



L'archive institutionnelle de l'EPFL

<http://infoscience.epfl.ch>

Contexte :

L'EPFL est une université technologique qui souhaite se situer parmi les meilleures en Europe et dans le monde.

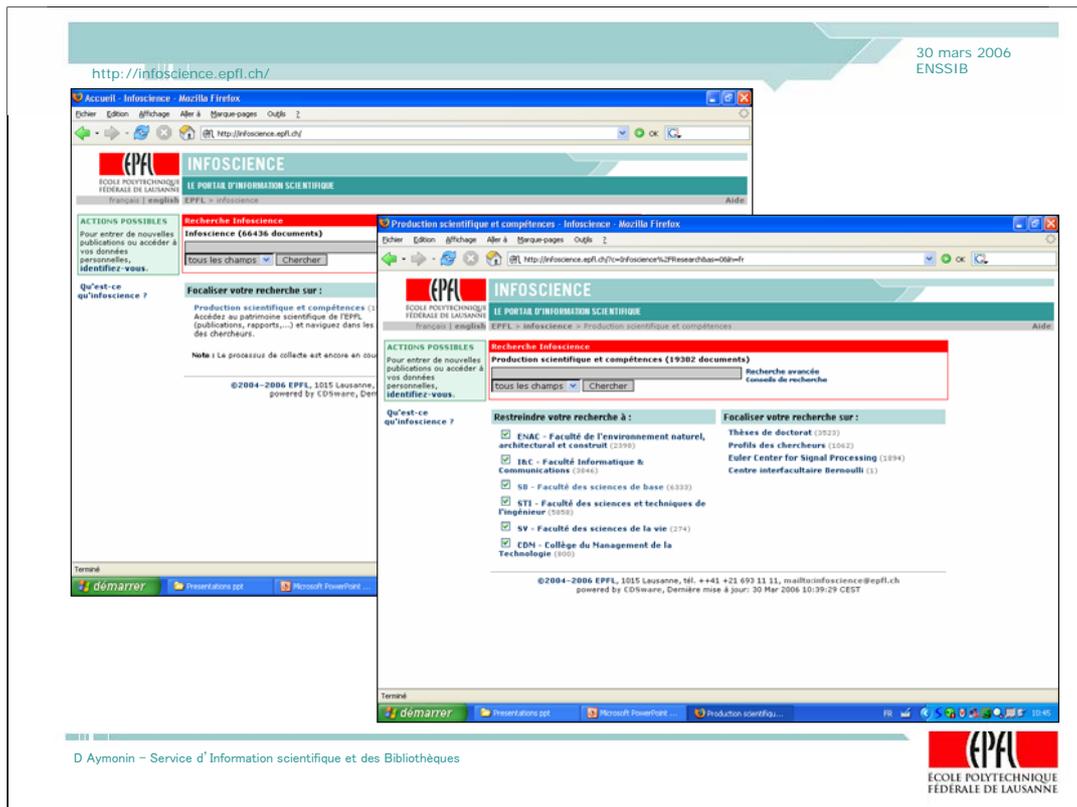
Elle accueille aujourd'hui près de 6000 étudiants, dont plus de 1000 doctorants. 3000 chercheurs et scientifiques y développent des recherches en mathématiques, physique, chimie, management de la technologie, sciences de l'ingénieur, sciences de la vie, architecture.

Infoscience :

En matière d'archives institutionnelles, chaque cas est unique, selon l'environnement administratif, technique, humain, financier de chaque université.

Je ne vais donc pas insister sur les aspects historiques d'Infoscience ni sur la machinerie interne, mais plutôt sur les principes à la base du projet et sur les réflexions et enseignements tirés de l'expérience.

Sachez tout de même qu'Infoscience a pu être mis en œuvre grâce à l'apport des compétences du service informatique central de l'EPFL et du service d'information scientifique et des bibliothèques, après l'expérience acquise lors du développement du rapport scientifique de l'EPFL, qui en présentait les publications principales et les laboratoires.



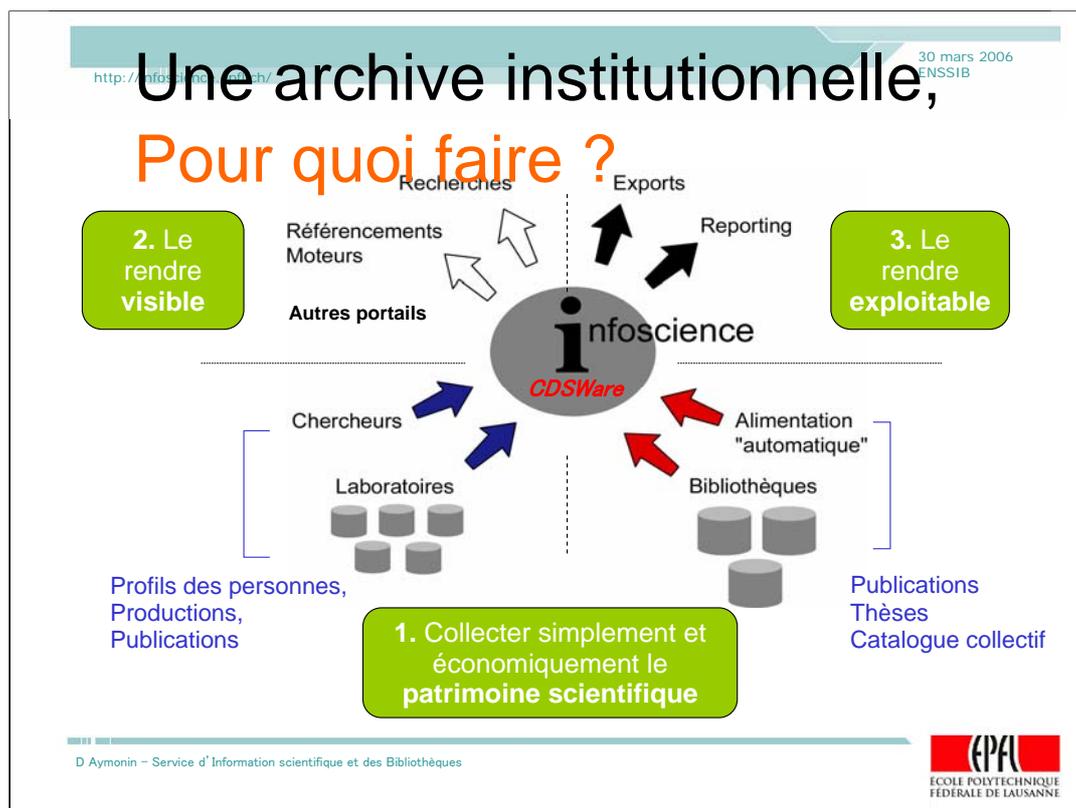
Avant de vous donner une définition d'Infoscience, une démonstration en ligne permet d'en présenter

Le contenu

- types de documents et statut des publications (pre, post print, peer reviewed) sont visibles dans l'écran de recherche avancée
- Les profils des personnes basés sur l'application People@epfl.ch et liés aux publications
- Les thèses numérisées depuis 1920, avec leur archivage sur le serveur de la bibliothèque centrale
- Le catalogue collectif des bibliothèque, toujours en cours de réalisation

Les services

- La réappropriation des données par les chercheurs dans leurs profils et par les laboratoires sur leurs sites web



Infoscience est donc SERVICE de collecte et de valorisation du patrimoine scientifique.

Le patrimoine scientifique est constitué selon nous, des productions scientifiques de l'EPFL, des collections de documents scientifiques acquises, et des personnes qui font la recherche, décrites par leurs profils, mentionnant aussi leurs activités et compétences.

Ce service utilise l'outil CDSWARE co-développé avec le CERN.

Sa valeur ajoutée vient de ce qu'il répond à des besoins identifiés chez des acteurs et publics de la recherche dans l'Ecole.

L'idée clé est que toute publication saisie une fois dans Infoscience est réutilisable par le chercheur qui l'y a mise, par sa faculté, ou par la Direction de l'école. Il n'y aura plus jamais à la ressaisir dans aucune autre application. Et elle sera conservée indéfiniment, avec ou sans full text, et pourra être enrichie au cours du temps.

La réappropriation des publications par les utilisateurs est un service essentiel sur lequel l'équipe est en permanence mobilisée pour répondre aux besoins nouveaux, même individuels.

Bénéfices attendus de l'archive institutionnelle

<i>Acteur</i>	<i>Effet direct</i>	<i>A court terme</i>	<i>A moyen terme</i>
Chercheur	Visibilité	+ de citations	Meilleur IF dans WOS(*)
Laboratoire	Visibilité Efficacité	Idem	Idem + économie de moyens + mémoire
Faculté	Visibilité Tableau de bord	Meilleure information	Meilleures décisions
Direction EPFL	Tableau de bord	Meilleure information	Meilleures décisions
	Analyse des données de publication	Anticipation et réactivité	Renforcement du dialogue avec le Conseil des EPF
	Cohérence des données dans WOS	Fiabilité	Influence sur les classements internationaux

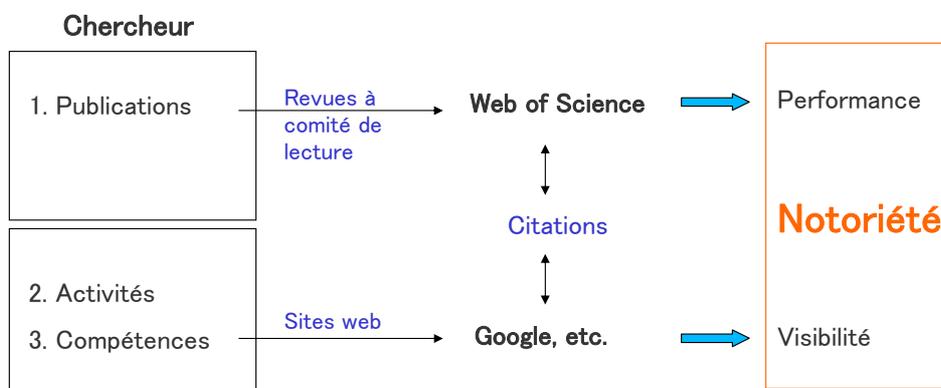
Les publics cibles sont des « CONSOMMATEURS ».

Ils contribuent à Infoscience parce qu'ils en tirent des bénéfices qui leur simplifient la vie ou même l'améliorent.

Chaque laboratoire qui entre dans Infoscience bénéficie d'un service personnalisé pour la reprise de son patrimoine existant (analyse de ses données, import et validation) et l'adaptation du service à ses besoins (site web, formats d'export, ...).

A l'échelle de l'EPFL, Infoscience doit permettre d'économiser du temps et de l'énergie consacrés à des tâches répétitives, tout en donnant des outils de pilotage aux différents acteurs.

Quels facteurs d'impact aujourd'hui ?



Petite réflexion personnelle :

Dans le monde d'avant le web, la performance des chercheurs pouvait s'analyser au travers du Science citation index, devenue WOS.

La notoriété était aussi un facteur important, traditionnellement dépendant des citations.

Avec le développement du web, la visibilité online devient un facteur essentiel de notoriété.

Or les pages web des professeurs ou des laboratoires, avec leurs CV et liste de publications disparates sont moins visibles qu'une archive institutionnelle comme Infoscience, qui structure l'information, et est indexée de manière systématique et en profondeur par les outils de recherche sur le web (Google Scholar, Scirus, etc...).

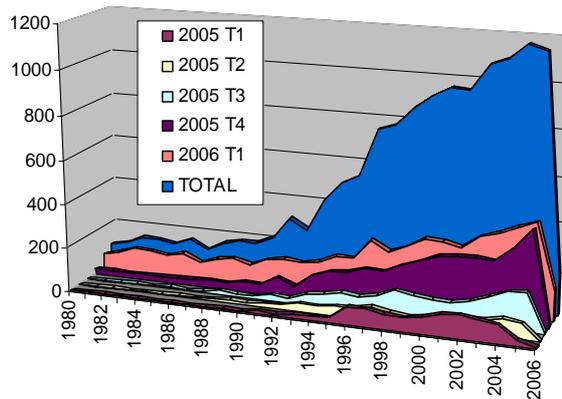
Le monde de la recherche est compétitif et les outils d'influence sont une nécessité. Les chercheurs vivent dans cet univers et leur attitude et leurs besoins vis-à-vis des services d'information scientifique ou de communication sont à analyser en ayant ceci en tête.

Ce qui distingue peut être l'EPFL d'autres institutions de recherche, c'est que cette compétition est reconnue explicitement. De manière pragmatique, la plupart des chercheurs, mais surtout la Direction de l'Ecole, cherchent à utiliser à leur profit les outils et pratiques de la mesure du mérite scientifique. Ces outils sont imparfaits mais ce sont ceux qui existent !

Dans le cas particulier de l'EPFL, l'apport d'Infoscience a été mesuré à cette aune par les chercheurs et décideurs de l'Ecole.

Participation des laboratoires

Nombre de publications signalées



Lancement du projet en janvier 2004.
Ouverture du serveur en sept 2004
Aujourd'hui 95 laboratoires (sur 200)
14000 publications environ

Le contact direct et la personnalisation des contacts avec les chercheurs et laboratoires, suivis du bouche à oreille et complétés de quelques messages diffusés à travers les listes de discussion internes ou les organes d'information internes ont été nos principaux outils de sensibilisation.

Aujourd'hui, je juge le développement d'Infoscience très satisfaisant : près de 100 laboratoires ont adhéré **volontairement**.

Ils apprécient les principes d'économie de moyens, de réappropriation des données saisies et l'adaptation de l'outil à leurs besoins spécifiques.

Pour ce qui est des Doyens de facultés nous n'avons pas assez de retours pour savoir si ils vont utiliser l'outil pour le pilotage.

Les membres de la Direction de l'Ecole viennent de donner, début mars 2006, leur soutien au projet et de valider son statut d'activité permanente.

Plan de développement 2006-2007

Délivrables	Méthode / Principe	Coût / Moyen
Annuaire des professeurs contenant 90% des profils en français et en anglais, structurés selon modèle commun	Obligation faite à chaque professeur de créer et maintenir son profil.	Directive EPFL à tous les professeurs. Adaptation de l'outil people@EPFL pour une structuration plus poussée
Dispositif de contrôle de mise à jour continue de l'annuaire avec garantie d'édition online annuelle à l'échéance souhaitée par la Direction	Détection automatisée des profils non mis à jour	Directive EPFL donnant la responsabilité aux facultés de contrôler la mise à jour des profils. Développement d'un outil de monitoring par l'équipe Infoscience.

D Aymonin - Service d'Information scientifique et des Bibliothèques



Grâce à cette décision de stabiliser l'activité, nous avons maintenant une visibilité de 3 à 4 ans.

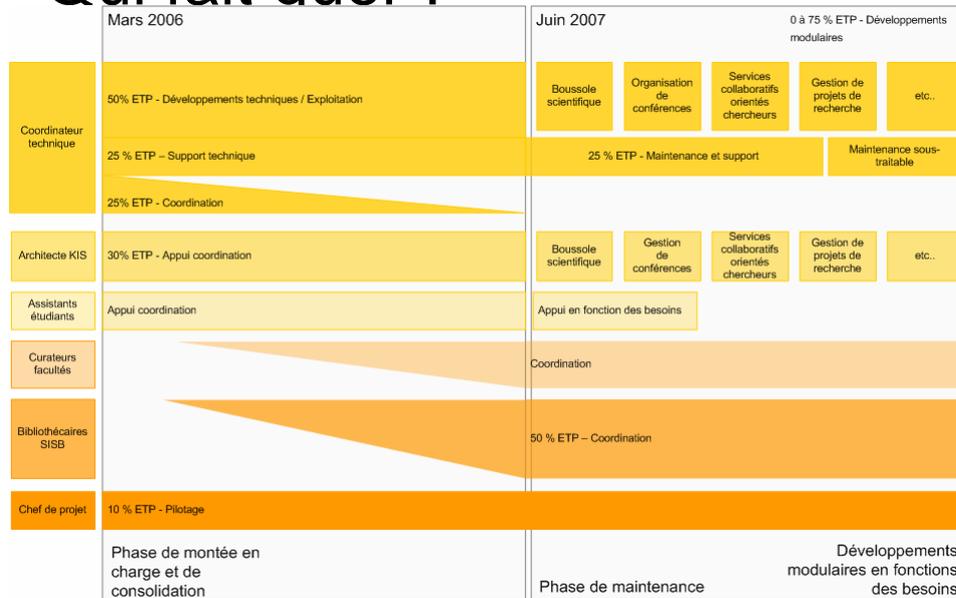
La Direction a accepté le plan de développement qui prévoit un certain nombre de « déivrables », qui prévoient, en substance :

- La constitution et la mise à jour de l'annuaire des scientifiques
- La réutilisation des données accumulées pour le rapport scientifique qui a précédé Infoscience
- Le contrôle de qualité et d'exhaustivité des publications de l'EPFL signalées dans WoS
- Les outils de contrôle de qualité et d'enrichissement des données

Plan de développement 2006-2007 (suite)

Délivrables	Méthode / Principe	Coût / Moyen
Liste exhaustive des publications EPFL signalées dans Web of Science 2000-2006	Récupération de la base des publications gérée jusqu'en 2006 par le service DAR	Travail commun entre le coordinateur technique Infoscience et le DAR.
Signalement complet et exact de toutes les publications EPFL dans Web of Science	Outil de contrôle de cohérence Infoscience et WoS. Demande de correction des erreurs à WoS.	Vérification des données par les Bibliothécaires scientifiques du SISB
Gestion partagée de la qualité des données Voir infoscience.epfl.ch/curator.py	Curateurs de faculté et Bibliothécaires SISB assurent les contrôles et nettoyages de données	Co-développements EPFL et CERN d'outils de contrôle et d'enrichissement (interface curator, DOI, ...)

Qui fait quoi ?



Infoscience, c'est une très petite équipe, mais des compétences pointues, notamment celles du coordinateur technique qui doit savoir dialoguer et coopérer avec de nombreux partenaires et utilisateurs (chercheurs, bibliothécaires, ...) pour entendre les besoins et les transformer en propositions techniques (développements logiciels, flux de traitement) et de services.

Recommandations envisagées

- **Publications admises :**
- **Catégorie A+** : articles signalés dans le Citation Index de Web Of Science
- **Catégorie A** : articles publiés dans des journaux à comité de lecture, communications de congrès à comité scientifique, documents indexés dans ISI index to Scientific Proceedings, Scopus, Inspec, Compendex, Mathscinet, CSA, Medline, Biosis, Avery Index of Architecture...
- **Catégorie B** : prépublications, "technical reports", articles publiés dans des revues sans comité de lecture, communications de congrès sans comité scientifique, brevets, normes, thèses.
- **Fulltext** : Le fulltext des publications est un facteur décisif de visibilité dans les moteurs de recherche de type Google, Google scholar, Citeseer, etc. Sont acceptées les versions préprints, selon les pratiques des facultés, et de postprints, dans le respect de la législation. **Un conseil en droit d'auteur et des informations à jour sur ces questions sont accessibles sur le site Infoscience.**
- **Publications non recommandées :**
- La pertinence de ces publications est très variable et elles ne font pas l'objet de peer reviewing scientifique. Exemple : les posters
- **Publications NON admises :**
- Les articles de vulgarisation ou de communication écrits par des chercheurs et publiés dans des revues non scientifiques ne sont pas admis dans Infoscience.

Le consensus presque trouvé au sein de la Direction sur les publications admises et leur catégorisation doit maintenant être soumis aux facultés qui auront certainement une vision différente des choses pour certains types de publications, selon leur domaine scientifique. Cette diversité est reconnue et admise par la Direction.

Ce transparent révèle cependant l'importance que l'EPFL accorde au fait qu'Infoscience doit donner une BONNE IMAGE de l'EPFL, et est donc un instrument politique sensible.

Par ailleurs nous avons détecté au cours du lancement du projet à quel point les chercheurs avaient besoin d'appui pour comprendre et respecter les règles du droit d'auteur dans le contexte nouveau du web généralisé. Avant même de parler d'Open access, il faut souvent répondre à des questions de base comme « à qui appartiennent mes publications ? ».

Le service Infoscience s'est donc enrichi d'une prestation de conseil en droit d'auteur, assuré par un spécialiste de l'édition scientifique, via une FAQ sur le site et un help desk accessible par e-mail. Des séminaires et conférences sur le sujet sont aussi à l'ordre du jour.

Dès la saisie d'une nouvelle publication, le chercheur est informé de son devoir de respecter le droit d'auteur et des moyens lui sont signalés pour l'aider à le faire (pré- post- prints, politique des éditeurs, open access, limitation de l'accès au pdf, etc.).

Infoscience, et au-delà une archive institutionnelle, est selon notre définition et notre pratique, un service global qui va de l'amont à l'aval de la production scientifique. De la production à la visibilité.