



**MINISTÈRE  
DE L'ENSEIGNEMENT SUPÉRIEUR  
ET DE LA RECHERCHE**

## Présentation du projet ISTEK

Dans le cadre des Initiatives d'excellence, l'ANR a signé avec le CGI, le MESR et le CNRS, une convention d'attribution d'aide au projet ISTEK, à hauteur de 60M€ sur trois ans.

ISTEK est un vaste programme d'acquisition d'archives scientifiques pour créer une bibliothèque numérique aux meilleurs standards internationaux, accessible à partir de toutes les universités du territoire. C'est également un moteur de recherche adapté aux besoins des chercheurs, offrant de nombreux services facilitant la recherche (extractions et fouilles de données, facilités de téléchargement...).

C'est donc un instrument puissant de compétitivité pour la recherche, qui garantira aux chercheurs, en quelques clics, un accès à tout le patrimoine scientifique disponible dans toutes les disciplines.

Cette vaste bibliothèque numérique, commune à l'ensemble des universités et des organismes de recherche, va réduire les inégalités territoriales en garantissant une équité d'accès à la connaissance sur tout le territoire.

### Le projet

ISTEK est un projet novateur qui offrira des services innovants d'accès à la connaissance et associe deux dimensions : une capacité sans précédent en France d'acquisition de ressources électroniques (55M€), et une plateforme nationale (5M€) pour les rendre accessibles et les exploiter. Ce programme sera un instrument au service de tous les chercheurs relevant de l'enseignement supérieur et de la recherche, quel que soit le statut de leur établissement de rattachement.

#### Un corpus d'excellence

Du fait du changement d'échelle qu'il implique, ISTEK offrira aux chercheurs l'accès à un corpus inédit de ressources. Ce réservoir exceptionnel regroupera un ensemble de collections rétrospectives, intéressant toutes les communautés scientifiques : collections rétrospectives de périodiques scientifiques généralistes, de périodiques scientifiques en sciences exactes et en techniques, en santé, en sciences humaines et sociales, corpus de textes et données bibliographiques. L'importance quantitative de cette offre permettra aussi une avancée remarquable en termes de qualité scientifique, la hissant au niveau des meilleurs standards internationaux : pour la première fois, des ressources habituellement disponibles sur les bases isolées des éditeurs pourront être regroupées, accessibles et traitées simultanément. Ce corpus, loin de se limiter à un ensemble de fichiers PDF image, sera constitué en grande partie par les fichiers plein texte correspondants.

La mise en place de cette initiative d'excellence sera l'occasion d'un dialogue renouvelé avec les éditeurs. L'importance des sommes qui seront consacrées à des acquisitions de collections rétrospectives permettra d'impliquer une grande variété d'éditeurs, divers par la taille et les profils. Par ailleurs, les modalités de négociations des licences nationales acquises stimuleront la qualité de l'offre : une large publicité sera faite à la fois auprès des chercheurs et des propriétaires de données, incitant de ce fait les éditeurs à offrir les produits les meilleurs tant sur un plan qualitatif que financier.

#### Une bibliothèque de ressources numérique mobile, accessible à tous et à distance

Dès les premières acquisitions, l'accès aux ressources sera identique pour l'ensemble des chercheurs, quel que soit leur statut. ISTEK est une réponse aux disparités territoriales et institutionnelles qui constituent

aujourd'hui un frein à la recherche française. Accessibles à distance, par tous les membres des établissements de l'enseignement supérieur et de la recherche, ces ressources seront en outre valorisées par des outils de recherche et des services à valeur ajoutée permettant de les exploiter pleinement. En premier lieu, un confort de navigation : l'accès à ces collections pourra être intégré à l'environnement numérique local de chaque chercheur (avec les autres ressources) donc accessible partout sans heurt ni rupture de navigation. Mieux encore, les données intégrées sur la plateforme dédiée pourront être manipulées par des outils de recherche portant sur l'ensemble du corpus, quelle que soit sa provenance éditoriale, de fouille de textes, d'enrichissements (par exemple ajout de tables de synonymes). La plateforme, en agrégeant les ressources acquises, favorisera le travail transdisciplinaire : les chercheurs disposeront de l'ensemble des données, sans frontières de champs disciplinaires. Enfin, appuyée sur une architecture ouverte, elle permettra aux établissements d'enrichir leurs portails propres.

## **Les bénéfices pour la communauté universitaire**

**Avant** : chaque université devait acheter elle-même l'accès aux archives auprès d'éditeurs pratiquant des prix souvent prohibitifs, avec deux conséquences :

- **Un budget considérable** que les universités devaient dégager pour acquérir ces données indispensables aux chercheurs. en 2009 : 8,7 M€ étaient consacrées par les universités et organismes de recherches aux achats d'archives électroniques.
- **Une inégalité territoriale** entre les universités ayant les moyens d'accéder à ces données et les autres

Istex permet de **remédier à ces inégalités, tout en renforçant la productivité de la recherche française.**

Les **bénéfices pour les chercheurs** :

- Une **équité d'accès à la connaissance sur tout le territoire**
- Un **accès à une base de données exceptionnelle** : capacité d'achat estimé à environ **50 à 60 millions de documents numériques** (articles, revues, livres...)
- une **augmentation de la productivité de la recherche avec un moteur de recherche adapté aux besoins des scientifiques**, des facilités de fouille et d'extraction de données, de téléchargements...
- un **confort de navigation** : l'accès à des collections sous licence sera intégré à l'environnement local de chaque chercheur, intégrable aux autres ressources, donc accessible partout sans heurt ou rupture de navigation.
- un moyen de **valoriser les productions scientifiques** : les données pourront faire l'objet de requêtes fédérées (identiques sur l'ensemble du corpus, quelle que soit sa provenance éditoriale), de fouille de textes, d'enrichissements (annotation, travail sémantique, par exemple ajout de tables de synonymes)
- Une incitation au **travail transdisciplinaire** : les chercheurs disposeront de l'ensemble des données, sans frontières de champs disciplinaires ce qui est un facteur favorable à la créativité scientifique.

Les **bénéfices pour les étudiants** :

- Un **accès équitable à une documentation scientifique du plus haut niveau**
- Une plus grande **autonomie de travail**, grâce à un accès 24h/24 à la base de données.
- Un **enrichissement des pédagogies universitaires** grâce à la documentation électronique

Les **bénéfices pour les universités** :

- Une **réduction significative des dépenses** : Istex soulagera le budget des établissements des sommes consacrées à l'acquisition d'archives (en 2009 : 8,7 M€)
- Une **réduction des inégalités entre les universités dans l'accès à la connaissance**

## Organisation et gouvernance

Doté d'un budget de 60 M€, ce projet sera mis en œuvre en trois ans. Pour lancer le projet, une campagne d'information sera organisée à destination des communautés de recherche en mai 2012, pour permettre la remontée des besoins des utilisateurs, qui sera croisée avec l'analyse des offres des éditeurs.

Le porteur institutionnel du projet sera le CNRS, et le projet sera piloté par un comité exécutif composé des 4 partenaires : l'ABES, Agence bibliographique de l'enseignement supérieur, le Consortium COUPERIN et l'Université de Lorraine, agissant pour le compte de la CPU.

Le pilotage d'ISTEX sera assuré dans le cadre du programme Bibliothèque scientifique numérique (BSN) inscrite dans la SNRI.

En termes d'organisation, le projet ISTEX est découpé en plusieurs ensembles fonctionnels. Plusieurs opérateurs ont naturellement vocation à prendre en charge le pilotage de ces différents segments en raison des missions qu'ils assument actuellement dans le paysage de l'IST en France. Le comité exécutif ISTEX assure la maîtrise d'ouvrage de ces segments, dont il délègue la réalisation (maîtrise d'œuvre) aux opérateurs retenus :

- Le consortium Couperin (avec les organismes hors périmètre Couperin) pour le recueil des besoins des communautés de recherche, l'évaluation des ressources, l'établissement des listes,
- L'ABES pour la négociation et l'acquisition des ressources, leur signalement, la gestion des accès et des droits (en partenariat avec le CNRS-INIST et en relation avec les opérateurs locaux),
- Le CNRS, en s'appuyant sur l'INIST, pour la conception et l'hébergement de la plateforme d'agrégation du texte intégral : hébergement et exploitation des données, développement des services à valeur ajoutée, analyse de l'utilisation et des usages
- L'Université de Lorraine agissant pour le compte de la CPU pour la recherche et les services.

L'archivage pérenne des données sera assuré par le CINES (Centre informatique national de l'enseignement supérieur).

## Calendrier

- avril 2012 : signature de la convention ANR/Etat/CNRS
- avril 2012 : présentation du dispositif de remontée des besoins dans le cadre de la journée des BU et de l'IST
- mai 2012 : séminaire à destination des acteurs de la recherche pour organiser la remontée des besoins
- mai 2012 à mai 2013 : développement du prototype de plateforme
- de mai à juillet 2012 : remontée des besoins
- 3<sup>ème</sup> trimestre 2012 : signature des conventions entre le CNRS et les partenaires (Université de Lorraine, ABES et Couperin)
- juillet 2012 : validation des listes de ressources à acquérir la première année
- été 2012 : lancement des premières négociations avec les éditeurs
- fin de l'automne 2012 : signature des premières licences
- de mai 2013 à mai 2014 : développement de la plateforme et des services à valeur ajoutée
- 2013 et 2014 : achat des autres ressources selon le même dispositif : remontée, évaluation, sélection, négociation.