

Appel à projet 2014 du Labex *Les Passés dans le présent*

Projet « PROFIL »

Partage, Reconstitution et Organisation de Fictions Identitaires en Ligne

Titre du projet :

PROFIL (Partage, Reconstitution et Organisation de Fictions Identitaires en Ligne)

Axe thématique du labex concerné :

Relations au passé, Mémoire collective, patrimonialisation et appropriations sociales

Institution responsable :

Laboratoire Dicen-IDF (Axe "Traçabilité, identités et mémoires numériques") Université Paris Ouest

Responsable du projet

nom & prénom : Louise MERZEAU

statut : MCF HDR en sciences de l'information de de la communication, directrice de Tactic

organisation : laboratoire Tactic / Dicen-IDF

courriel : louise@merzeau.net

tel fixe : +33 (0)1 42 78 38 09

tel portable : +33(0) 6 48 07 59 08

site internet : <http://merzeau.net>

Coreponsable(s) du projet (le cas échéant)

nom & prénom : Nicolas SAURET

statut : Responsable projets média et éditorialisation

organisation : IRI, Institut de Recherche et d'Innovation, Centre Pompidou, Paris

courriel : nicolas.sauret@iri.centrepompidou.fr tel fixe : +33 (0)1 83 87 63 25

tel portable :

site internet : <http://www.iri.centrepompidou.fr/pied/credits/>

Composition du comité de projet

- Lionel BARBE, MCF en sciences de l'information de de la communication, Paris Ouest Nanterre, Dicen-IDF
- Gilles BERTIN, Ingénieur Cnam, Dicen-IDF
- Haud GUEGUEN, MCF en philosophie, Cnam, Dicen-IDF

Yannick MAIGNIEN, ancien Directeur du TGE ADONIS, Responsable de la création de la plateforme ISIDORE, professeur de philosophie, chercheur associé Dicen-IDF

- Camille PALOQUE-BERGES, Ingénieure de recherche au laboratoire HT2S, chercheuse associée à Dicen-IDF
- Nicolas SAURET, Chef de Projet - IRI, Institut de Recherche et d'Innovation, Centre Pompidou
- Cheikh BA, doctorant ED139 université Paris Ouest Nanterre, Dicen-IDF
- Philippe RYGIEL, Professeur d'histoire à Paris Ouest Nanterre
- Moustapha MBENGUE, Enseignant-chercheur Ph D. à l'École des Bibliothécaires, Archivistes et Documentalistes (EBAD) de l'université Cheikh Anta Diop de Dakar

- Sylvia FREDRIKSSON, designer, enseignante en Design graphique (BTS Lycée Jacques Prévert, Boulogne-Billancourt, La Fonderie de l'image) et en Sciences et Techniques de l'Exposition (université Paris 1 Sorbonne).

La composition du comité, qui associe des chercheurs en sciences humaines (SIC, philosophie, histoire) à des ingénieurs et designers spécialistes du Web et des écritures multimédia, reflète notre volonté d'aborder la problématique de la mémoire et des traces par une approche transdisciplinaire se réclamant des humanités numériques.

I. Le projet dans le labex

1. Résumé

Le projet consiste à concevoir un dispositif de construction de profils à partir de traces individuelles extraites de ressources conservées dans des collections patrimoniales numérisées ou "nativement" numériques. Les profils seront construits par extraction de données, enrichissements documentaires et prolongements fictionnels, selon des modalités empruntant à la fois au Web social (*crowdsourcing*) et au Web des données (*datamining*), en appliquant des techniques de narration documentaires et fictionnelles.

La finalité d'un tel projet est triple :

- revisiter le concept de profil tel que l'ont modélisé les industries numériques de l'identité. Entièrement anonymisé, le profil sera envisagé non comme l'empreinte enregistrée d'un individu, mais comme un écosystème transmédiatique résultant d'une éditorialisation collective.
- favoriser l'appropriation sociale du passé, en explorant les possibilités de redocumentarisation de fonds d'archives, dans une démarche impliquant chercheurs en sciences humaines, archivistes et documentalistes, auteurs et amateurs.
- expérimenter de nouveaux modèles de collecte et de constitution de corpus, adaptés à la granularité et la mobilité de l'environnement numérique.

2. Thématiques et objectifs

Interroger la présence numérique à l'aune de l'histoire

Le projet s'appuie sur les recherches menées au laboratoire Dicen-IDF, visant à mettre en évidence les logiques de reconfiguration des identités à l'heure du Web social et du Web des données. Combinant des réflexions sur les enjeux éthiques, politiques, économiques, cognitifs et sociétaux des nouvelles modalités de présence en ligne, ces travaux portent sur des problématiques qui traversent tous les milieux (académiques, professionnels, politiques) et qui atteignent aujourd'hui le grand public : ouverture ou protection des données personnelles, indexation des individus et droit à l'oubli, vie en ligne et nouvelles logiques affinitaires, etc.

Pour faire face à de tels défis, il nous paraît indispensable de replacer ces questionnements dans une perspective longue, afin de confronter les modalités contemporaines de profilage et de traitement algorithmique des personnes aux systèmes antérieurs de fichage, d'identification et d'archivage des individus. La critique nécessaire de l'industrialisation des données personnelles ne peut en effet se dispenser d'un retour sur la notion même de trace, et d'une interrogation sur la diversité des manières de *fabriquer de l'identité*.

Repenser la mémoire numérique comme expérience transmédia

Après avoir été longtemps éludée au profit des seuls aspects ethnographiques ou psychologiques, la dimension technique ou industrielle de la mémoire doit être aujourd'hui interrogée. Les mécanismes inédits de traçabilité, le brouillage des frontières entre public et privé et le *Big Data* n'induisent pas qu'un

changement d'échelle : en bouleversant les modalités de conservation, de gestion, d'accès, de partage ou d'effacement de nos traces, ils affectent la mémoire elle-même.

La pensée de l'environnement numérique doit donc s'affranchir de la vision court-termiste, centrée sur les seuls effets d'immédiateté, d'instantanéité, de vitesse ou d'usages tendanciers. C'est l'un des apports des humanités numériques, que de permettre cette réhabilitation des enjeux mémoriels de la vie connectée.

Après avoir salué la croissance illimitée des capacités d'enregistrement et de conservation, la société souffre d'une surcharge informationnelle et se rend compte que le stockage des données *ne fait pas mémoire*, et même, qu'il la menace peut-être. Accaparée par les logiques de surveillance et de marketing, l'administration des données profilaires produit une mémoire "par défaut", machinique, verrouillée et hors projet collectif.

Réinvestir *culturellement* cette mémoire numérique, en articulant traces individuelles et dépôts collectifs, est donc un enjeu majeur. Non pas par des politiques coercitives étrangères aux usages, mais par une médiation qui favorise les pratiques *bottom-up* de collecte et d'exploitation des données, et les protège des stratégies marchandes.

Relevant moins du stockage que du *savoir copier-agréger-relier-inventer*, l'écriture des identités est une expérience transmédiatique, encore trop souvent dépendante de systèmes de boîte-noire qui la privent de toute maîtrise et la séparent de son histoire. Notre projet vise à restaurer cette culture mémorielle en faisant du profil non plus un ADN automatiquement capturé, mais un écosystème non programmé de données, de documents et de liens, ancré dans le patrimoine et partageable.

Réenchanter la mémoire numérique par la fiction

Dans la suite des travaux conduits par [l'atelier VériFiction](#) accueilli par notre laboratoire, notre projet renoue avec la question historiographique du rapport entre histoire et récit. Après le mythe, l'épopée, le théâtre, le roman et le cinéma, nous faisons l'hypothèse qu'il revient au transmédia de prendre en charge la mise en fiction de la mémoire collective.

Cela suppose que la pratique des espaces numériques puisse être "réenchantée" par une forme de présence qui résiste à la calculabilité des identités. Notre proposition vise à réhabiliter cette part d'invention intrinsèque à toute pratique mémorielle, en ménageant dans le Web des données un espacement où l'exploitation des data soit affranchie des objectifs de rationalisation capitalistique.

Appréhendée comme capacité à soutenir une activité narrative dans laquelle sont mises en œuvre des opérations de réécriture, de montage, d'oubli et de création, la présence numérique n'est plus réductible à une somme de traces : elle devient elle-même une *re-connaissance* active du passé, dans laquelle documentation et fiction s'enrichissent l'une l'autre. Entre "arts de faire", activité documentaire et exploitation de données, le mode de traitement des traces que nous proposons ne doit pas être confondu avec la reconstitution ou la simulation historique telle qu'elle se traduit dans le docufiction. L'objectif est ici de montrer les effets miroirs et récursifs du calcul et du traitement éditorial, en questionnant le statut d'attestation des marqueurs identifiants et en redonnant une place aux incertitudes que la raison algorithmique exclut.

Expérimenter de nouvelles logiques d'archivage

À une pensée du support (où le document se définit par ses contours matériels), nous proposons de substituer une pensée de l'environnement transmédiatique (où le contenu se recompose au gré des redocumentarisations). Ce changement de paradigme rend les techniques traditionnelles de collecte et d'archivage en partie caduques. Le modèle du catalogue, hérité de la bibliothèque, et celui du prélèvement d'échantillons, hérité de l'archivistique, s'avèrent mal adaptés à l'architecture et la mobilité des informations en ligne. La construction d'un patrimoine numérique – *par des méthodes et des outils numériques* – nécessite donc d'imaginer de nouveaux modèles.

L'expérimentation que nous proposons sur les profils peut en ce sens aider à explorer un plusieurs pistes :

- envisager la séquence données-document-métadonnées non comme une chaîne successive, mais comme un cycle récursif où la ressource est constamment enrichie ;

- croiser des sources hétérogènes pour documenter une ressource et conserver cette hétérogénéité dans le dispositif de restitution ;
- exploiter la sémantisation des données pour fabriquer des "fictions heuristiques" permettant de contextualiser une ressource ;
- faciliter la navigation dans l'archive par des systèmes d'hypertagging participatif (voir les travaux sur les herbiers numériques menés par Dicen-IDF dans le cadre de l'infrastructure nationale E-Recolnat) ;
- travailler sur la plasticité de l'archive, en adaptant la granularité de la collecte au flux ;
- organiser une navigation temporelle dans l'archive, pour restituer l'instabilité des contenus en ligne (cf. modèle du dépôt légal du Web à l'Ina).

REFERENCES BIBLIOGRAPHIQUES

- Ertzscheid O., Gallezot G., Simonnot B. (2013), "À la recherche de la « mémoire » du web : sédiments, traces et temporalités des documents en ligne", in C. Barats (dir.), *Manuel d'analyse du web*, Armand Colin, p53-68.
- Kaplan F. (2013), "Lancement de la *Venise Time Machine*", [en ligne] <http://fkaplan.wordpress.com/2013/03/14/>
- Maignien Y. (2004), "Vérité et fiction sur Internet. Le statut des documents publiés en ligne" in J-M. Salaün et C. Vandendorpe (dir.), *Les défis de l'édition numérique*, Presses de l'ENSIB, Lyon.
- Merzeau, L. (2013), "L'Intelligence des traces", *Intellectica*, n°59, p.115-135
- Merzeau, L. (2012). "Faire mémoire des traces numériques", *E-dossiers de l'audiovisuel* [en ligne] <http://www.ina-expert.com/e-dossier-de-l-audiovisuel-sciences-humaines-et-sociales-et-patrimoine-numerique/faire-memoire-des-traces-numeriques.html>
- *Mémoires & Internet*, MEI 32 (2011), sous la dir. de Pignier Nicole et Lavigne Michel
- Rouvroy A. et Berns Th. (2010). "Le nouveau pouvoir statistique", *Multitudes* 1/2010 (n° 40), p. 88-103.
- *Traçabilité et réseaux*, *Hermès* 53 (2009), sous la direction de M. Arnaud et L. Merzeau
- Zacklad, M. (2007). "Réseaux et communautés d'imaginaire documédiatisées", Peter Lang (Ed.) p. 279-297 [en ligne] http://archivesic.ccsd.cnrs.fr/sic_00180185

3. Durée du projet

12 mois

4. Attendus du projet par rapport au labex *Les passés dans le présent*

Le projet PROFIL contribuera aux recherches du labex à divers titres :

- par sa visée, il enrichira la réflexion interdisciplinaire sur la présence du passé dans la société contemporaine, en interrogeant les mutations de la mémoire induites par les mécanismes de la traçabilité numérique et les nouvelles modalités de la présence en ligne ;
- par sa méthodologie, il introduira des formes d'expérimentation innovantes, reposant sur la conception de dispositifs d'écriture et d'observation transmédiatique et sur le recours aux modalités participatives de l'élaboration et de la diffusion des savoirs ;
- par ses résultats, il apportera aux partenaires du labex des préconisations, des études de cas et des modélisations permettant d'améliorer les méthodes d'archivage et de navigation dans des corpus numériques hétérogènes.

5. Partenaires du labex *Les passés dans le présent* impliqués ou à impliquer

- Bibliothèque de Documentation Internationale Contemporaine (BDIC)

Interlocuteurs : Valérie Tesnière (directrice), Lionel Maurel (conservateur)

La BDIC fournira le matériau documentaire à partir desquels les profils seront élaborés. Elle orientera l'échantillonnage en fonction de ses propres attentes et du recensement des fonds susceptibles d'être convoqués pour l'enrichissement des profils. Elle remplira enfin une fonction de conseil pour l'anonymisation des traces.

Si les attendus du projet ont déjà fait l'objet d'une concertation, le mode opératoire de la collaboration avec la BDIC reste à préciser et à planifier.

- **BnF : dépôt légal du Web et Gallica**

Interlocuteur : Clément Oury (chef du service)

À ce stade de préfiguration, le partenariat avec la BnF est encore en cours de négociation. Au vu des contraintes juridiques attachées au statut de dépôt légal de ses collections, le partenariat devrait moins porter sur la réutilisation de ressources que sur la réflexion à mener en amont et en aval de la constitution d'un web archivé. Des ateliers méthodologiques et des enquêtes d'usage, notamment, pourraient être envisagés.

Les modèles d'architecture et de navigation élaborés dans le cadre de PROFIL pourraient par ailleurs être testés sur des corpus fermés de la BnF, à des fins de validation d'hypothèses. Sur la question des écritures identitaires, notre projet pourrait enfin proposer un retour critique sur l'expérience de la BnF en matière d'[archivage des journaux personnels en ligne](#), en croisant le modèle de la bibliothèque avec celui de l'archive et du flux.

6. Partenaire(s) extérieur(s) impliqués ou à impliquer

- **Institut de Recherche et d'Innovation (IRI)**

Interlocuteurs : Vincent Puig (directeur exécutif), Nicolas Sauret

L'IRI interviendra sur la conception des scénarios et des dynamiques de contributions en s'appuyant sur ses travaux récents sur l'éditorialisation d'archive journalistique (ANR Périplus) et sur les dispositifs d'indexation et d'annotation collaborative de l'image (projet Egonomy), de la vidéo (projet Cinecast) et le texte (partenariat avec la bibliothèque de l'Université de Tokyo). Plusieurs outils orientés indexation, annotation collaborative, agrégation et éditorialisation sont ainsi à disposition pour mettre en œuvre et expérimenter les objectifs du projet.

Par ailleurs, les réflexions et débats du séminaire Ecritures numériques et nouvelles formes d'éditorialisation, dont l'IRI est co-organisateur avec la revue Sens-Public et l'Université de Montréal, alimenteront en parallèle les aspects théoriques avancés dans le projet PROFIL

- **Ina - service du dépôt légal du Web**

Interlocuteur : Claude Mussou (chef du service)

Le partenariat avec l'Ina s'appuierait sur une collaboration de longue date menée au sein des ateliers de recherche méthodologiques du dépôt légal du web à l'Ina, copilotés par Claude Mussou et Louise Merzeau depuis 2010. En tenant compte des restrictions juridiques qui pèsent là aussi sur la mise en ligne des collections, la coopération avec l'Ina pourrait prendre la forme d'un retour réflexif sur un usage ciblé des outils de recherche et de navigation développés à l'Ina. Outre qu'il pourra susciter de nombreuses recherches à effectuer dans l'archive, notre projet pourra en effet servir de cadre à une évaluation des modèles de prospection, de structuration et de restitution adoptés par l'Ina.

7. Mots-clés du projet

PROFIL, IDENTITÉ, TRACE, MÉMOIRE, FICTION

II. Mise en œuvre du projet

1. Description des actions

1.1. Période de préfiguration

Pour le lancement du projet, le choix s'est arrêté sur les fonds relatifs aux réseaux militants post-1945. La richesse et la diversité des documents conservés par la BDIC (mais aussi par la BnF), la structure communautaire des activités documentées et leurs possibles prolongements dans la période contemporaine font en effet de ces collections un matériau idéal pour notre expérimentation.

Le travail de préfiguration impliquera 4 modes opératoires :

- réunions de coordination et de gestion de projet ;
- ateliers de restitution des recherches exploratoires individuelles ;
- ateliers destinés à discuter et préciser avec les partenaires les contours et modes opératoires de chaque action envisagée (état de l'art, constitution du corpus, redocumentarisation, construction de profils et fictions transmédias) ;
- un dernier atelier portera sur la rédaction d'un document de synthèse et la programmation d'une journée d'étude destinée à diffuser les résultats de cette phase de préfiguration.

1.2. Étapes envisagées pour la phase opérationnelle du projet

1.2.1. État de l'art

Recensement et analyse :

- des dispositifs d'éditorialisation et de médiation numériques de la mémoire ;
- des modèles d'archivage et de restitution de corpus nativement numériques ;
- des dispositifs participatifs appliqués à la documentation et l'enrichissement de ressources patrimoniales

1.2.2. Constitution du corpus

Cette étape consistera à rechercher et extraire des collections de la BDIC des échantillons documentaires qui serviront de base à la construction des profils "vérificationnels" (écosystèmes documentaires et fictionnels)

1.2.3. Redocumentarisation des traces

Cette phase consistera à enrichir et croiser les traces prélevées à l'étape précédente avec :

- des ressources issues d'autres documents conservés dans des fonds analogues
- des ressources identifiées par exploration des archives du Web (BnF, Ina, Internet Archive)
- des données extraites de jeux de données ouverts relatifs aux contextes dans lesquels les traces sont susceptibles d'avoir été captées ou déposées
- des ressources hétérogènes issues du *crowdsourcing* (objets, images, témoignages, lettres...)

Le mode opératoire pourra combiner plusieurs méthodes :

- ateliers collaboratifs associant des étudiants, des documentalistes des institutions patrimoniales impliquées et des représentants de communautés scientifiques interdisciplinaires ;
- appels à contributions sur les réseaux sociaux ;
- fouille et agrégation de données.

1.2.4. Construction de profils et fictions transmédias

La génération des profils implique trois activités qui feront dans cette phase l'objet d'un travail d'analyse :

- construction par liens et agrégation d'écosystèmes cohérents à partir des extractions de ressources effectuées à l'étape précédente (pêle-mêle numériques, playlists, albums, graphes relationnels...) ;
- traduction de ces écosystèmes dans un dispositif d'annotation, de visualisation et de navigation en ligne ;
- écriture de prolongements fictionnels transmédias (scénarios, clips, blogs, webdocumentaires...) traduisant ces profils en récits de vie.

En appui de l'équipe, nous ferons ici appel à des artistes/auteurs et des communautés d'internautes de différents pays. Ces analyses seront conduites sur des dispositifs par les membres de l'équipe ou les partenaires dans le cadre d'expérimentations antérieures. Voir notamment :

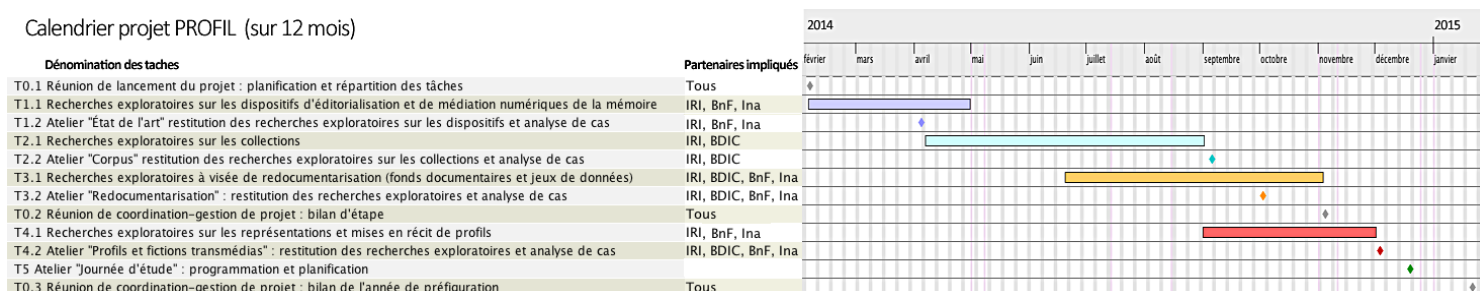
- le dispositif d'éditorialisation des [Entretiens du nouveau monde industriel](#) conçu par l'IRI ;
- l'outil [Renkan](#) développé par l'IRI, permettant d'associer des ressources web par des liens typés ;
- les outils de gestion de métadonnées, d'annotation et d'agrégation de fragments vidéos élaborés par l'IRI, comme [Hashcut](#) ou [MD-Composer](#) ;
- l'outil d'hypertagging participatif développé dans le cadre de l'infrastructure nationale [E-Recolnat](#).

1.2.5. Modélisation de logiques d'archivage numérique

Sur la base des recherches effectuées dans les deux étapes précédentes, une modélisation des nouvelles logiques de prospection, d'indexation, d'agrégation et de représentation des archives numériques sera proposée.

2. Calendrier détaillé correspondant à chacune des phases du projet

Calendrier projet PROFIL (sur 12 mois)



3. Budget prévisionnel des dépenses et recettes

DÉPENSES PRÉVISIONNELLES

RESSOURCES HUMAINES	TOTAL
---------------------	-------

Vacations de recherche	<i>coût/heure</i>	<i>nbre heures</i>	
<i>diplômé de M1 (tests et analyse de dispositifs)</i>	15,91 €	45	715,95 €
<i>doctorants (exploration des fonds, tests et analyse de dispositifs)</i>	18,60 €	230	4278,00 €

TOTAL GÉNÉRAL DES DÉPENSES	4993,95 €
----------------------------	-----------



RECETTES PRÉVISIONNELLES

	TOTAL
Contribution financière demandée au labex	4400,00 €

Autres ressources : crédits laboratoire Dicen-IDF	593,95 €
---------------------------------------------------	----------

TOTAL GÉNÉRAL DES RECETTES	4993,95 €
----------------------------	-----------