

**Journées nationales  
Bibliothèques Numériques de Référence  
1<sup>er</sup> décembre 2016 - Grenoble**

**Expérimentation sur  
l'accessibilité de la presse ancienne**



# Présentation de l'Arald

- Agence Rhône-Alpes pour le livre et la documentation
- Une structure régionale pour le livre (SRL)
- Au carrefour des différents métiers du livre et de leurs acteurs : auteurs, éditeurs, libraires, bibliothécaires, archivistes, etc.
- Des actions de coopération, d'information, de formation, de conseil et de promotion en faveur du livre et de la lecture
- Une association financée par la Région Auvergne – Rhône-Alpes et le Ministère de la Culture et de la Communication / DRAC Auvergne – Rhône-Alpes

# L'Arald et le patrimoine écrit

- Une campagne de sauvegarde et de valorisation de la presse ancienne (1996 – 2012)
  - > 54 titres microfilmés et/ou numérisés (+ d'1 million de fichiers)
- Une mise en ligne sur le portail *Mémoire et actualité en Rhône-Alpes* depuis 2009
  - > 520 000 pages interrogeables en texte intégral
- Des projets soutenus par le Service du Livre et de la Lecture (SLL), Ministère de la Culture et de la Communication
  - > Appels à projet Patrimoine écrit en 2007, 2011, 2015 et 2016

# Le projet

- **Une expérimentation innovante** pour ouvrir le patrimoine écrit à de nouvelles formes de lecture :
  - > plus accessibles
  - > pour tous les utilisateurs
  - > pour répondre à des besoins de recherche en perpétuelle évolution
- **Publics ciblés :**
  - > personnes porteuses d'un handicap visuel
  - > mais aussi : chercheurs, généalogistes, scolaires, grand public
- **Pour rendre accessible la presse ancienne il faut :**
  - > un OCR en qualité éditoriale (99,99%) → **reprise manuelle**
  - > une structuration fine de la page : détection géographique des blocs et identification du type de contenu (article, illustration, feuilleton, publicité, etc.) → structuration au format **XML ALTO / METS**
  - > un format pivot accessible → **EPUB 3 REFLOW**

# Étape 1 – Conception du projet (1)

## Le corpus

- Un corpus test :
  - > 10 journaux de Rhône-Alpes + 1 journal d'Auvergne
  - > 39 numéros
  - > 152 pages
  - > grande variété des titres : contenu, structuration, format, etc.
- Un corpus de masse :
  - > 2 journaux de Savoie
  - > *Les Alpes Illustrées*
  - > *Les Alpes Pittoresques*
  - > 6500 pages
- Établissements associés :
  - > centres d'archives départementaux ou municipaux
  - > bibliothèques municipales

# Étape 1 – Conception du projet (2)

## L'équipe projet

- A l'Arald : 2 personnes en charge du projet
- Un groupe de travail : 3 réunions en novembre, décembre 2015 et janvier 2016 / 15 personnes
  - > responsables de fonds patrimoniaux
  - > experts de la presse ancienne
  - > experts en numérisation
  - > experts des questions d'accessibilité
- Un comité d'experts associés : consultation à distance / 9 personnes
  - > experts institutionnels
  - > responsables de projets numériques sur la presse ancienne
  - > responsables de fonds

# Étape 2 – Les opérations de retraitement des fichiers numériques

- Rédaction d'un cahier des charges technique (Arald)
- Lancement d'une consultation auprès de prestataires techniques

## Principaux problèmes rencontrés :

- Peu de prestataires en mesure de répondre à la consultation : nouveaux développements à créer + outils à imaginer pour une production de masse
- Coûts élevés : nouveaux développements à créer + reprise manuelle nécessaire pour l'OCR et la structuration

# Étape 3 – Valorisation

- Mise en ligne des fichiers retraités sur le nouveau portail *Lectura+*, en cours de développement par l'Arald
- Publication de recommandations techniques pour les professionnels du patrimoine autour de la mise en accessibilité de la presse ancienne
- Valorisation en direction du grand public avec des ateliers de présentation des nouveaux outils disponibles auprès des personnes en situation de handicap et des associations liées
- Un projet soutenu par :
  - > la DRAC Auvergne – Rhône-Alpes
  - > le SLL/MCC dans le cadre de l'Appel à projet national Patrimoine écrit



# Merci pour votre attention !

**Contact :**

**Edith GOUDIN**

e.goudin@arald.org - 04 72 00 00 34

**Priscille LEGROS**

p.legros@arald.org - 04 72 00 00 35

**Arald**

[www.arald.org](http://www.arald.org)

[www.memoireetactualite.org](http://www.memoireetactualite.org)