

# Application BASTRI

## Fiches Equipes

### VALDA (SR0824OR)

Valeur à partir des données  
VALDA (SR0765JR) □ VALDA

**Statut:** Décision signée

**Responsable :** Pierre Senellart

**Mots-clés de "A - Thèmes de recherche en Sciences du numérique - 2023" :** A3.1. Données, A3.1.1. Modélisation, représentation, A3.1.2. Gestion, interrogation et stockage, A3.1.3. Données distribuées, A3.1.4. Données incertaines, A3.1.5. Contrôle d'accès, confidentialité, A3.1.6. Optimisation de requêtes, A3.1.7. Données ouvertes, A3.1.8. Données massives (production, stockage, acheminement), A3.1.9. Bases de données, A3.1.10. Données hétérogènes, A3.1.11. Données structurées, A3.2. Connaissances, A3.2.1. Bases de connaissances, A3.2.2. Extraction de connaissances, nettoyage, A3.2.3. Inférence, A3.2.4. Web sémantique, A3.2.5. Ontologies, A3.2.6. Données liées, A3.3. Analyse de données et de connaissances, A3.3.1. Analyse en ligne, A3.3.2. Fouille de données, A3.3.3. Analyse de données massives, A3.4.3. Apprentissage par renforcement, A3.4.5. Méthodes bayésiennes, A3.5.1. Analyse de grands graphes, A4.7. Contrôle d'accès, A7.2. Logique, A7.3. Calculabilité et computabilité, A9.1. Connaissances, A9.8. Raisonnement

**Mots-clés de "B - Autres sciences et domaines d'application - 2023" :** B2. Santé, B3.3. Géosciences, B4. Energie, B4.2. Production d'énergie nucléaire, B9.3. Médias, B9.5.6. Science des données, B9.6.5. Sociologie, B9.6.10. Humanités numériques, B9.7.2. Données ouvertes, B9.9. Ethique, B9.10. Confidentialité, vie privée

**Domaine :** Perception, Cognition, Interaction

**Thème :** Représentation et traitement des données et des connaissances

**Période :** 01/01/2018 -> 31/12/2024

**Dates d'évaluation :** 03/10/2019 ,

**Etablissement(s) de rattachement :** ENS PSL, CNRS

**Laboratoire(s) partenaire(s) :** DI-ENS (UMR8548)

**CRI :** Centre Inria de Paris

**Localisation :** Ecole Normale supérieure Paris

**Code structure Inria :** 021144-1

**Numéro RNSR :** 201622223R

**N° de structure Inria:**SR0824OR

### Présentation

Les recherches de Valda portent sur les *aspects fondamentaux et systèmes de la gestion de données* complexes, en particulier les données produites par l'activité humaine. Les données qui nous intéressent sont généralement hétérogènes, massivement distribuées, d'évolution rapide, coûteuses d'accès et souvent subjectives, peut-être erronées, imprécises, incomplètes. Dans ce cadre, Valda s'intéresse à l'optimisation de ressources complexes telles que le temps et l'espace de calcul, ou des budgets de communications, financiers, ou de privauté. Le but est d'extraire de la *valeur à partir des données*, au-delà d'une simple réponse à des requêtes.

### Axes de recherche

- Fondements de la gestion de données.** Les systèmes qui nous intéressent pour, p. ex., manipuler des données hétérogènes et confidentielles, changeant rapidement et distribuées de façon massive, sont par nature sensibles aux erreurs. De plus, à cause du coup de l'accès aux données, il est important d'optimiser les ressources nécessaires à leur manipulation. Cela ne peut être accompli qu'avec des fondements solides de ces systèmes de gestion de données. Ces fondements sont une base pour des spécifications appropriées (règles de confidentialité, propriétés de robustesse, etc.), pour des vérifications formelles et en temps réelles des spécifications, pour la conception de langages de requêtes adaptés (avec une bonne expressivité, un usage limité des ressources), pour la conception de bons index (pour l'évaluation optimisée), etc.

### Contact

- **Responsable :** Pierre Senellart
- **Tél :** + 33. 1. 4.4 .32. 2.0 .61
- **Secrétariat Tél :** + 33. 1. 8.0 .49. 4.0 .33

### En savoir plus

- Site de l'équipe
- Site sur [inria.fr](http://inria.fr)
- Site du [responsable](#)
- Derniers Rapports d'Activité : [2017](#), [2018](#), [2019](#), [2020](#), [2021](#), [2022](#), [2023](#)

### Documents sur la structure

- [Intranet](#)
- [Privés](#)

### Décisions

- [12569](#) (18/12/2017) : création
- [14331](#) (30/07/2020) : prolongation
- [16994](#) (26/04/2024) : prolongation

### Localisation

- **Adresse postale :** École Normale supérieure 45 rue d'Ulm 75005 Paris France
- **Coordonnées GPS :** 48.841898, 2.345021

2. **Incertitude et provenance des données.** Cet axe de recherche traite de la modélisation et gestion efficace de données accompagnées d'incertitude (distributions de probabilité, incomplétude logique, etc.) et d'informations de provenance (indiquant là d'où les données proviennent). Les outils et fondements pour la gestion d'incertitude et de provenance sont souvent similaires.
3. **Systèmes de gestion d'informations personnelles (Personal Information Management System).** Un PIMS est un système qui permet à un utilisateur d'intégrer ses propres données, p. ex., ses méls et d'autres types de messages, son calendrier, ses contacts, ses recherches web, ses réseaux sociaux, ses informations de voyage, ses projets de travail, etc. De telles informations sont souvent réparties sur des services distincts. Le but est de rendre à un utilisateur le contrôle de ses informations, en lui permettant de formuler des requêtes telles que « Quel type d'interaction ai-je eu récemment avec Alice B. ? », « Où ont été mes dix derniers déplacements professionnels, et qui m'a aidé à les préparer ? » Le système doit orchestrer des requêtes vers différents services et intégrer l'information résultante, p. ex., aligner une localisation GPS de l'utilisateur avec une adresse d'un commerce ou un endroit mentionné dans un mél, ou un événement du calendrier avec un événement mentionné dans une recherche Web.

Relations industrielles et internationales