des résultats explicables

Spécification des règles métier pour l'appariement macroscopique d'échantillons

Exemple : Deux échantillons, initialement *classés comme proches chimiquement*, s'avèrent *être éloignés* si *leurs formes chimiques ne sont pas les mêmes*.

estEloigneDe(x,y) ←

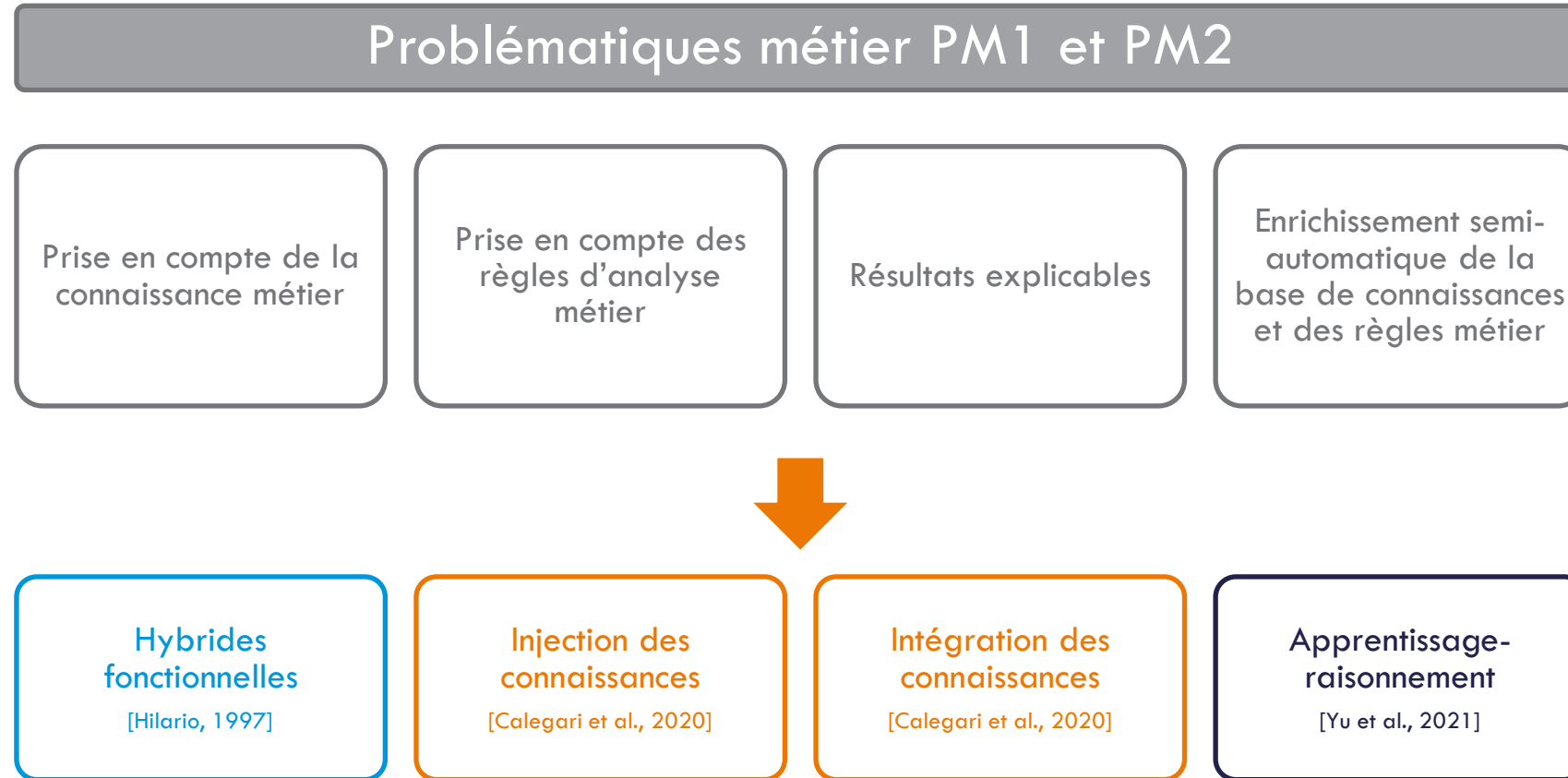
estProcheChimiquementDe(x, y) ∧
aPrincipeActif(x, pa1) ∧ aPrincipeActif(y, pa2) ∧
aFormeChimique(pa1, fc1) ∧ aFormeChimique(pa2, fc2) ∧
typeFormeChimique(fc1) ≠ typeFormeChimique(fc2)



```

INSERT
{
  ?e1 stups:estEloigneDe ?e2
}
WHERE
{
  ?e1 stups: estProcheChimiquementDe ?e2 .
  ?e1 stups:aPrincipeActif ?p1 .
  ?e2 stups:aPrincipeActif ?p2 .
  ?p1 stups:aFormeChimique ?f1 .
  ?p2 stups:aFormeChimique ?f2 .
  FILTER (?f1 != ?f2)
}
  
```

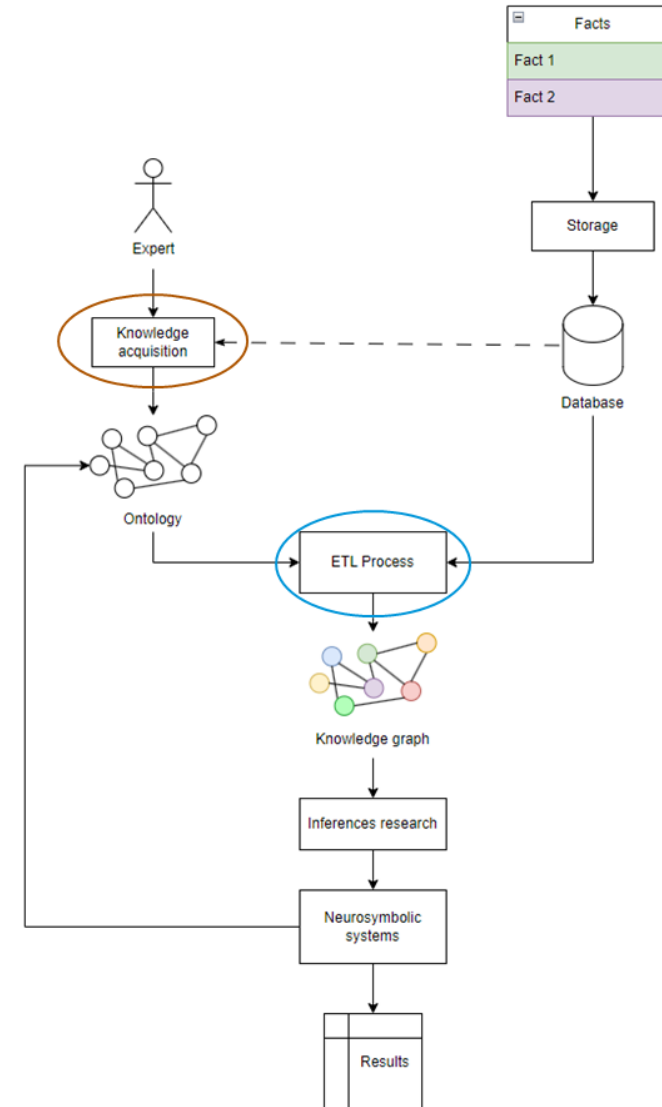
(PS2) Coupler des approches d'IA symbolique et d'IA connexionniste pour la prédiction et l'aide à la décision



- Etude des BD
- Construction de la Tbox de l'ontologie
- Processus ETL
- Ajout des règles logiques
- Déploiement sur GraphDB

Processus ETL disponible à l'adresse :

<https://github.com/SebastienGuillemin/etl>



Conclusion

- Article prospectif présentant les travaux réalisés durant cette première année de thèse

Travaux futurs

- Étude des apports et des limites de l'IA connexionniste dans l'appariement automatique d'échantillons
- Spécification de modèles neuro-symboliques pour l'appariement automatique d'échantillons
- Spécification de modèles prédictifs pour caractériser les produits stupéfiants circulant en France
- Spécifier les liens entre informatique granulaire et problématiques métier

Merci pour votre attention

sebastien.guillemine@u-bourgogne.fr

- [Calegari et al., 2020] R. Calegari, G. Ciatto, A. Omicini, On the integration of symbolic and sub-symbolic techniques for XAI: A survey, *Intelligenza Artificiale*, Vol.14(1), pp.7-32, 2020.
- [Dastres et al., 2021] R. Dastres, M. Soori, Artificial Neural Network Systems. *International Journal of Imaging and Robotics (IJIR)*, 21 (2), pp.13-25, 2021.
- [Flasiński et al., 2016] M. Flasiński, Symbolic Artificial Intelligence. In: *Introduction to Artificial Intelligence*. Springer, Cham. pp. 15-22, DOI : 10.1007/978-3-319-40022-8_2, 2016.
- [Hilario, 1997] M. Hilario, 1997, An overview of strategies for neurosymbolic integration, *Connectionist-Symbolic Integration: From Unified to Hybrid Approaches*, Psychology Press, 1997.
- [Hitzler et al., 2022] P. Hitzler, A. Eberhart, M. Ebrahimi, M. K. Sarker, Lu Zhou, Neuro-symbolic approaches in artificial intelligence, *National Science Review*, Volume 9(6), 2022.
- [Russell et al., 2021] S. Russell, P. Norvig, *Intelligence Artificielle – Une approche moderne*, Traduit par L. Miclet, F. Popineau, C. Cadet, 4e éd. Pearson, 2021.
- [Studer et al., 1998] R. Studer, V.R. Benjamins, D. Fensel, Knowledge engineering: principles and methods, *Data & knowledge engineering*, Vol. 25(1), Pp.161–197, 1998.
- [Yu et al., 2021] D. Yu, B. Yang, D. Liu, H. Wang, S. Pan, Recent Advances in Neural-symbolic Systems: A Survey, arXiv:2111.08164, 2021.