



HAL
open science

VALIDER UN MODÈLE DE DONNÉES AVEC UN OUTIL D'ANNOTATION : LE CAS DU PROJET READ-IT

François Vignale, Julian Gonggrijp, Alex Hebing, Quentin Morcrette

► **To cite this version:**

François Vignale, Julian Gonggrijp, Alex Hebing, Quentin Morcrette. VALIDER UN MODÈLE DE DONNÉES AVEC UN OUTIL D'ANNOTATION : LE CAS DU PROJET READ-IT. *Humanistica* 2020, May 2020, Bordeaux, France. hal-02877907

HAL Id: hal-02877907

<https://univ-lemans.hal.science/hal-02877907>

Submitted on 22 Jun 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

VALIDER UN MODÈLE DE DONNÉES AVEC UN OUTIL D'ANNOTATION : LE CAS DU PROJET READ-IT.

François VIGNALE¹, Julian GONGGRIJP², Alex HEBING², Quentin MORCRETTE¹

¹3L.AM, Le Mans Université (France) ; ²DH Lab, Utrecht Universiteit (Pays-Bas)

Au cours des dernières décennies, les connaissances sur l'histoire de la lecture (pratiques, usages, habitudes) ont considérablement augmenté, mais des questions fondamentales demeurent comme « pourquoi » (motivations, dispositions ...) et « comment » (circonstances, environnement, position de lecture ...) on lit. En explorant des expériences de lecture contenues dans les sources numérisées, le projet READ-IT (Reading Europe Advanced Data Investigation Tool) vise à apporter des réponses à ces questions.

En combinant différentes conceptions (Jauss 1982 ; Iser 1978), nous avons défini un modèle où l'expérience de lecture est vue comme un phénomène temporel précédé de prémisses et suivi d'effets dans lesquels une personne interagit avec un contenu écrit à travers un support (Vignale et al. 2019). Cette construction théorique, fondée sur des exemples, a été validée par des spécialistes du domaine à l'aide d'une plateforme d'annotation développée pour le projet.

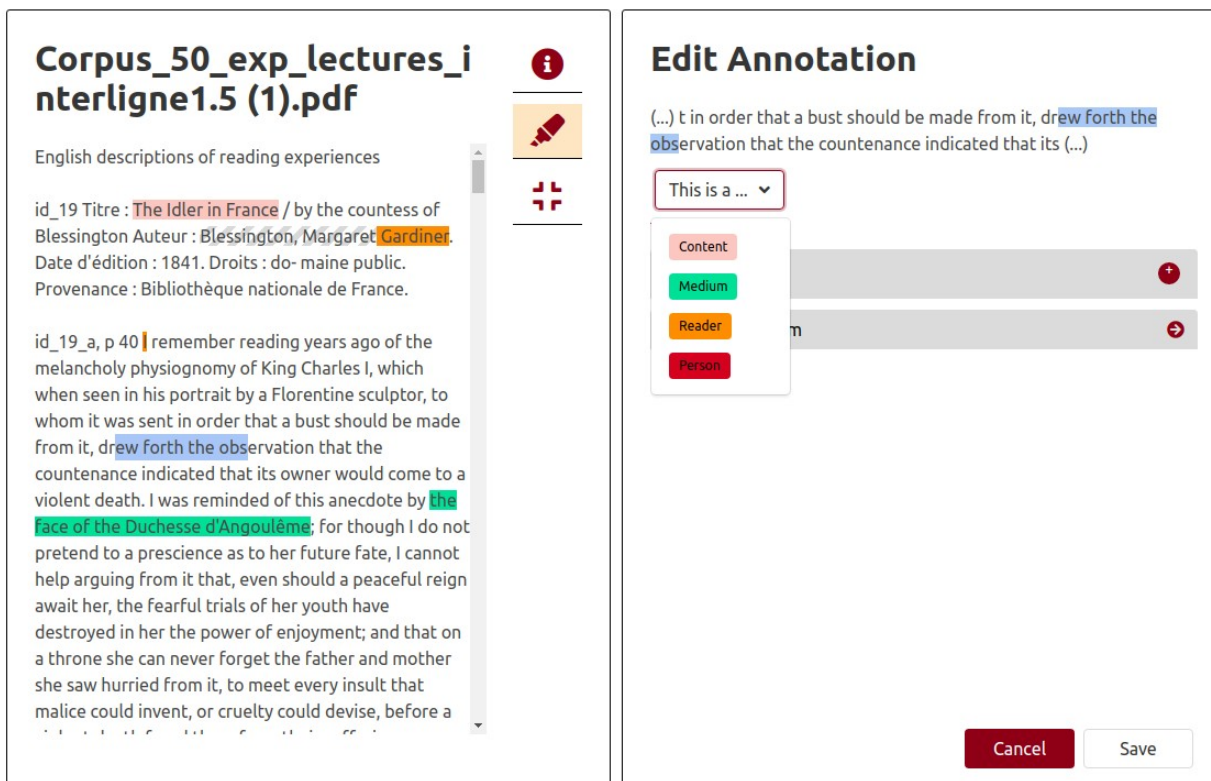
Des textes provenant de diverses sources (correspondance, commentaires en ligne, journaux intimes, etc.) ont été annotés à l'aide d'un guide afin de valider les concepts de base en comparant la similarité ou non des annotations. Les termes sélectionnés sont utilisés pour vérifier la pertinence des concepts, pour améliorer les performances de l'ontologie mais aussi pour fournir un vocabulaire structuré pour dans différents thésaurus.

Nous avons choisi de créer un outil d'annotation personnalisé permettant l'intégration de n'importe quelle ontologie dans l'outil, tant qu'elle est conforme à RDF. Aucun outil d'annotation existant n'offre une telle flexibilité dans la personnalisation des catégories et des relations disponibles pour les annotateurs au moyen de RDF.

Techniquement, le backend est basé sur le standard Open Annotation (oa). L'un des défis est de traduire les sélections de texte en oa:Selectors qui sont également reproductibles en cas de surlignage multiples. Ainsi, les détails de la sélection (nœud de début et de fin et indices de début/fin) pour chaque annotation sont stockés à la fois comme oa:TextQuoteSelector et deux oa:XpathSelectors (dans lesquels les indices mentionnés sont stockés comme une requête de sous-chaîne, par exemple : 'substring(//*[3]/text(), 15)').

Du côté client, l'interface d'annotation est développée en collaboration avec l'équipe du projet. A chaque réunion du projet, les développeurs invitent les chercheurs de l'équipe à concevoir une partie de l'application et à proposer des scénarios d'utilisation qui traduisent leur vision de l'interface. De ces sessions est née la conception actuelle de l'outil, où l'outil d'annotation est constitué du texte à gauche et des outils nécessaires pour annoter le texte à côté à droite, avec éventuellement un troisième " panneau " avec des outils insérés à droite. L'outil d'annotation est entièrement intégré dans le mode explorateur de l'interface utilisateur, de sorte que les utilisateurs peuvent modifier ou ajouter des annotations tout en explorant la base de données liée.

Comme l'outil d'annotation est complètement intégré avec l'explorateur de données de READ-IT, il ne sera pas ouvert et libre au moins pour le moment. C'est une suggestion intéressante, cependant, et nous la prendrons en considération le moment venu.



The screenshot displays the READ-IT Explorer interface. On the left, a document viewer shows the title "Corpus_50_exp_lectures_1terligne1.5 (1).pdf" and a list of English descriptions of reading experiences. The main text area shows a passage from "The Idler in France" by Margaret Gardiner, with several words highlighted in blue and green. On the right, an "Edit Annotation" dialog box is open, showing a dropdown menu with "This is a ..." and a list of categories: Content, Medium, Reader, and Person. The dialog also includes "Cancel" and "Save" buttons.

RÉFÉRENCES

- Iser, Wolfgang. 1978. *The act of reading: a theory of aesthetic response*. London, Royaume-Uni: Routledge and Kegan Paul.
- Jauss, Hans Robert. 1982. *Ästhetische Erfahrung und literarische Hermeneutik*. Frankfurt am Main, Allemagne: Suhrkamp.
- Vignale, François, Guillaume Gravier, Brigitte Ouvry-Vial, et Alessio Antonini. 2019. « The Model of Reading Modelling Principles, Definitions, Schema, Alignments ». <https://hal-univ-lemans.archives-ouvertes.fr/hal-02301611>.