

# Évaluation de la qualité des données dans l'IoT à l'aide de graphes de connaissances

Oumaima AMAL<sup>1</sup>, Nathalie HERNANDEZ<sup>2</sup>, Thierry MONTEIL<sup>1</sup> prenom.nom@irit.fr

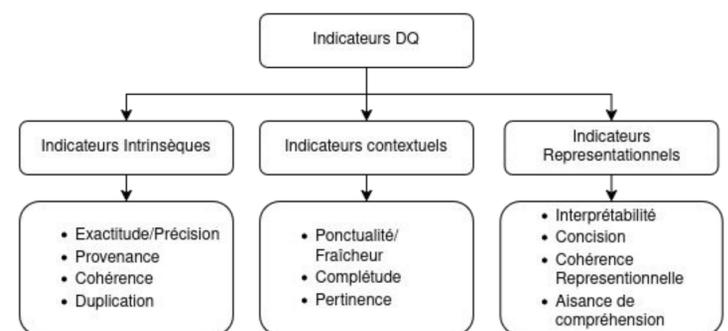
<sup>1</sup> IRIT, UT3, INSA Toulouse, France <sup>2</sup> IRIT, UT2, Toulouse, France

## Contexte - Objectifs

- Comment garantir la **qualité des données IoT** malgré l'hétérogénéité et la complexité des environnements distribués ?
- Fournir une solution générique, réutilisable et adaptable aux besoins de l'IA distribuée dans les réseaux IoT.
- Contribuer à la gestion efficace des données de l'énergie dans le cadre du projet **AI-NRGY** du programme **PEPR TASE**.

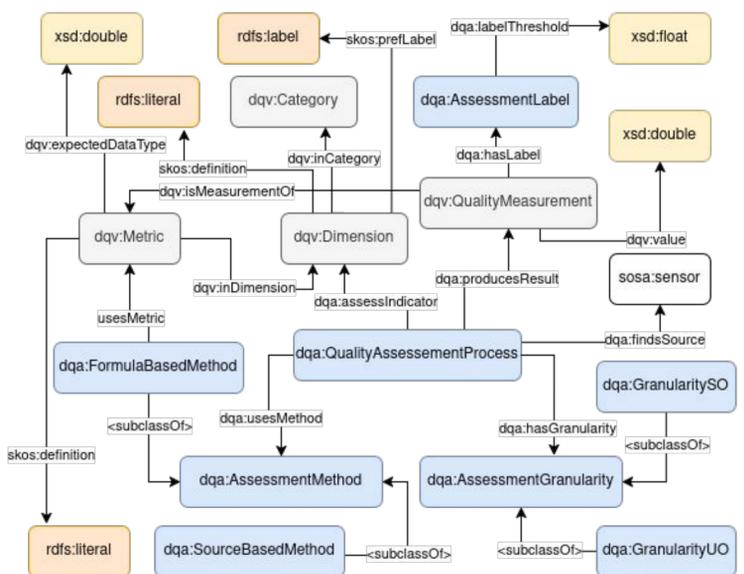
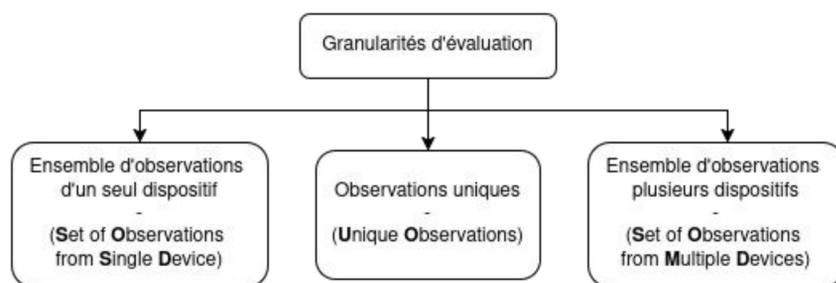
## État de l'art

- Étude des indicateurs de qualité des données proposés dans la littérature, notamment des indicateurs **intrinsèques**, **contextuelles** et **représentationnelles**.
- Revue des ontologies proposées pour la description des observations : **SSN/SOSA**<sup>1</sup> et **OBOE**<sup>2</sup>, et pour l'évaluation de la qualité des données : **DQV**<sup>3</sup>.
- **Limites identifiées** :
  - Approches dédiées à un nombre restreint d'indicateurs de qualité.
  - Aucune analyse sur la pertinence des indicateurs en fonction de la granularité des données (donnée unique, historique des données, agrégation non temporelle de données).
  - Pas d'ontologie permettant la description des processus d'évaluation de QD.

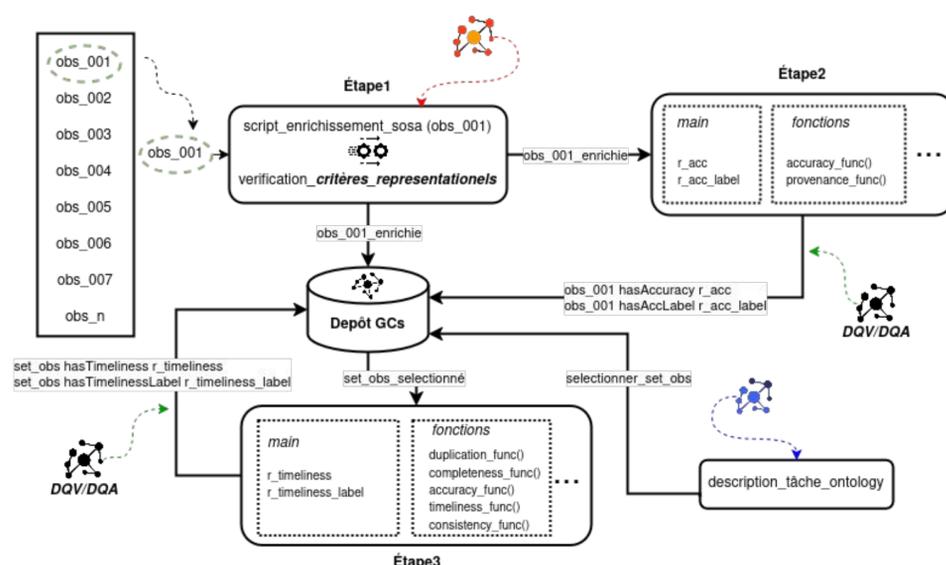


## Contributions

- Proposition de trois **niveaux de granularités** permettant la structuration des indicateurs d'évaluation de la QD : **UO**, **SO-SD** et **SO-MD**.
- Croisement et analyse des indicateurs de QD en fonction de ces trois niveaux de granularité.
- Proposition d'une extension de l'ontologie DQV, nommée **DQA** pour **Data Quality Assessment** (Code QR).



## Approche



## Conclusion - Perspectives

- **Conclusion** : Une approche générique, granulaire et traçable pour évaluer la qualité des données issues de l'IoT en exploitant des ontologies existantes et des graphes de connaissances.
- **Perspectives** :
  - Validation prévue dans un environnement énergétique réel avec une architecture centralisée.
  - Adaptation de l'approche pour une mise en œuvre au sein d'une architecture Fog distribuée.
  - Prise en compte de la mobilité des dispositifs dans des scénarios plus dynamiques.
  - Intégration des dimensions liées à la vie privée et à la sécurité.
- **Remerciements** : Ce travail a bénéficié d'une aide de l'État gérée par l'Agence Nationale de la Recherche au titre de France 2030 portant la référence 22-PETA-0004 - projet AI-NRGY.



<sup>1</sup> <https://github.com/w3c/sdw>

<sup>2</sup> <https://www.w3.org/ns/sosa/oboe>

<sup>3</sup> <https://www.w3.org/TR/vocab-dqv/>