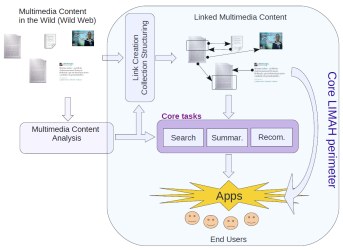
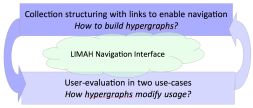
EXPLORATION D'ARCHIVES MULTIMÉDIAS : QUELLES ATTENTES ?

Le projet LIMAH : en texte ...

Le projet *Linking Media in Acceptable Hypergraphs* (LIMAH) a pour objectif applicatif de **proposer une interface d'exploration d'archives multimédias**. Le paradigme d'exploration s'appuie sur la **création automatique de liens entre (fragments de) documents** à partir de l'analyse des contenus. Ces liens reflètent une proximité sémantique : contenu similaire, proximité thématique, évolution chronologique d'un événement, opinion exprimée, réaction, etc. L'analyse des contenus, ainsi que la création des liens et leur typage, permettent de proposer de **nouvelles fonctionnalités** pour explorer une collection documentaire ou pour rechercher des informations. **Ces fonctionnalités sont-elles pertinentes et/ou utiles pour des professionnels de l'information ?** Quels liens peut-on créer automatiquement ? Quels types de liens amènent une valeur ajoutée ? En quoi ces nouvelles fonctionnalités changent-elles les pratiques ? Telles sont quelques-unes des interrogations auxquelles LIMAH cherche à répondre.

... et en images





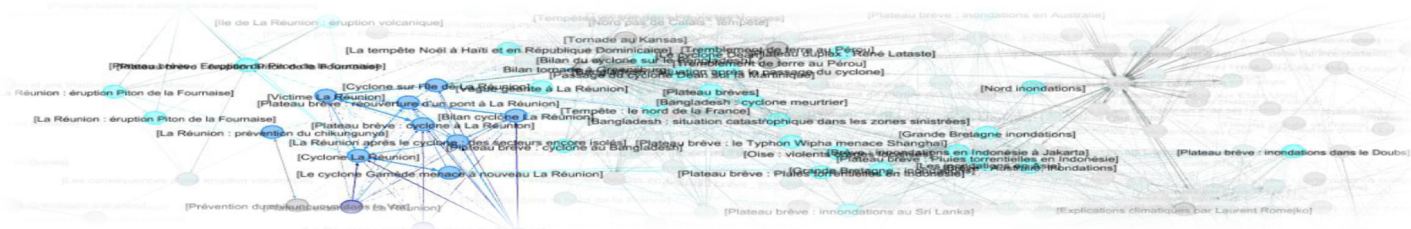
Données

Données relatives à l'actualité

- Plusieurs sources récupérées sur Internet
  - TV vidéos : news, débats, *etc.*
  - Journaux en ligne
  - Podcasts radio
  - Tweets et Facebook
- Période : 3 semaines, automne 2015
- Variété de contenus : modalités multiples, des news journalières de 5 min. à des débats de 2h
- Données + métadonnées + hyperliens

Données relatives à la e-éducation

- Quelques cours enregistrés
  - Vidéos, livres, transparents, Wikipédia, etc.
  - Étude de la segmentation pour l'apprentissage



Scénarios

1. Rechercher une information : Qui a dit « Moi Président » ? Quand a commencé le conflit israélo-palestinien ?
2. Acquérir une vue globale sur un sujet, e.g., pour rédiger une synthèse : dates de début et fin plus chronologie, faits principaux, causes et conséquences, personnes/organismes impliqués, opinion publique et critiques principales, *etc.*

Méthodologie

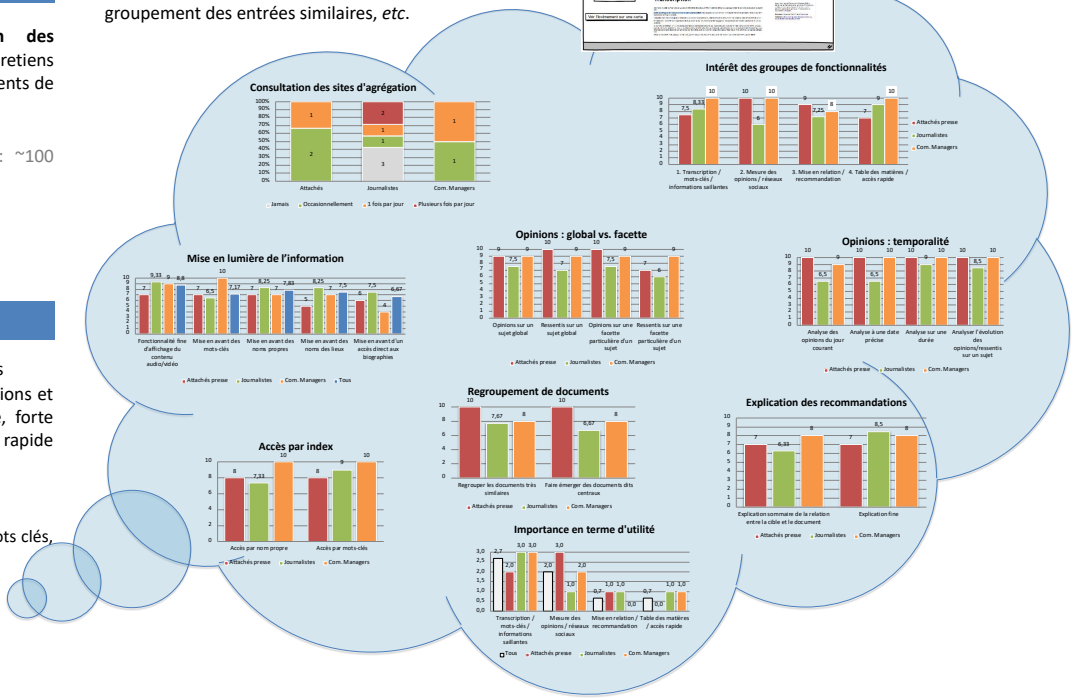
1. **Interviews s'appuyant sur la description des fonctionnalités à l'aide de *mockups*** : 13 entretiens individuels de 1h30 avec des journalistes, des agents de presse et des *community managers*
   - *compréhension des pratiques et des attentes*
2. Enquêtes en ligne concernant l'acceptabilité : ~100 participants
   - *définir les fonctionnalités attendues*
3. Tests d'usage de l'interface d'exploration LIMAH
   - *valider et améliorer les fonctionnalités*

Quelques tendances

- Intérêt marqué pour les fonctionnalités proposées
- Variations selon les profils : importance des opinions et des réseaux sociaux pour les agents de presse, forte valeur ajoutée de la transcription et de l'accès rapide pour les journalistes
- Suggestion de nouvelles fonctionnalités
  - Classement et organisation des mots clés
  - Filtres intelligents pour les commentaires, les mots clés, les sources, *etc.*
  - Tracabilité des sources, fiabilité et popularité
  - Propositions de synthèses automatiques

Fonctionnalités proposées

1. **Mise en lumière de l'information** -> transcription, mots-clés, entités nommées, lien vers des données externes (biographies, définitions, *etc.*)
2. **Mise en lumière de l'opinion** -> opinions et sentiments sur les réseaux sociaux, globalement ou par aspect, évolution temporelle, synthèse, *etc.*
3. **Accès rapide** -> table des matières, segmentation thématique, organisation chronologique, accès par entités nommées ou mots clés, résumé, *etc.*
4. **Liens et recommandations** -> mise en évidence de liens entre contenus, typage et explication des liens, groupement des entrées similaires, *etc.*



Ce travail a bénéficié d'une aide de l'État attribuée au labex COMIN LABS et gérée par l'Agence Nationale de la Recherche au titre du programme « Investissements d'avenir » portant la référence ANR-10-LABX-07-01