

PRESTATIONS ET ACCOMPAGNEMENT EN DATA SCIENCE



DESCRIPTION DE L'ÉTABLISSEMENT

L'Abes - Agence Bibliographique de l'Enseignement Supérieur- est une agence nationale qui administre une gamme d'outils et de services mutualisés au profit de l'ensemble des établissements documentaires de l'ESR : Sudoc (catalogue collectif des BU, 16M de notices), theses.fr (500K thèses), IdRef (référentiels de 5M d'auteurs, de concepts, etc.) et autres bases. Elle est aussi acheteuse de ressources électroniques pour le compte de l'ESR. 90 agents à Montpellier. [En savoir plus.](#)



CONTEXTE

Les catalogues et référentiels que gère l'Abes sont de grandes bases de données très structurées, alimentées par des milliers d'experts données soucieux de qualité. En complément de ce catalogage de masse par les BU, des millions de notices d'articles ou de chapitres sont récupérées auprès des éditeurs, modélisées, converties, enrichies et exposées sur le web (y compris le web de données). L'Abes est assise sur une montagne de données très structurées qu'il s'agit d'explorer, d'évaluer et d'améliorer, pour mieux les exploiter.



IRRITANT A RESOUDRE

Les équipes de l'Abes sont essentiellement constituées de bibliothécaires et d'informaticiens.

Les principales **technologies orientées** données que l'Abes maîtrise sont les bases de données relationnelles (depuis toujours), la famille XML (XSLT, XQuery, schémas, Bases XML), les technologies du web sémantique (Bases RDF, SPARQL, schémas, ontologies) et les moteurs d'indexation (Lucene).

Cette palette de technologies nous donne de la polyvalence pour convertir ou interroger nos données, mais ne suffisent pas pour effectuer d'autres tâches : exploration des données, restructuration des données, enrichissement des données.

L'Abes est handicapée par son **manque de compétences** dans deux domaines qui sont désormais incontournables, a fortiori pour une organisation dont les données est le cœur de métier : le **traitement du langage naturel** et ce qu'on peut regrouper sous l'expression « science des données », en y incluant le **data mining** et le **machine learning**.

Le marché de l'emploi et la grille salariale du public étant ce qu'ils sont, on ne peut guère envisager le recrutement de data scientists. Inversement, la formation continue n'est pas une voie réaliste pour

transformer rapidement nos bibliothécaires et nos informaticiens en data scientists productifs. On se contente alors de petits pas et de timides essais, en ayant le sentiment de passer à côté d'une révolution technologique et surtout de sous-exploiter nos données, et donc de dégrader notre offre de services.



L'ENJEU DE L'APPEL A PROJET

Nous sommes arrivés à la conclusion suivante : nous avons besoin de nous faire accompagner par une **entreprise locale, spécialisée en TAL et science des données**. Cet accompagnement comprendrait deux volets : **prestations et transfert de compétences**. Notre objectif est double : à court terme, réaliser des opérations précises sur nos données (extraction, classification automatique, entity linking, détection d'anomalies, etc.); à moyen terme, à l'occasion des prestations, faire monter en compétences nos équipes.

Cette prestation d'accompagnement à deux faces suppose de la durée dans le temps et de la proximité spatiale.

Les métadonnées bibliographiques (et celles qui gravitent autour) sont des données complexes qui demandent un réel investissement pour les appréhender. D'où l'importance d'une relation commerciale durable : pour être efficace, notre prestataire ne doit pas redécouvrir nos spécificités à chaque prestation. De plus, le transfert de compétences suppose un temps d'acculturation réciproque entre nos équipes et celles de l'entreprise.

De même, travailler avec une entreprise locale donne la possibilité de se voir rapidement, et souvent si nécessaire. C'est une condition très favorable pour des prestations et un transfert de compétences de qualité.

Cet accompagnement a deux faces mais elles ne font qu'un : il ne s'agit pas de réaliser des prestations d'un côté, et de dispenser des formations de l'autre. C'est bien à travers les prestations sur telle tâche que le transfert de compétences devrait se faire.

Le contexte de collaboration envisagé

L'Abes cherche à nouer une relation commerciale durable et locale, dans le cadre de la commande publique. Une première prestation permettrait de tester la faisabilité et de calibrer les modalités d'une telle relation.



THEMATIQUE ET MOTS CLES

Formation et vie de l'étudiant (pédagogie et gestion associée)	Recherche	X Conservation culturelle (bibliothèques)
--	-----------	---

Métiers supports/transverses (précisez)

RH Finances Gestion des missions Achats Juridique Vie de l'établissement/Vie du campus	Gestion du SI Gestion de la Sécurité : coté SI et coté établissement Moyens généraux Immobilier :	Gestion de la documentation et de la connaissance (GED, archivage etc...) Gestion des données Pilotage Autre : précisez
---	--	--

Technologies, usages (précisez)

Autour de l'enseignement et de la formation Serious game Podcast Learning Spaces Réalité virtuelle Réalité augmentée Téléprésence Autre : précisez	Autour de la recherche Réalité virtuelle la Réalité augmentée Iot Autre : précisez	Autour de la donnée (transverse) : Démat GED Editique Open Data <input checked="" type="checkbox"/> Data mining Data Vizualisation Big Data Décisionnel <input checked="" type="checkbox"/> Autre : TAL, machine learning
Autour du SI : Cloud Sécurité SI Administration SI, supervision hypervision Coffre forts BlockChain Forge Assistance Autre : précisez Service accompagné auto formé	Sécurité hors SI Télé-alerte SMS Autre : précisez	Autour du patrimoine Iot BIM Réalité augmentée Autre : précisez
Sujets transverses		
e-reputation Télétravail Chatbot <input checked="" type="checkbox"/> IA	Collaboration Intelligence collective RPA Interactif display Autre : précisez	