

Frédéric WEISS & Najib FAWZI

REI (23) – *La nouvelle base CTLF*
[1]

corpus de textes linguistiques fondamentaux

Notices Bibliographie Images Textes Articles

Recherche dans les articles de l'équipe CTLF
| Actualité du CTLF | Actualité de la base d'articles | Fonctionnement du portail |

titre

auteur

rubrique

accès réservé

archive
 non oui

présentation
 abrégée détaillée

par titres par rubriques par dates

1 Pour citer cet article :

WEISS Frédéric & Najib FAWZI. «REI (23) – La nouvelle base CTLF». *CTLF – Articles*. Colloque 'Refonte et extension internationale du CTLF: Corpus de textes linguistiques fondamentaux', Université Paris Diderot, UMR 7597, «Histoire des théories linguistiques», Programme Action structurante 2015-2018, 31 mai et 1er juin 2018. Mise en ligne le 01/10/2018 à l'adresse: <http://ctlf.ens-lyon.fr/documents/articles/colloque/REI-23.Ctlf-2.pdf>.

Nouvelle base CTF

Bref historique

Avant le CTLF

CRGTL Corpus représentatif des grammaires et des traditions linguistiques

Ensemble de notices descriptives paru sous forme papier :

- *Histoire Epistémologie Langage*, hors-série n° 2 (1998) et n° 3 (2000)
- Notices composées de 14 rubriques
- Deux tomes accompagnés d'une bibliographie

Bref historique

Projet CTLF (2003-2006)

CPER Contrat de plan État-Région

Partenaires :

- HTL (CNRS, U. Paris 7)
- CID (Centre d'Ingénierie Documentaire, ENS Lettres et Sc. humaines)
- MOM (Maison de l'Orient et de la Méditerranée, CNRS, U. Lyon 2)

Bilan :

- Architecture du portail CTLF mise en place
- Chaîne de traitement des fac-similés jGalith (MOM)
- 4 bases de données : Notices, Bibliographie, Images, Articles
- 130 volumes numérisés

Bref historique

Projet DETCOL (2007-2011)

ANR Agence nationale de la recherche

DETCOL Développement et exploitation textuelle d'un corpus d'œuvres linguistiques

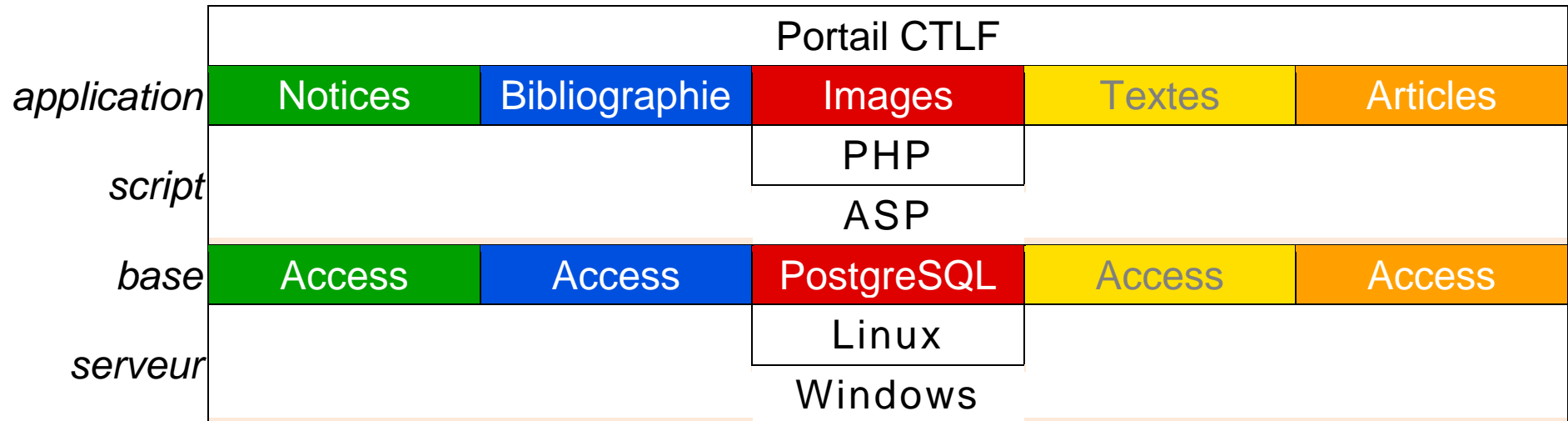
Partenaires : HTL, CID, ATILF (CNRS, U. de Lorraine)

Bilan :

- Enrichissement et amélioration des applications existantes
- Réalisation de l'application Textes : 600 textes structurés et corrigés
- Cahier des charges (XML) pour la base Frantext CTLF

Etat des lieux

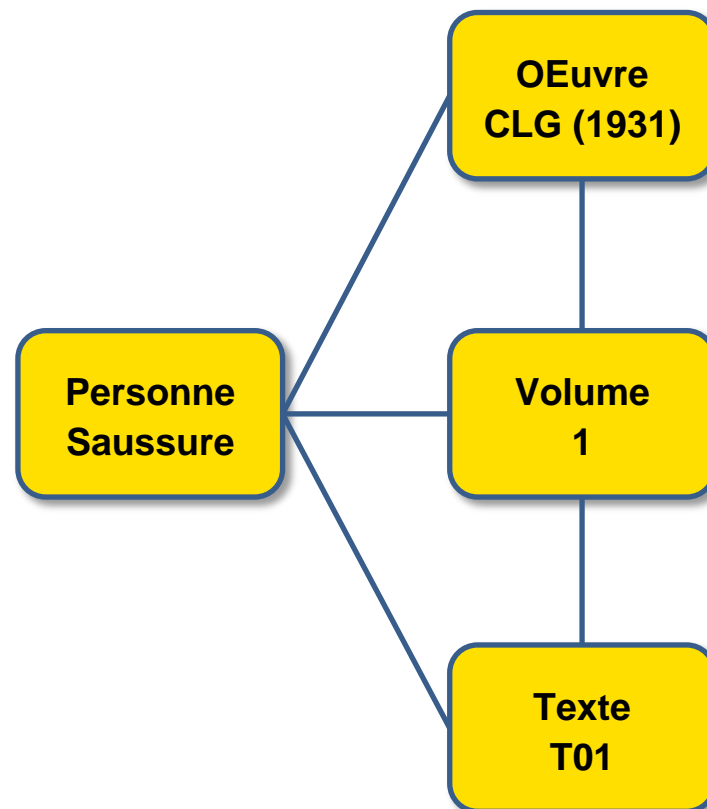
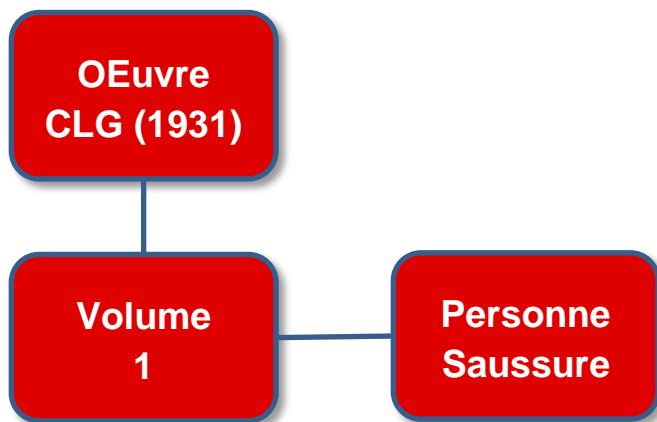
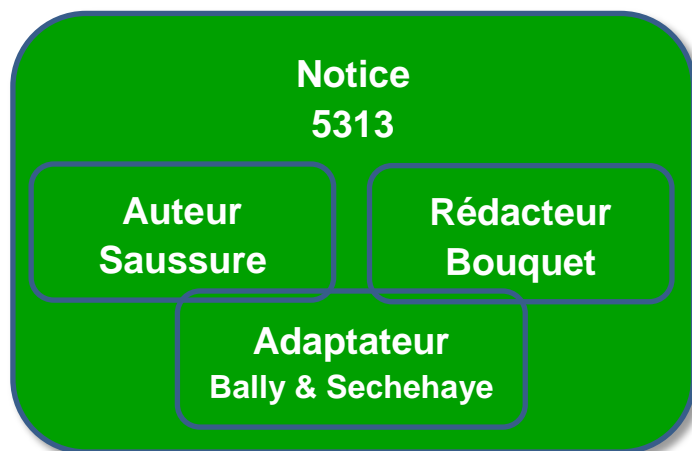
Structure générale du CTLF



5 applications, 5 bases de données

Etat des lieux

Gestion des personnes : exemple de Saussure



Redondance "Saussure" dans Notices (vert), Images (rouge), Textes (jaune)

Etat des lieux

Conclusion

Disparition programmée des serveurs Windows à l'ENS de Lyon

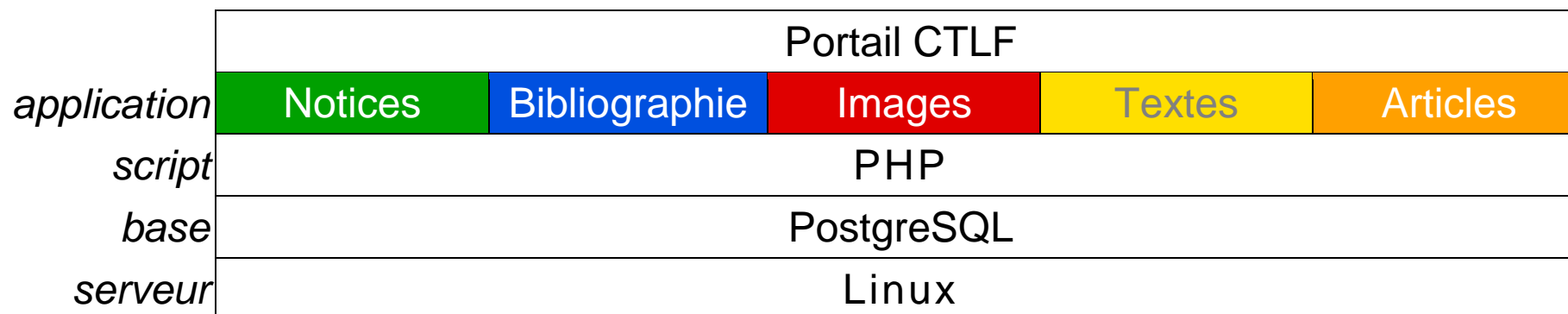
Technologies hétérogènes compliquant inutilement la maintenance et les évolutions

Cinq bases de données :

- Redondance d'information
- Risque d'erreurs de saisie
- Complexité au niveau de la gestion

Nouvelle base unifiée

Structure générale simplifiée



5 applications, 1 base de données

Nouvelle base unifiée

Gestion des personnes : exemple de Saussure

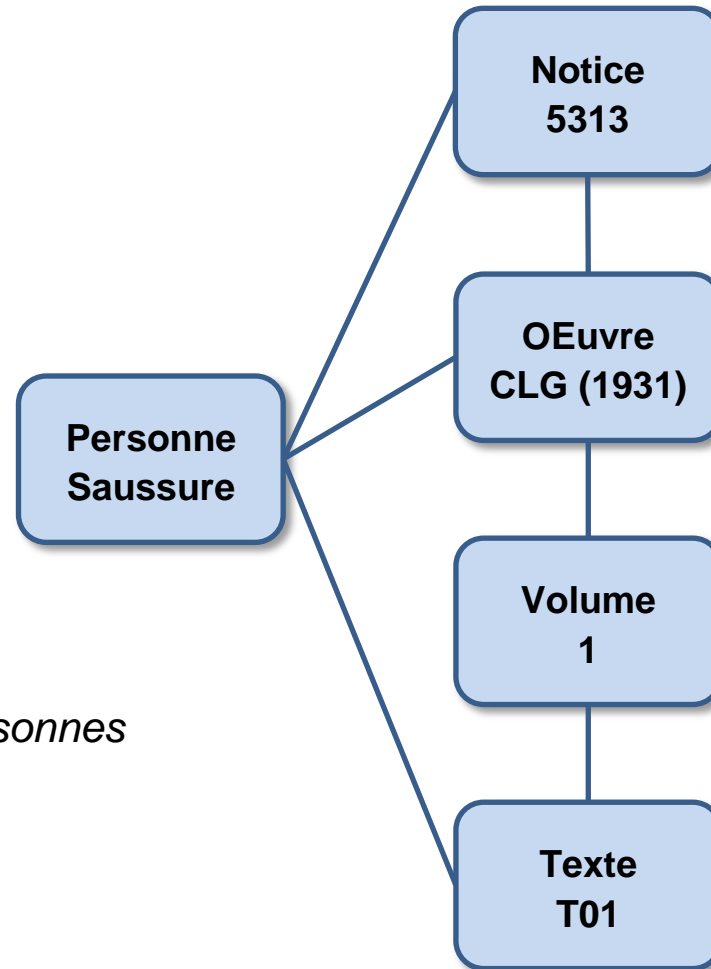


Table unique pour les personnes

Nouvelle base unifiée

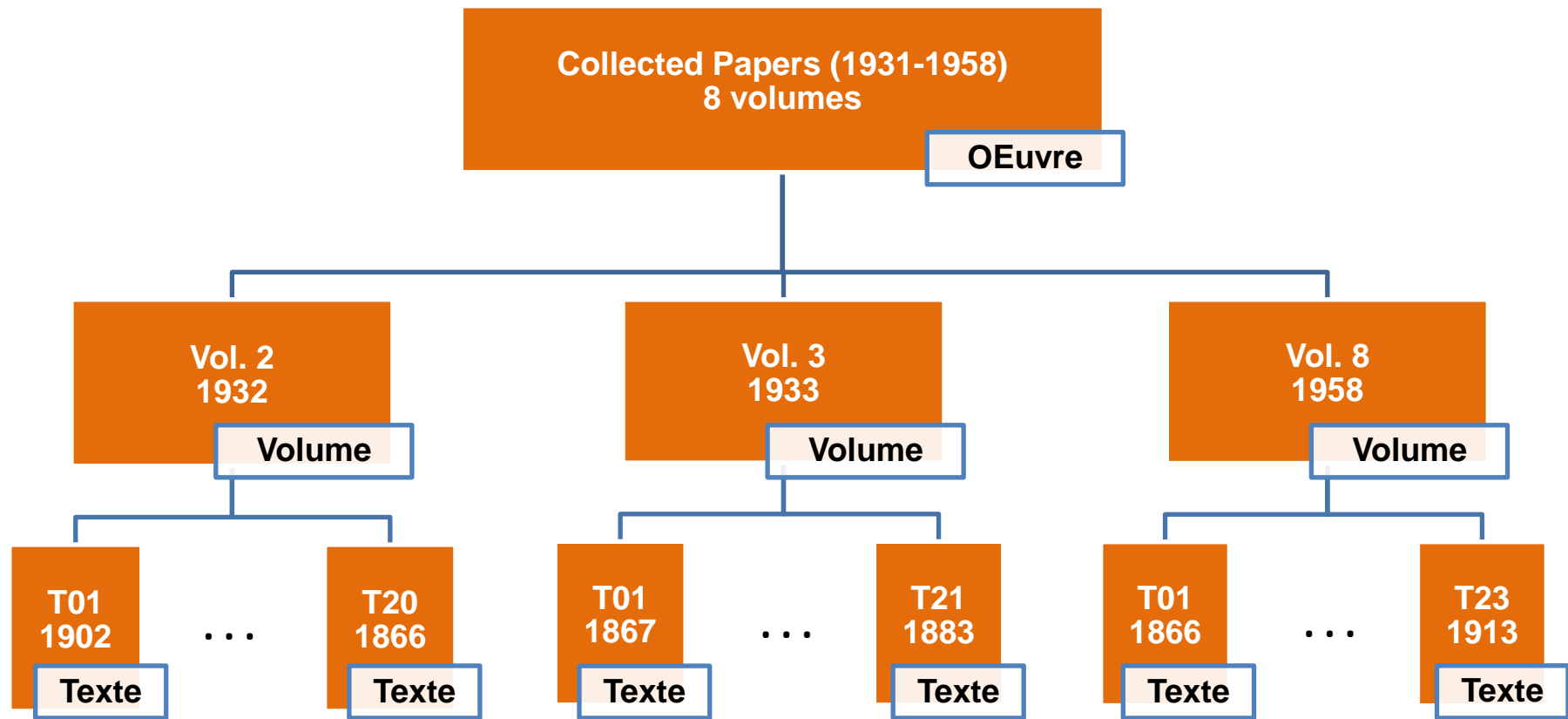
Gestion des œuvres numérisées : exemple de Peirce (5305)

CTLF actuel

- Informations gérées dans les bases Images et Textes
- Redondance (titre et date d'édition d'un volume)
- Vue "atomisée" de l'œuvre :
liste de volumes (appli. Images), liste de textes (appli. Textes)

CTLF unifié

- une 1^e table gère les œuvres (en un ou plusieurs volumes)
- une 2^e table reliée à la première gère les volumes
- une 3^e table reliée à la seconde gère les textes



Perspectives

Rappel des objectifs de l'unification

- Assurer la pérennité technologique du CTLF
- Réduire les risques d'erreur et optimiser les procédures
- Privilégier les technologies en usage dans le monde universitaire

Développements ultérieurs envisageables

- CTLF bilingue
- Affichage coordonné entre fac-similés, textes et table de matières
- Exploitation de l'enrichissement des textes