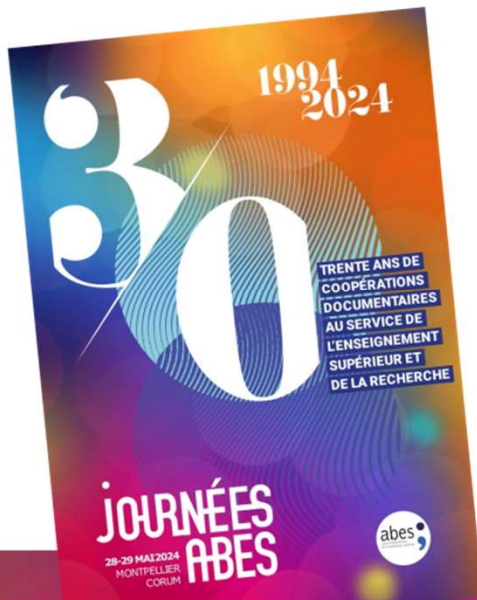


Des métadonnées pour les organisations de l'ESR

La première réinformatisation de l'Abes

Stéphane Gully, responsable du
département informatique de l'Abes

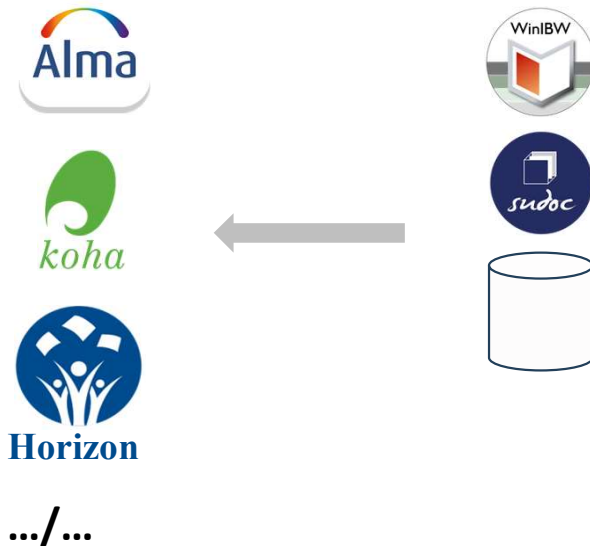


Au programme

- Le système d'information actuel de l'Abes
- Objectifs généraux de la réinformatisation
- Le projet SGM2024 : calendrier
- L'étude du marché
 - Les objectifs du « sourcing » et du « benchmark »
 - Les tendances observées : tailles & stratégies
 - Les tendances observées : données
 - Les tendances observées : technique
- Expression des besoins
- Suite des travaux

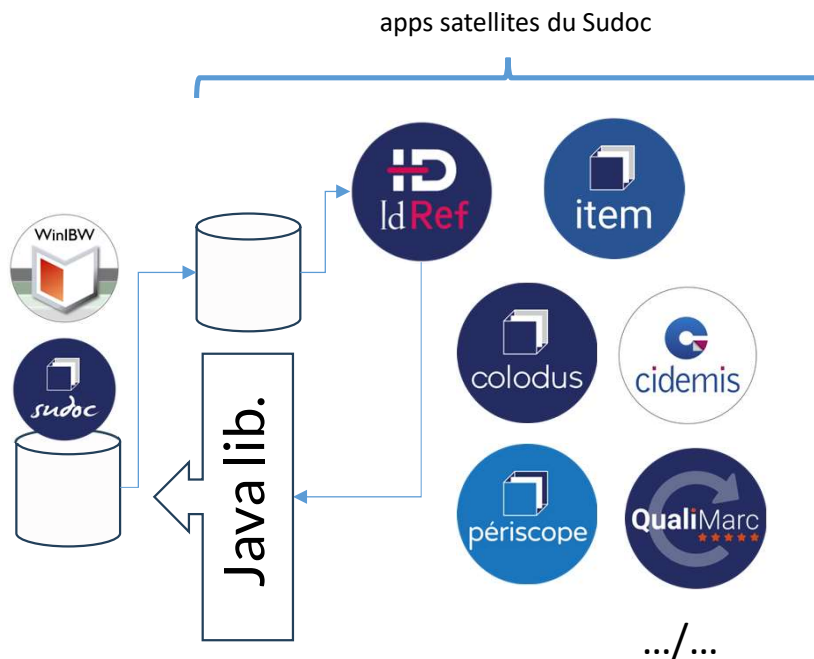
Le système d'information actuel : le cœur

BU - SIGB



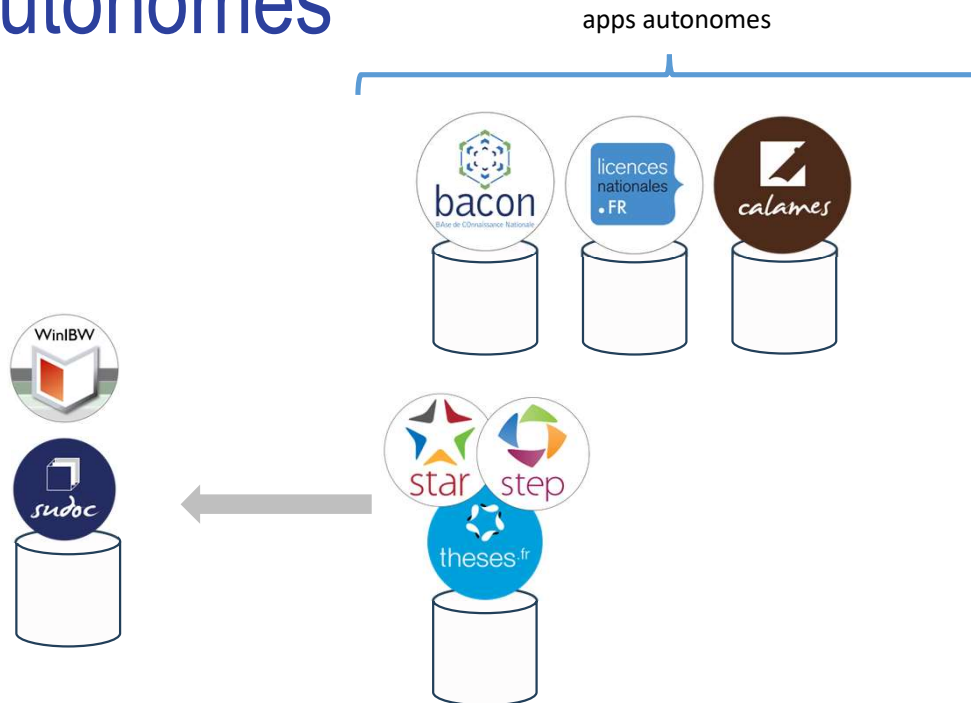
- Depuis l'an 2000
- Base de données : OCLC CBS, format PICA+, ~18m notices
- OPAC OCLC (le sudoc public) : interface en HTML pour la consultation
- Outil de catalogage : OCLC WinIBW, catalogage centralisé en Unimarc
- Notices et exemplaires envoyés dans les SIGB :
 - Batch
 - (S)FTP
 - Unimarc / Marc21

Le système d'information actuel : les applications satellites (du Sudoc)



- Depuis environ 2010, extension du cœur
- Base de données Oracle, copie du CBS
- Émulation de WinIBW pour écrire dans CBS
- Développements maison d'applications non fournies par CBS, mais CBS reste une « base maître », exemples :
 - Gestion avancée des notices d'autorités (IdRef)
 - Donner la main aux bibliothèques pour créer et éditer des exemplaires en masse (item)
 - Contrôle qualité (QualiMarc)
 - .../...

Le système d'information actuel : les applications autonomes



4 apps « autonomes » avec développements maison

Exemple avec apps des thèses :

- Outils de catalogage des thèses :
 - soutenues (star) & en préparation (step)
 - imports massifs de données depuis des logiciels d'écoles doctorales (ex. ADUM)
- Site web public : theses.fr
- Exports (unimarc) vers CBS / Sudoc
 - ~450.000 notices (TEF),
 - ~150.000 pleins textes (PDF)

Objectifs généraux de la réinformatisation

- Pour les bibliothèques membres :
 - plus de métadonnées (plus de volume, fourni dans les canaux de données)
 - plus faciles à utiliser, à moindre coût (de nouvelles interfaces de catalogage plus modernes, ou mieux intégrées)
- Pour l'Abes :
 - simplifier notre architecture globale (tant côté logiciel que côté données)
 - élargir le champ des données que nous gérons
- Pas un produit « fait maison », un produit du marché, mais « souveraineté » et autonomie :
 - Imports, manipulations, et exports de données
 - Capacité à « étendre » le système et à le connecter au paysage informatique global de l'environnement de l'enseignement supérieur français : API first
 - Open Source serait intéressant, mais pas de fork
- 1ère phase du projet, à définir, en production en 2027 : il nous faut un « cliquet »

Voir aussi « La première réinformatisation de l'Abes » Arabesques n°113

Le projet SGM2024 : calendrier général

Objectif : acheter un produit du marché pour réinformatiser l'Abes

<https://projet2024.abes.fr/docs/2.4/projet2024-calendrier>

- **T1+T2 2024 : *sourcing* + *benchmark* + 1^{ère} expression des besoins**
- T3+T4 2024 : rédaction d'un cahier des charges
- T1 2025 – T3 2025 : mise en concurrence et choix d'un prestataire
- T4 2025 – T1 2027 : projet tranche #1
- T1 2026 – T1 2027 : intégration des « satellites » au nouveau SGM
- T2 2027 – T3 2028 :
 - projet tranches #2 (et #3?) avec le prestataire
 - refonte satellites et mise en place de nouveaux services aux membres



- Équipe sourcing / benchmark : 8 personnes
- Équipe expression des besoins : 9 personnes
- Participation des extérieurs : ACEF, AULE, SSI ADBU et KOHALA

Les objectifs du “sourcing”

- Informer les prestataires du projet de l'Abes et les motiver à répondre ;
- Solliciter les prestataires pour obtenir leur avis : notre projet est-il réalisable et à quelles conditions ?
- Identifier les prestataires, leurs solutions et leur feuille de route ;
- Identifier les facteurs de coûts et le modèle économique des fournisseurs ;
- Adapter le montage contractuel et la procédure de passation ;
- Réduire les facteurs de risques identifiables.



Le *sourcing* n'est pas une pré-sélection



CCP en 2016 « *sourcing* » est la possibilité pour un acheteur « d'effectuer des consultations ou réaliser des études de marché, de solliciter des avis ou d'informer les opérateurs économiques du projet et de ses exigences » afin de préparer la passation d'un marché public (article 4 du décret 25 mars 2016)

Les objectifs du « benchmark »

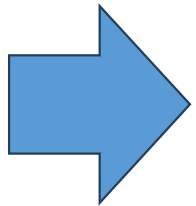
Le parangonnage ou « benchmark » est une approche complémentaire du « sourcing » qui consiste à rencontrer les clients de ses fournisseurs ou de ses fournisseurs potentiels afin d'échanger sur les bonnes pratiques, les stratégies tarifaires (achat ou location par exemple) adoptées et les difficultés rencontrées.

- Entretiens avec des organisations ayant des projets de réinformatisation comparables à l'Abes
 - Types de structures : consortiums américains ou allemands, bibliothèques nationales têtes de réseau
 - Particularités projets : volume de données, de gestion consortiale ou de type de logiciel acheté



Sources d'informations

- Connaissances existantes du marché
- Entretiens avec prestataires (sourcing)
- Entretiens avec des organisations ayant des projets de réinformatisation comparables à l'Abes (benchmark)



Observation des tendances

Les tendances observées : tailles, présence sur le marché

Panel de prestataires sur le marché

- ETP très variables (facteur de 1 à 100) & parfois sous-traitance
- CA très variables (facteur de 1 à 1000)
- Présence :
 - Seulement en France
 - Dans un petit nombre de pays européens (<5)
 - Internationale
- Différentes implications dans le travail normatif sur les métadonnées
- Différentes spécialisations : ESR et/ou lecture publique et/ou GLAM
- Activités plus ou moins diversifiées : types de logiciels pour les bibliothèques, offre de réservoirs de données, domaines d'activités hors bibliothèques

Les tendances observées : stratégies

Projet

- Majorité des produits sont déjà en production chez des clients
- Moyenne des projets de réinformatisation : de 18 à 24 mois
- Hétérogénéité de l'accompagnement
 - avec ingénierie documentaire (données et paramétrage)
 - avec dev spécifiques autour de l'outil du marché (ex : ajout d'API)
 - avec une stratégie de remplacement de logiciel à un instant T ou mise en place progressive d'une nouvelle solution sur plusieurs années
 - sociétés compétentes et efficaces, parfois « rouleau compresseur »
 - effet bénéfique d'une communauté d'usage importante & produit dynamique

Souveraineté et autonomie

- Majorité de systèmes propriétaires mais opensource aussi présent sur le marché
- Ouverture des données via API

Gestion consortiale

- Peu présente dans l'outil
- Développée spécifiquement pour un pays
- Implémentée en natif dans l'outil

Les tendances observées : données

- Perspectives vers des entités / relations :
 - BibFrame domine les préoccupations et est en cours d'implémentation dans les systèmes
 - RDA / RDA-FR n'est pas une perspective envisagée aujourd'hui pour la grande majorité, mais pas d'impossibilité technique à terme
- Ecart entre fonctionnalités et formats :
 - évolution/souplesse de gestion formats
 - aspects fonctionnels du logiciel
- Constante pour la gestion des formats :
 - imports / exports avec mapping multi-formats,
 - utilisation d'un format pivot interne (masqué)
- Grand écart pour la gestion de la doc élec : soit pas gérée dans un circuit intégral, soit gérée au niveau titres, soit associée à un DT (indexés d'articles).

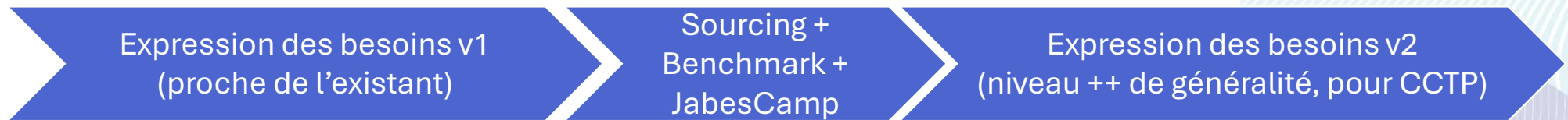
Les tendances observées : technique

- Hébergement :
 - SaaS systématique, voire exclusif et parfois *on premise* possible
 - coût du SaaS $\simeq$ coût du *on premise*
- Base de données : format pivot & SGBDR systématique (ex : postgresql, oracle)
- Accès aux données :
 - APIs systématiques mais variations dans la granularité et la complétude de l'offre d'API (parfois « API first »)
 - Accès direct à la base de données systématiquement interdit
- Extensibilité :
 - Interne au système avec *paramétrage*
 - Dans le système avec *plugins*
 - Externe au système avec *APIs*
- Authentification : SAML très courant
- Pilotage et stats : extraction de données dans un outil de BI très courant

Retour sur les objectifs du “sourcing”

- Informer les prestataires du projet de l'Abes et les motiver à répondre
Notre projet suscite de l'intérêt et on aura des réponses...
- Solliciter les prestataires pour obtenir leur avis : notre projet est-il réalisable et à quelles conditions ?
... à condition d'être raisonnables dans nos demandes : notre projet est jugé ambitieux
- Identifier les prestataires, leurs solutions et leur feuille de route
Le marché est très étroit, avec des gros et des petits acteurs
- Identifier les facteurs de coûts et le modèle économique des fournisseurs
Nous aurons majoritairement des solutions techniques « standard »,
et des offres économiques *ad hoc* du fait du positionnement national de l'Abes
- Adapter le montage contractuel et la procédure de passation
Les produits existent sur le marché, mais des échanges seront probablement nécessaires pour
que notre demande soit correctement évaluée

Expression des besoins



Résumé des besoins vs sourcing/benchmark

- Diminuer le nombre de bases de données et de logiciels de l'Abes
Faisable
- Sortir à terme de la logique des formats Marc → système entités-relations
Atteignable, plutôt via Bibframe
- Faciliter l'export et l'exposition des données
Faisable, avec API et protocoles/formats standards
- Augmenter le volume de métadonnées et faciliter les flux : synchronisation entre systèmes locaux et système Abes, pour éviter la double saisie
**Oui pour synchro avec API mais bémols pour anciens SIGB,
Ça dépend pour le volume des données : DT et granularité +/- fine**

Mercredi
29/05
JabesCamp

Suite des travaux

- CA du 27 juin et ses suites
 - Ajustement des orientations générales
 - Mandat pour rédaction du cahier des charges
 - Mise en place d'un montage financier pour le projet

« À ce stade, on sait que le projet de l'Abes est faisable, et que notre objectif de 2027 reste réaliste »