



HAL
open science

**Compte-rendu de fin de projet - Projet
ANR-07-CORP-009 BOUVARD - Les Dossiers de
Bouvard et Pécuchet de Flaubert. Enrichissement,
valorisation, documentation d'un corpus multi supports**

Stéphanie Dord-Crouslé

► **To cite this version:**

Stéphanie Dord-Crouslé. Compte-rendu de fin de projet - Projet ANR-07-CORP-009 BOUVARD - Les Dossiers de Bouvard et Pécuchet de Flaubert. Enrichissement, valorisation, documentation d'un corpus multi supports: Programme "Corpus et outils de la Recherche en Sciences Humaines et Sociales" 2007. [Rapport de recherche] ANR (Agence Nationale de la Recherche - France). 2012. halshs-00760914

HAL Id: halshs-00760914

<https://shs.hal.science/halshs-00760914>

Submitted on 4 Dec 2012

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.



Distributed under a Creative Commons Attribution 4.0 International License

Projet ANR-07-CORP-009

BOUVARD

Programme « Corpus et outils de la Recherche en
Sciences Humaines et Sociales » 2007

A	IDENTIFICATION.....	2
B	RÉSUMÉ CONSOLIDÉ PUBLIC	4
B.1	Instructions pour les résumés consolidés publics	4
B.2	Résumé consolidé public en français	4
B.3	Résumé consolidé public en anglais.....	7
C	MÉMOIRE SCIENTIFIQUE.....	9
C.1	Résumé du mémoire	9
C.2	Enjeux et problématique, état de l'art	10
C.3	Approche scientifique et technique.....	11
C.4	Résultats obtenus	13
C.5	Exploitation des résultats.....	15
C.6	Discussion	16
C.7	Conclusions.....	17
C.8	Références.....	18
D	LISTE DES LIVRABLES.....	20
E	IMPACT DU PROJET	21
E.1	Indicateurs d'impact	21
E.2	Liste des publications et communications.....	22
E.3	Liste des éléments de valorisation.....	28
E.4	Bilan et suivi des personnels recrutés en CDD (hors stagiaires) ...	29

Ce document est à remplir par le coordinateur en collaboration avec les partenaires du projet. L'ensemble des partenaires doit avoir une copie de la version transmise à l'ANR.

Ce modèle doit être utilisé uniquement pour le compte-rendu de fin de projet.

A IDENTIFICATION

Acronyme du projet	BOUVARD
Titre du projet	Les Dossiers de <i>Bouvard et Pécuchet</i> de Flaubert. Enrichissement, valorisation, documentation d'un corpus multi supports
Coordinateur du projet (société/organisme)	CNRS – UMR LIRE Littérature, idéologies, représentations aux XVIIIe et XIXe siècles, ENS de Lyon
Période du projet (date de début – date de fin)	28/12/2007 – 28/06/2012
Site web du projet, le cas échéant	http://dossiers-flaubert.ish-lyon.cnrs.fr/

Rédacteur de ce rapport	
Civilité, prénom, nom	Mme Stéphanie Dord-Crouslé
Téléphone	0437376251
Adresse électronique	Stephanie.DordCrousle@ens-lyon.fr
Date de rédaction	11/2012

Si différent du rédacteur, indiquer un contact pour le projet	
Civilité, prénom, nom	
Téléphone	
Adresse électronique	

Liste des partenaires présents à la fin du projet (société/organisme et responsable scientifique)	1- CNRS – UMR LIRE Littérature, idéologies, représentations aux XVIIIe et XIXe siècles, ENS de Lyon, resp. sc. Stéphanie Dord-Crouslé
---	---

Liste des personnes ayant participé au projet	<ul style="list-style-type: none"> • Stéphanie Dord-Crouslé, CR, CNRS, LIRE • Arahara Yukiko, Doctorante, Univ. Paris 8 • Arlin Nathalie, IE, Vacataire • Bâisse-Macchi Élisabeth, AI, CNRS, LIRE • Bara Olivier, PR, Univ. Lyon 2, LIRE • Barel-Moisan Claire, CR, CNRS, LIRE • Berthaud Christine, IR, CNRS, ISH • Burghart Marjorie, IE, EHESS, CIHAM • Champeau-Adjalian Stéphanie, MCF, Univ. de Rouen, CEREdI • Chemin-Roberty Anne, IE, ENS de Lyon, UNIS • Cordier Cécile, IA, CDD • Davi Maria Francesca, MCF, Univ. de Messine • Di Matteo Daniela, Doctorante, Univ. de Messine • Églin Véronique, MCF, INSA, LIRIS • Faure Laetitia, AI, CNRS, LIRE • Giguet Claire, IA, CDD • Gleizes Delphine, MCF, Univ. Lyon 2, LIRE • Goutaland Carine, Doctorante, Univ Lyon 2, LIRE • Lagravière Jérémie, IE, CDD • Le Calvez Éric, PR, Georgia State Univ., Atlanta
---	---

- Le Rolle Vanessa, Master, Stagiaire
- Leclerc Yvan, PR, Univ. de Rouen, CEREdI
- Lemius Christophe, IA, CNRS, LIRE
- Magaudda Biagio, Docteur, Univ. de Messine
- Malleron Vincent, Doctorant, Univ. Lyon 2, LIRE/LIRIS
- Mangiapane Stella, MCF, Univ. de Messine
- Mercier Félicie, Master, Univ. Lyon 3
- Micale Simonetta, MCF, Univ. de Messine
- Mombert Sarah, MCF, ENS de Lyon, LIRE
- Morlock-Gerstenkorn Emmanuelle, IE, CNRS, ISH
- Nakajima Taro, Docteur, Univ. Paris Est
- Notter-Truxa Françoise, IA, CNRS, LIRE
- Ogane Atsuko, PR, Univ. de Kanto Gakuin
- Panarello Saverio, Doctorant, Univ. de Messine
- Patanè Salvatore, PR, Univ. de Messine
- Pernoo Marianne, Conservateur, BIU Denis Diderot, LIRE
- Petit Nathalie, Doctorante, Univ. de Rouen, CEREdI
- Portier Pierre-Édouard, MCF, INSA, LIRIS
- Robert Joëlle, Chercheur associé, Univ. de Rouen, CEREdI
- Rosa Maria Palermo, PR, Univ. de Messine
- Schuler Hugo, IE, CDD
- Seter Barbara, Doctorante, Univ. de Messine
- Siler Douglas, Chercheur indépendant, Belgique
- Stelitano Sara, Doctorante, Univ. de Messine
- Sugaya Norioki, PR, Univ. Rikkyo, Tokyo
- Tola Martial, IE, CNRS, ISH
- Tournoy Raphaël, IE, CNRS, ISH
- Vatan Florence, MCF, Univ. of Wisconsin, Milwaukee
- Wada Mitsumasa, PR, Univ. Seinan-Gakuin, Fukuoka
- Wustner Stéphane, Master, Stagiaire
- Yamazaki Atsushi, MCF, Univ. Chukyo, Nagoya
- Zumbo Rosarianna, MCF, Univ. de Messine

B RÉSUMÉ CONSOLIDÉ PUBLIC

Ce résumé est destiné à être diffusé auprès d'un large public pour promouvoir les résultats du projet, il ne fera donc pas mention de résultats confidentiels et utilisera un vocabulaire adapté mais n'excluant pas les termes techniques. Il en sera fourni une version française et une version en anglais. Il est nécessaire de respecter les instructions ci-dessous.

B.1 INSTRUCTIONS POUR LES RÉSUMÉS CONSOLIDÉS PUBLICS

Les résumés publics en français et en anglais doivent être structurés de la façon suivante.

Titre d'accroche du projet (environ 80 caractères espaces compris)

Titre d'accroche, si possible percutant et concis, qui résume et explicite votre projet selon une logique grand public : il n'est pas nécessaire de présenter exhaustivement le projet mais il faut plutôt s'appuyer sur son aspect le plus marquant.

Les deux premiers paragraphes sont précédés d'un titre spécifique au projet rédigé par vos soins.

Titre 1 : situe l'objectif général du projet et sa problématique (150 caractères max espaces compris)

Paragraphe 1 : (environ 1200 caractères espaces compris)

Le paragraphe 1 précise les enjeux et objectifs du projet : indiquez le contexte, l'objectif général, les problèmes traités, les solutions recherchées, les perspectives et les retombées au niveau technique ou/et sociétal

Titre 2 : précise les méthodes ou technologies utilisées (150 caractères max espaces compris)

Paragraphe 2 : (environ 1200 caractères espaces compris)

Le paragraphe 2 indique comment les résultats attendus sont obtenus grâce à certaines méthodes ou/et technologies. Les technologies utilisées ou/et les méthodes permettant de surmonter les verrous sont explicitées (il faut éviter le jargon scientifique, les acronymes ou les abréviations).

Résultats majeurs du projet (environ 600 caractères espaces compris)

Faits marquants diffusables en direction du grand public, expliciter les applications ou/et les usages rendus possibles, quelles sont les pistes de recherche ou/et de développement originales, éventuellement non prévues au départ.

Préciser aussi toute autre retombée : partenariats internationaux, nouveaux débouchés, nouveaux contrats, start-up, synergies de recherche, pôles de compétitivités, etc.

Production scientifique et brevets depuis le début du projet (environ 500 caractères espaces compris)

Ne pas mettre une simple liste mais faire quelques commentaires. Vous pouvez aussi indiquer les actions de normalisation

Illustration

Une illustration avec un schéma, graphique ou photo et une brève légende. L'illustration doit être clairement lisible à une taille d'environ 6cm de large et 5cm de hauteur. Prévoir une résolution suffisante pour l'impression. Envoyer seulement des illustrations dont vous détenez les droits.

Informations factuelles

Rédiger une phrase précisant le type de projet (recherche industrielle, recherche fondamentale, développement expérimental, exploratoire, innovation, etc.), le coordonnateur, les partenaires, la date de démarrage effectif, la durée du projet, l'aide ANR et le coût global du projet, par exemple « Le projet XXX est un projet de recherche fondamentale coordonné par xxx. Il associe aussi xxx, ainsi que des laboratoires xxx et xxx). Le projet a commencé en juin 2006 et a duré 36 mois. Il a bénéficié d'une aide ANR de xxx € pour un coût global de l'ordre de xxx € »

B.2 RÉSUMÉ CONSOLIDÉ PUBLIC EN FRANÇAIS

Titre du projet

Les seconds volumes possibles de *Bouvard et Pécuchet*, roman inachevé de Flaubert

Enjeux et objectifs

L'objectif était de construire, à partir de l'édition des documents préparatoires laissés par l'auteur, des fins possibles pour ce roman inachevé.

La réalisation du projet BOUVARD a permis d'éditer sur un site dédié (<http://dossiers-flaubert.ish-lyon.cnrs.fr/>) un ensemble patrimonial fragile et complexe (mélangeant les supports manuscrits, imprimés et mixtes), d'importance scientifique et culturelle reconnue : les dossiers documentaires constitués par Flaubert en vue de la rédaction de son dernier roman, conservés aujourd'hui à la bibliothèque municipale de Rouen sous la forme de huit recueils de documents divers auxquels s'ajoutent deux recueils concernant le *Dictionnaire des idées reçues*, soit 2400 feuillets. La mise à disposition sur Internet (sous forme d'images et de transcriptions) de ce corpus auparavant très difficile d'accès s'accompagne d'un enrichissement scientifique conséquent (moteur de recherche, métadonnées, annotations et bibliothèques) dont le point d'orgue est un dispositif informatique original permettant de créer des agencements paramétrables de citations extraites des documents édités. Les internautes peuvent ainsi produire, entre autres agencements, des reconstitutions conjecturales du second volume de *Bouvard et Pécuchet*, un volume dont Flaubert avait partiellement conçu le plan et commencé à réunir la matière mais que la mort l'a empêché de rédiger.

Méthodes ou technologies utilisées

Grâce à l'encodage XML-TEI intégral du corpus, le site propose à la fois une édition en ligne multi-format et un outil de production d'agencement de citations.

Transcrit par l'équipe scientifique du projet, le corpus a été encodé en XML-TEI, un langage informatique libre et ouvert qui s'attache à décrire la structure logique des documents et à identifier les différents éléments constitutifs. Ce choix permet de proposer sur le site d'édition, à partir d'un même fichier numérique, plusieurs types de visualisation pour chaque document. Il offre aussi une réponse technique à l'autre versant du projet visant à produire des reconstitutions conjecturales du second volume de *Bouvard et Pécuchet* : il a permis de découper virtuellement le corpus en fragments textuels autonomes (selon la logique que Flaubert avait déjà partiellement mise en œuvre en vue de la rédaction de son « encyclopédie critique en farce »), tout en veillant à ce que chaque fragment reste relié à la fois à la zone d'image dont il présente la transcription et aux différentes unités textuelles dont il est partie prenante (comme la page où se situe le fragment). Stockés dans une base de données relationnelle, les fragments peuvent être convoqués et organisés selon diverses hypothèses de recherche. Les agencements de citations produits (les seconds volumes possibles) sont exportables au format XML ou PDF.

Résultats majeurs

Le corpus des dossiers documentaires de *Bouvard et Pécuchet* (texte et images) est intégralement disponible sur un site web dédié. Accompagné de plusieurs outils (moteur de recherche et bibliothèques), le site donne accès à la transcription des 2400 pages du corpus sous quatre formats : ultra-diplomatique, diplomatique, normalisé et enrichi. Il permet aussi de produire des agencements de citations à la demande, dont des seconds volumes possibles pour le roman inachevé de Flaubert. Outil de composition et de structuration de l'œuvre éditée elle-même, cette réalisation présente une tentative pour étendre le champ de l'édition critique aux contenus textuels fragmentaires.

Production scientifique

La réalisation du projet s'est accompagnée de la tenue de cinq colloques (organisés en France, Italie et Japon) et est à l'origine de plus de 50 publications diverses. Les résultats sont à la disposition de la communauté scientifique (dépôt des articles dans l'archive ouverte HAL, accès libre aux données sur le site web du projet) et on a veillé à la pérennité et à l'interopérabilité des données produites en utilisant pour l'encodage du corpus le standard international, libre et ouvert, XML-TEI.

Illustration



La chaîne de production des seconds volumes possibles de *Bouvard et Pécuchet*

Informations factuelles

Le projet BOUVARD est un projet de recherche fondamentale réalisé par une équipe scientifique internationale sous la direction de Stéphanie Dord-Crouslé, chargée de recherche dans l'UMR 5611 LIRE (Littérature, Idéologies, Représentations, XVIII^e-XIX^e siècles). Sa mise en œuvre technique a été opérée par l'ISH (Institut des Sciences de l'Homme). Le projet a commencé en janvier 2008 et a duré 54 mois. Il a bénéficié de la part de l'ANR d'une aide de 150 000 euros pour un coût complet de l'ordre de 900 000 euros.

B.3 RÉSUMÉ CONSOLIDÉ PUBLIC EN ANGLAIS

Project title

The possible second volumes for Flaubert's unfinished novel *Bouvard et Pécuchet*

Challenges and objectives

The aim was to use the online publication of the preparatory documents left by the author to explore possible endings for this unfinished novel.

The BOUVARD project made it possible to publish a fragile and complex patrimonial set (combining manuscripts, printed and mixed materials) of high scientific and cultural significance on a dedicated website (<http://dossiers-flaubert.ish-lyon.cnrs.fr/>). This set is comprised of the documentation files that Flaubert gathered to write his last novel. These 2400 sheets are now kept at the Rouen library in the form of eight collections of various documents and two additional collections dedicated to the *Dictionnaire des idées reçues*. The online availability (through images and transcriptions) of a corpus that used to be very hard to access is accompanied by noteworthy scientific enrichments (search engine, metadata, annotations, and libraries). Its main asset is an original computing device producing configurable arrangements of quotations extracted from the published documents. Among other arrangements, web users can produce hypothetical reconstructions of *Bouvard et Pécuchet's* second volume. Flaubert had partially planned this volume and started to gather material, but death prevented him from completing the work.

Methods or technologies

Thanks to the full XML-TEI encoding of the corpus, the site offers both a multi-format online edition and a tool producing arrangements of quotations.

The corpus has been transcribed by the scientific project team and then encoded in XML-TEI, a free and open computer language that seeks to describe the logical structure of documents and to identify their different components. Thanks to this choice, the edition site offers several ways of visualizing each document from a single digital file. The site also offers a technical solution for the other part of the project geared toward the hypothetical reconstructions of the second volume of *Bouvard et Pécuchet*: it allowed to cut virtually the corpus into autonomous text fragments (according to the logic that Flaubert had already partially implemented when he began to write his "critical and farcical encyclopedia"), while ensuring that each fragment is connected both to the image area which it transcribes, and to the different textual units to which it belongs, such as the page where the fragment appears. Stored in a relational database, fragments can be gathered and organized according to various research hypotheses. The resulting arrangements of quotations (possible second volumes) may be exported in XML or PDF.

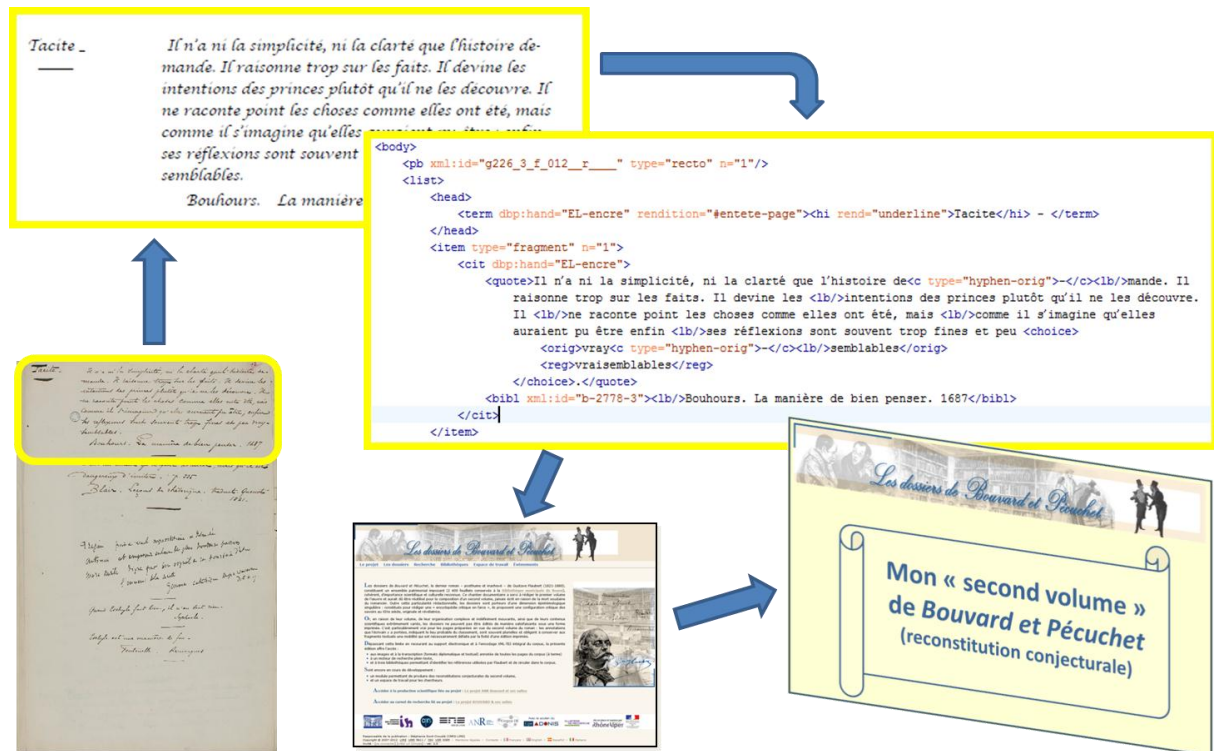
Major results

Bouvard et Pécuchet documentary files (text and images) are fully available on a dedicated website. Besides several tools (search engine and libraries), the site provides access to a corpus of 2400 pages transcribed in four formats: ultra-diplomatic, diplomatic, normalized and enriched. It also allows readers to produce arrangements of quotations on demand, including possible second volumes of Flaubert's unfinished novel. As a composition and structuring tool of the published work itself, this project seeks to extend the benefits of a critical edition to fragmentary textual contents.

Scientific production

Five conferences (organized in France, Italy and Japan) took place during the project which was also the source of more than 50 publications. Results are available to the scientific community at large (articles are deposited in the open access archive HAL, free data access is offered on the project's website) and attention was paid to sustainability and interoperability of corpus data through the use of the international, free and open encoding standard XML-TEI.

Illustration



Workflow for Bouvard et Pécuchet's possible second volumes

Factual information

The BOUVARD project is a basic research project completed by an international team of scientists and scholars under the leadership of Stéphanie Dord-Crouslé, a researcher in the UMR 5611 LIRE (Literature, Ideologies and Representations in the 18th and 19th centuries). Its technical implementation was carried out by the ISH (Institute for Human Sciences). The project began in January 2008 and lasted 54 months. It was supported by an ANR grant in the amount of 150 000 Euros for a full cost of about 900 000 Euros.

C MÉMOIRE SCIENTIFIQUE

Maximum 5 pages. On donne ci-dessous des indications sur le contenu possible du mémoire. Ce mémoire peut être accompagné de rapports annexes plus détaillés.

Le mémoire scientifique couvre la totalité de la durée du projet. Il doit présenter une synthèse auto-suffisante rappelant les objectifs, le travail réalisé et les résultats obtenus mis en perspective avec les attentes initiales et l'état de l'art. C'est un document d'un format semblable à celui des articles scientifiques ou des monographies. Il doit refléter le caractère collectif de l'effort fait par les partenaires au cours du projet. Le coordinateur prépare ce rapport sur la base des contributions de tous les partenaires. Une version préliminaire en est soumise à l'ANR pour la revue de fin de projet.

Un mémoire scientifique signalé comme confidentiel ne sera pas diffusé. Justifier brièvement la raison de la confidentialité demandée. Les mémoires non confidentiels seront susceptibles d'être diffusés par l'ANR, notamment via les archives ouvertes <http://hal.archives-ouvertes.fr>.

Mémoire scientifique confidentiel : non

C.1 RÉSUMÉ DU MÉMOIRE

Titre du projet

Les seconds volumes possibles de *Bouvard et Pécuchet*, roman inachevé de Flaubert

Enjeux et objectifs

L'objectif était de construire, à partir de l'édition des documents préparatoires laissés par l'auteur, des fins possibles pour ce roman inachevé.

La réalisation du projet BOUVARD a permis d'éditer sur un site dédié (<http://dossiers-flaubert.ish-lyon.cnrs.fr/>) un ensemble patrimonial fragile et complexe (mélangeant les supports manuscrits, imprimés et mixtes), d'importance scientifique et culturelle reconnue : les dossiers documentaires constitués par Flaubert en vue de la rédaction de son dernier roman, conservés aujourd'hui à la bibliothèque municipale de Rouen sous la forme de huit recueils de documents divers auxquels s'ajoutent deux recueils concernant le *Dictionnaire des idées reçues*, soit 2400 feuillets. La mise à disposition sur Internet (sous forme d'images et de transcriptions) de ce corpus auparavant très difficile d'accès s'accompagne d'un enrichissement scientifique conséquent (moteur de recherche, métadonnées, annotations et bibliothèques) dont le point d'orgue est un dispositif informatique original permettant de créer des agencements paramétrables de citations extraites des documents édités. Les internautes peuvent ainsi produire, entre autres agencements, des reconstitutions conjecturales du second volume de *Bouvard et Pécuchet*, un volume dont Flaubert avait partiellement conçu le plan et commencé à réunir la matière mais que la mort l'a empêché de rédiger.

Méthodes ou technologies utilisées

Grâce à l'encodage XML-TEI intégral du corpus, le site propose à la fois une édition en ligne multi-format et un outil de production d'agencement de citations.

Transcrit par l'équipe scientifique du projet, le corpus a été encodé en XML-TEI, un langage informatique libre et ouvert qui s'attache à décrire la structure logique des documents et à en identifier les différents éléments constitutifs. Ce choix permet de proposer sur le site d'édition, à partir d'un même fichier numérique, plusieurs types de visualisation pour chaque document. Il offre aussi une réponse technique à l'autre versant du projet visant à produire des reconstitutions conjecturales du second volume de *Bouvard et Pécuchet* : il a permis de découper virtuellement le corpus en fragments textuels autonomes (selon la logique que Flaubert avait déjà partiellement mise en œuvre en vue de la rédaction de son « encyclopédie critique en farce »), tout en veillant à ce que chaque fragment reste relié à la fois à la zone d'image dont il présente la transcription et aux différentes unités textuelles dont il est partie prenante (comme la page où se situe le fragment). Stockés dans une base de données relationnelle, les fragments peuvent être convoqués et organisés selon diverses hypothèses de

recherche. Les agencements de citations produits (les seconds volumes possibles) sont exportables au format XML ou PDF.

Résultats majeurs

Le corpus des dossiers documentaires de *Bouvard et Pécuchet* (texte et images) est intégralement disponible sur un site web dédié. Accompagné de plusieurs outils (moteur de recherche et bibliothèques), le site donne accès à la transcription des 2400 pages du corpus sous quatre formats : ultra-diplomatique, diplomatique, normalisé et enrichi. Il permet aussi de produire des agencements de citations à la demande, dont des seconds volumes possibles pour le roman inachevé de Flaubert. Outil de composition et de structuration de l'œuvre éditée elle-même, cette réalisation présente une tentative pour étendre le champ de l'édition critique aux contenus textuels fragmentaires.

Production scientifique

La réalisation du projet s'est accompagnée de la tenue de cinq colloques (organisés en France, Italie et Japon) et est à l'origine de plus de 50 publications diverses. Les résultats sont à la disposition de la communauté scientifique (dépôt des articles dans l'archive ouverte HAL, accès libre aux données sur le site web du projet) et on a veillé à la pérennité et à l'interopérabilité des données produites en utilisant pour l'encodage du corpus le standard international, libre et ouvert, XML-TEI.

C.2 ENJEUX ET PROBLÉMATIQUE, ÉTAT DE L'ART

Présenter les enjeux initiaux du projet, la problématique formulée par le projet, et l'état de l'art sur lequel il s'appuie. Présenter leurs éventuelles évolutions pendant la durée du projet (les apports propres au projet sont présentés en C.4).

En 2007, en s'inscrivant dans la dynamique créée par l'annonce d'une Bibliothèque numérique européenne, le projet se donnait pour objectif de favoriser – à son niveau – la visibilité et l'accessibilité nationale et internationale de la littérature et des productions scientifiques françaises, en éditant en ligne une partie de l'œuvre d'un auteur phare de la culture française. Visant des documents pour l'essentiel inédits et manuscrits qui ne pouvaient relever du traitement automatique requis par une numérisation documentaire de masse, le projet BOUVARD visait l'enrichissement, la valorisation et la documentation d'un corpus doublement original, à la fois hétérogène – en raison de la nature diverse des documents qui le composent (pages manuscrites, imprimées et complexes) ; et polymorphe (ses éléments constitutifs doivent pouvoir être organisés selon des configurations diverses et mouvantes).

Le fonds concerné est un ensemble patrimonial cohérent, d'importance scientifique et culturelle reconnue, conservé à la bibliothèque municipale de Rouen et connu sous le nom de « dossiers de *Bouvard et Pécuchet* ». Cet ensemble considérable de documents divers a été rassemblé par Flaubert en vue de la rédaction de son dernier roman, *Bouvard et Pécuchet*. Son intérêt est double : il tient d'abord à la portée épistémologique singulière de l'ouvrage qui est une « encyclopédie critique en farce » ; il vient ensuite du fait que la mort de l'auteur a interrompu l'écriture de l'œuvre : un seul des deux volumes projetés a été rédigé (sans être d'ailleurs complètement achevé), et le second est resté à l'état de chantier documentaire.

D'utiles éditions de ce « second volume » existent déjà, qu'elles tendent à l'exhaustivité ou présentent avec clarté et méthode la sélection succincte qu'elles opèrent. Néanmoins, jamais une édition imprimée ne pourra donner une image fidèle du chantier documentaire que sont les dossiers de *Bouvard et Pécuchet*, sans réduire *ipso facto* leur complexité. Car les dossiers sont le lieu d'une extrême hétérogénéité typologique. Les notes prises en vue de la rédaction du premier volume y coexistent avec celles destinées au second, puisqu'elles ont été ou devaient être partiellement réutilisées pour sa composition. De plus, comme les documents préparatoires du second volume (incomplets, dans la mesure où Flaubert devait encore y travailler) n'ont pas été résorbés dans le processus de rédaction, rien ne permet d'opérer une partition stricte entre ceux qui auraient été utilisés et ceux qui se seraient trouvés finalement rejetés. Enfin, même dans le cas des citations qui semblent les plus conformes à l'état que Flaubert pouvait chercher à atteindre, les catégories de classement se

révèlent instables et obligent à conserver aux fragments textuels une mobilité qui est nécessairement défaite par la fixité d'une édition imprimée.

Le projet BOUVARD avait donc pour ambition d'expérimenter les possibilités d'extension du champ de l'édition critique à des contenus textuels fragmentaires. En ce sens, le projet s'inscrivait dans un mouvement qui initiait un véritable changement de paradigme, puisqu'il supposait le passage du régime du livre, dans lequel s'inscrivent les pratiques éditoriales héritées, au régime de la collection numérique. Il avait pour ambition d'éditer une « collection », c'est-à-dire un rassemblement d'objets numériques mis en relation et destinés à produire du sens. Le projet BOUVARD visait donc moins à éditer des pages en ligne (reproduisant sur écran ce qui existe sur papier) qu'à permettre le surgissement et l'analyse de configurations critiques complexes, à la fois singulières et interdépendantes, qui dépendraient de la définition d'une ontologie.

C.3 APPROCHE SCIENTIFIQUE ET TECHNIQUE

Pour mettre en mouvement les fragments textuels composant le corpus tout en gardant chacun d'entre eux relié à la fois au fragment d'image dont il présente la transcription et aux différentes unités textuelles dont il est partie prenante, il a fallu organiser parallèlement :

1. l'acquisition et le traitement des images des manuscrits, y compris leur découpage en régions ou zones correspondant aux différents fragments textuels qui les composent ;
2. la production et l'enrichissement scientifique de la transcription intégrale du corpus ;
3. la structuration des données au sein d'une base relationnelle et l'encodage XML-TEI du corpus ;
4. la réalisation d'un site web répondant aux deux objectifs liés du projet : éditer le corpus et le rendre manipulable.

1- Acquisition et traitement des images

Avant même la sélection et le financement effectif du projet par l'ANR, une version numérisée du microfilm de sauvegarde des documents visés par le projet a été acquise. Ainsi, le retraitement de ces images a permis de mettre le corpus très rapidement à la disposition de l'équipe scientifique, sans attendre l'issue malheureuse des négociations qui ont été menées avec l'institution de conservation en vue de la numérisation intégrale du corpus. Cette opération aurait en effet nécessité la restauration préalable des documents et donc un important financement dédié qu'il n'a pas été possible de trouver. L'essentiel des images utilisées par le projet (au nombre de 3500) se trouve donc être les images issues de la numérisation du microfilm de sauvegarde en noir et blanc, d'une qualité moyenne mais en générale suffisante pour la lecture et l'analyse intellectuelle des spécificités manuscrites. Néanmoins, pour fournir une idée des documents plus proche de la réalité (papiers de couleur, utilisation concomitante d'encre et de crayon, etc.) et pour permettre les tests d'un doctorant informaticien travaillant sur l'extraction automatisée des lignes et des fragments dans les images des manuscrits, un échantillon de 300 images en couleur, représentatif des différents types de documents se trouvant dans le corpus, a été acquis auprès du service de numérisation de la bibliothèque municipale de Rouen aux conditions relativement onéreuses offertes au public. L'institution a en revanche mis gracieusement à la disposition du projet la numérisation des deux dossiers concernant le *Dictionnaire des idées reçues* (soit 130 images).

Un cadre juridique clair a d'emblée été donné à la diffusion des images sur le site web du projet ainsi qu'à leur utilisation par les membres du projet dans le cadre de leurs recherches personnelles. Une convention quadripartite (Ville de Rouen / CNRS / ENS de Lyon / responsable du projet) autorisant la reproduction, la diffusion et la représentation à titre gracieux des images visées par le projet a été signée le 07/07/2010.

Le découpage des images des manuscrits en régions correspondant aux différents fragments textuels a fait l'objet de plusieurs approches successives. Il s'agissait de tracer les contours de chaque zone et d'en récupérer les coordonnées SVG pour pouvoir ensuite les lier, dans les fichiers XML-TEI (voir infra), aux fragments textuels concernés. Un programme dédié visant à automatiser le processus

a été développé par un doctorant informaticien mais les tests finaux ne se sont pas révélés concluants : l'analyse des seules propriétés physiques des manuscrits ne permettait pas d'identifier les zones pertinentes avec suffisamment de précision. Cette piste a donc été abandonnée au profit d'une solution manuelle assistée par un logiciel, en l'occurrence une configuration *ad hoc* de la [plateforme DINAH](#) : chaque image est découpée manuellement en autant d'« imassettes » que l'analyse y a décelé de fragments textuels. Dans chaque fichier TEI, les fragments de texte comportent un identifiant numérique qui est mis en correspondance avec le même chiffre reporté dans le nom des régions d'images découpées par le logiciel DINAH.

2- Production et enrichissement scientifique de la transcription intégrale du corpus

La transcription et l'enrichissement scientifique du corpus ont été pris en charge par une équipe internationale (France, Japon, Italie et États-Unis) d'une trentaine de personnes, composée à la fois de flaubertistes et de spécialistes d'autres champs dix-neuviémistes dont la présence était requise en raison des fortes dimensions encyclopédiques du corpus. En un premier temps, un protocole de transcription et d'enrichissement scientifique (annotations critiques et métadonnées) a été établi. Puis, un flux de travail a été mis en place permettant de traiter le corpus au niveau des « pages ». Livrées par les collaborateurs scientifiques sous la forme d'un fichier de traitement de texte, les transcriptions ont d'abord été transformées en fichiers PDF qui offrent une visualisation ultra-diplomatique de chaque page. Les spécificités du document original y sont respectées autant que faire se peut : mise en page, fautes d'orthographe et abréviations mais aussi présence de traits de liaison ou d'indications en marge. Parallèlement, les annotations critiques (notes d'édition et de commentaire) ont été rédigées et stockées.

3- La base de données relationnelle

La base de données MySQL a été structurée en fonction des différents types d'objets éditoriaux qui ont été progressivement identifiés ; aussi la détermination des niveaux de granularité pertinents a-t-elle été une dimension essentielle du travail mené en synergie par les équipes scientifique et technique du projet. En un premier temps, chaque page a été reliée à son image et à sa transcription. Mais cette organisation ne permettait pas de rendre leur mobilité aux énoncés du corpus. Pour pouvoir insérer les fragments dans différents contextes de classement, la base de données a dû passer de l'échelle de la page à celle du fragment, ce qui impliquait d'enregistrer dans la base l'ensemble des fragments du corpus. Cette opération permettant d'identifier et de délimiter chaque unité jugée signifiante a été menée grâce à un encodage XML-TEI. Réalisé pour l'ensemble du corpus à des fins d'édition, d'exploitation en recherche, d'échange et de pérennisation, ce balisage offre une méthode efficace de découpage des pages selon une perspective logique. Le fragment peut alors être défini comme une unité textuelle qui renvoie à un passage précis d'une source imprimée ou manuscrite exogène. Il correspond à une entité de la base de données et est en relation avec un segment de la transcription TEI et avec une zone d'image. Deux bibliothèques interconnectées permettent d'attribuer aux fragments textuels leur référence bibliographique exacte : chacun d'entre eux est répertorié dans la « bibliothèque de Bouvard et Pécuchet » (qui regroupe tous les fragments textuels du corpus) et est relié à une unité documentaire appartenant à la « bibliothèque de Flaubert » (qui rassemble les références bibliographiques des ouvrages lus par Flaubert). Toutes les informations qui concernent un fragment sont accessibles sur une page dédiée dotée d'une URL pérenne.

Mais le processus rédactionnel particulier dont les dossiers de *Bouvard et Pécuchet* sont le résultat amène souvent à ce qu'un même contenu textuel se trouve recopié à plusieurs endroits dans le corpus. Ces différents fragments doivent pouvoir être classés typologiquement (ils appartiennent à des stades rédactionnels différents) et surtout ils doivent pouvoir être manipulés conjointement. À cette fin, un objet éditorial particulier a été créé : la citation qui est un regroupement de fragments textuels partageant une même référence bibliographique exacte et dont l'un a été choisi pour être le fragment de référence de la citation. Toutes les informations la concernant sont accessibles sur une page dédiée dotée d'une URL pérenne. Les configurations plurielles produites à partir des éléments du corpus le sont au niveau des citations, que celles-ci regroupent un seul ou plusieurs fragments.

4- La réalisation d'un site web

Le site web a été développé avec des logiciels libres ou open-source. L'essentiel du développement a été fait avec le langage [PHP](#). Les transcriptions XML-TEI sont stockées dans une base de données XML native ([eXist-db](#)) qui permet d'interroger les textes grâce au langage XQuery et de formuler des requêtes complexes. Les affichages multiples des transcriptions sont créés avec le langage de transformation XSLT.

Les autres données non XML sont enregistrées dans une base de données relationnelle [MySQL](#). Un moteur de recherche « [Apache SolR](#) » permet de faire des requêtes plein texte sur les transcriptions, les références bibliographiques et les métadonnées qui leur sont associées. La recherche peut se faire à l'aide d'équations booléennes et de facettes.

Les données du projet sont sauvegardées sur la grille du TGE-ADONIS.

Toutes les opérations préalables avaient pour but de permettre la lecture du corpus et de rendre techniquement mobiles les fragments textuels qui le composent afin de remédier à la brusque immobilisation du chantier documentaire causée par la mort de Flaubert. Restait donc à permettre au lecteur d'organiser l'agencement de ces fragments en fonction de ses propres hypothèses et à en tester les résultats. Ceci a été rendu possible grâce au développement de plusieurs interfaces décrites dans le point suivant.

C.4 RÉSULTATS OBTENUS

Positionner les résultats par rapports aux livrables du projet et aux publications, brevets etc. Revisiter l'état de l'art et les enjeux à la fin du projet.

1- L'édition intégrale du corpus complexe des dossiers documentaires de *Bouvard et Pécuchet*

Sur le site « [Les dossiers de Bouvard et Pécuchet](#) » (<http://dossiers-flaubert.ish-lyon.cnrs.fr>), les deux objectifs conjoints du projet ont été réalisés. Le premier visait à éditer, c'est-à-dire à donner librement accès, sous forme images et texte, au corpus des dossiers documentaires constitué par Flaubert en vue de la rédaction de son dernier roman. Pour cela, deux interfaces de lecture ont été créées. La première offre une entrée dans le corpus par l'unité « page » et selon différents modes de classement des documents (classement [patrimonial](#), classement [typologique](#), classement [chronologique](#) et classement [par scripteurs](#)). Chaque page peut être visualisée de quatre manières différentes : outre l'affichage [ultra-diplomatique](#) procuré par le fichier PDF, trois autres affichages HTML sont générés grâce à l'encodage des fichiers en XML-TEI : [diplomatique](#) (visant la reproduction du document originel au plus près), [normalisé](#) (proposant une version lisible par tous : fautes d'orthographe rectifiées, abréviations développées, etc.) et [enrichi](#) (permettant de visualiser les différences existant entre les affichages diplomatique et normalisé). Au niveau de la page sont articulés plusieurs enrichissements scientifiques, comme les annotations critiques et les [métadonnées](#). Au nombre de ces dernières figurent les références bibliographiques qui permettent d'accéder aux deux unités fondamentales du corpus : le fragment textuel et la citation (qui est un regroupement de fragments textuels partageant la même référence bibliographique exacte). Un second point d'entrée dans le corpus est fourni par l'unité « texte ». Un texte est un ensemble de pages résultant d'une construction critique. On accède aux textes par l'intermédiaire d'un [sommaire](#) qui les liste selon leur type et permet leur lecture en continu sous forme diplomatique ou normalisée.

Outre ces deux interfaces de lecture ont été développées deux interfaces d'interrogation qui leur sont complémentaires et permettent d'autres types d'entrées dans le corpus. La première consiste en un [moteur de recherche](#) plein-texte dans les fichiers encodés en XML-TEI et dont les résultats peuvent être affinés grâce à un tri par facettes. Parallèlement, on peut accéder aux résultats de [requêtes complexes](#) prédéfinies formulées au moyen du langage XQuery. L'interrogation du corpus peut aussi se faire par le biais de [bibliothèques](#) qui sont au nombre de trois. Indépendamment de la « [bibliographie critique](#) » qui rassemble les travaux scientifiques en rapport avec le champ du projet, la « [bibliothèque de Flaubert](#) » propose les références bibliographiques exactes de l'ensemble des

ouvrages qui sont, à un moment ou à un autre, passés sur la table de travail de Flaubert et/ou qui sont mentionnés d'une manière ou d'une autre dans le corpus des dossiers documentaires. La « [bibliothèque de Bouvard et Pécuchet](#) », quant à elle, regroupe tous les fragments textuels qui se trouvent dans les dossiers documentaires et les met en relation avec une référence bibliographique exacte appartenant à la « bibliothèque de Flaubert ».

2- Un outil spécifique de production d'agencements de citations dont des « seconds volumes » possibles

Rendu accessible, le corpus peut aussi être manipulé, ce qui était l'objectif ultime visé. A été développé un module de production à la demande d'agencements (regroupements et ordonnancements) de citations qui se décompose en trois interfaces : la première (« [Plans](#) ») sert à construire le squelette de l'agencement à partir des termes d'indexation (ou catégories) présents dans le corpus – ou permet de modifier tout ou partie d'un plan déjà existant (édition « papier » antérieure, classement matériel des documents, etc.) ; la deuxième interface (« [Matrice](#) ») permet de sélectionner les citations qui vont composer l'agencement à partir du croisement de différents critères (statut typologique des fragments, périmètre génétique, etc.) ; et la troisième est une interface de lecture et d'édition (« [Agencements](#) ») : elle permet de visualiser et d'affiner l'agencement des citations qui est ensuite exportable aux formats PDF et XML. Les reconstitutions conjecturales du second volume de *Bouvard et Pécuchet* forment un cas d'agencement de citations, particulièrement intéressant mais auquel ne se réduisent pas les nombreuses utilisations possibles du module développé : il peut produire tout type de corpus constitués à partir des filtres proposés par la matrice et/ou directement alimentés par l'intermédiaire d'un « panier » de résultats de recherche.

La structuration de la base de données, la rédaction du manuel d'encodage XML-TEI dédié au projet et la conception générale de l'architecture du site web sont le fruit d'une élaboration progressive menée en synergie par les équipes scientifique et technique. Nourrie par les différents colloques et journées d'études qui se sont tenus à intervalles réguliers, cette réflexion s'est exprimée dans les publications qui ont accompagné l'avancée du projet et qui formulent des propositions théoriques plus larges.

3- Un exemple d'édition de corpus fragmentaire

En effet, les agencements de citations créent des objets qui n'existent pas indépendamment des liens qui les génèrent : ce sont des objets éditoriaux entièrement nouveaux, pluriels, résultant d'une agrégation structurée de citations. On peut leur conférer le titre d'« œuvres » singulières si l'on prend la précaution indispensable de préciser que leur auteur n'est pas Flaubert et que le sens que chacune d'entre elles délivre ne peut pas être saisi indépendamment des autres et doit toujours être replacé au sein d'une pluralité de possibles. L'invention du module de production à la demande d'agencements de citations propose ainsi une réponse au défi que représente l'édition d'un corpus si particulier qu'il requiert l'invention d'une forme éditoriale adaptée. Puisqu'on ne peut pas reconstituer le film de la composition de l'œuvre (désir qui sommeille secrètement en tout généticien...), c'est la mise en mouvement des possibilités d'agencement que cette œuvre recèle qui dessine ce qu'ont pu être ou ce à quoi auraient pu conduire les intentions de l'auteur, les chemins qu'elles auraient pu emprunter au vu des traces qu'il a laissées – sans introduire de biais latent ni restreindre la complexité originale du corpus.

L'image du kaléidoscope illustre bien le processus. Dispositif optique constitué d'un nombre fini d'éléments contenus dans un espace lui-même fini, mais autorisant un nombre indéfini de combinaisons entre ses éléments, le kaléidoscope présente la particularité de créer du nouveau par simple réagencement de ce qui existait déjà auparavant. Un nombre fini d'éléments produit ainsi un nombre indéfini de figures différentes. Ce ne sont donc pas les éléments qui font le tout, mais la forme que prend leur combinaison. Comme le tube du kaléidoscope, l'outil électronique de lecture permet de faire l'expérience que le tout est plus que la somme des parties, et l'édition de fragments donne naissance à une forme d'énonciation éditoriale qui s'exprime implicitement, en creux, par l'agencement des fragments : recueil, sélection et mise en série. L'édition électronique d'un corpus tel que celui des dossiers documentaires de *Bouvard et Pécuchet* ne pouvait donc se contenter d'être un

outil de lecture et un outil de diffusion, elle devait également se faire outil de composition et de structuration de l'œuvre éditée elle-même. De ce fait, elle représente une tentative pour étendre le champ de l'édition critique aux contenus textuels fragmentaires.

Symétriquement, l'apparition de ces nouveaux objets implique l'avènement d'un nouveau lecteur, promu dans un rôle où il se voit conférer une partie des compétences qui étaient jusque-là l'apanage exclusif de l'éditeur. Le lecteur devient auteur du texte, ou du moins de l'organisation du texte qu'il se donne à lire. Et c'est justement en faisant ainsi l'expérience nécessairement réitérée et plurielle de la mise en collection et de l'agencement des citations, que le lecteur peut approcher et tenter d'appréhender de l'intérieur le processus éditorial imaginé par Flaubert pour la composition de son second volume, un processus inabouti mais dont le chantier sans fin des dossiers documentaires de *Bouvard et Pécuchet* présente les prodromes.

C.5 EXPLOITATION DES RÉSULTATS

Outre les publications mentionnées en E2 et E3 (dont trois thèses portant directement sur des sections du corpus édité, soutenues devant l'université Paris Est-Marne-la-Vallée pour l'une et devant l'université de Messine pour les deux autres), le site web issu du projet est dorénavant ouvert aux investigations de la communauté flaubertiste, et plus largement aux chercheurs s'intéressant aux rapports de la littérature et des savoirs ou aux problèmes de l'édition de corpus fragmentaires. Il devrait donc rapidement servir à produire la ou les grandes études d'ensemble, transversales, dépassant les monographies portant sur un domaine particulier traité dans le roman encyclopédique de Flaubert (comme la littérature, la médecine ou la pédagogie), auxquelles les chercheurs étaient jusqu'ici contraints de se limiter, en raison des difficultés qu'ils rencontraient pour maîtriser ne serait-ce qu'un seul champ disciplinaire. La mise à disposition intégrale du corpus et les outils qui accompagnent son édition ont fait sauter le verrou que représentait pour les chercheurs la complexité de ce corpus.

Pour améliorer la visibilité des ressources structurées produites par le projet *Bouvard et Pécuchet* et favoriser leur exploitation, un entrepôt [OAI-PMH](http://dossiers-flaubert.ish-lyon.cnrs.fr/oai/) (Open Archives Initiative Protocol for Metadata Harvesting) a été développé (<http://dossiers-flaubert.ish-lyon.cnrs.fr/oai/>). Il permet le moissonnage des métadonnées du corpus par différents moteurs de recherche et en particulier par [ISIDORE](#) (Intégration de Services, Interconnexion de DONnées de la Recherche et de l'Enseignement), la plateforme de recherche créée par le [TGE-ADONIS](#), qui collecte, indexe et enrichit les métadonnées et les données numériques structurées accessibles selon des protocoles informatiques standardisés interopérables couramment utilisés par les communautés des sciences humaines et sociales.

Enfin, diverses collaborations sont en cours. La plus avancée concerne le projet d'édition des manuscrits du premier volume de *Bouvard et Pécuchet*. En effet, outre les pages préparées pour le second volume du roman, l'édition des dossiers documentaires de *Bouvard et Pécuchet* met à la disposition des chercheurs et du public les matériaux ayant servi à préparer la rédaction du premier volume du roman (en particulier les notes de lecture et les « notes de notes »). Le corpus entretient donc des relations étroites avec le corpus des manuscrits rédactionnels du premier volume en cours d'édition par le Centre Flaubert (CERÉdI) de l'université de Rouen sous la direction d'Yvan Leclerc (voir la [présentation du projet](#)). Sont visés des manuscrits qui sont eux aussi conservés à la bibliothèque municipale de Rouen :

- Ms gg10 – les plans et scénarios du premier et du second volume (72 feuillets) ;
- Ms g224 (1) et (2) – le manuscrit définitif du premier volume (300 feuillets) ;
- Ms g225 (1) à (9) – les brouillons rédactionnels du premier volume (1203 feuillets).

La collaboration entre les deux entreprises est effective depuis longtemps : de nombreux transcripteurs font partie concomitamment des deux équipes scientifiques et il est prévu que des pages conservées matériellement dans un corpus mais appartenant intellectuellement à l'autre puissent être virtuellement réintégrées. En outre, et bien que les technologies utilisées de part et d'autre soient différentes, tout sera fait pour que des passerelles aussi nombreuses et pertinentes que

possibles soient établies entre les deux sites et permettent aux chercheurs d'accéder efficacement à l'intégralité des documents touchant à la genèse du dernier roman de Flaubert.

Une collaboration interdisciplinaire prometteuse est aussi en cours avec un laboratoire d'informatique lyonnais. Développé par [Pierre-Édouard Portier](#) (INSA, LIRIS) dans le cadre du [projet Desanti](#), le [logiciel DINAH](#) outille, grâce à une plateforme philologique, le processus de construction des documents multistructurés. Une configuration simplifiée du logiciel adaptée au projet Bouvard permet déjà le découpage et l'archivage de l'imagette correspondant à chaque fragment textuel du corpus (accessible au niveau de la métadonnée « Référence bibliographique de fragment »). Mais un projet de plus grande envergure vise à développer un mécanisme de type « bureau de travail » où les chercheurs pourront tester des hypothèses de classement et de datation. Ils pourront opérer des agencements libres ou guidés à partir du contenu d'un panier ; ils pourront sauvegarder ces agencements en vue de leur analyse et récupérer les enrichissements produits. Divers usages sont déjà envisagés, comme de pouvoir procéder à des analyses thématiques, linguistiques ou génétiques de sous-corpus variés, proposer des datations relatives pour les pages et les fragments du corpus, ou reconstituer des pages dans leur état premier, antérieur aux découpages effectués par Flaubert (fonction « puzzle »). Mais d'autres usages restent à inventer. Des réponses communes à différents appels à projets ont été rédigées ([ARC 5](#) de la région Rhône-Alpes et [IXXI](#)-Institut rhônalpin des systèmes complexes).

Enfin, grâce à l'interopérabilité des données produites (encodage XML-TEI), le corpus Bouvard a pu bénéficier d'une importation dans la [plateforme TXM](#) (Serge Heiden et Alexei Lavrentev, CNRS-ICAR) qui permet de construire et d'analyser différents corpus textuels structurés à partir des dossiers dans une [perspective textométrique](#). Lors de l'import, chaque mot des transcriptions a été automatiquement lemmatisé et annoté avec une étiquette grammaticale (avec un taux d'erreur d'environ 5%) pour offrir des moyens d'indexation supplémentaires. La plateforme permet de construire des concordances triables et exportables à partir d'une recherche de mots ou de séquences textuelles effectuée sur tout ou partie de ce corpus, des listes de fréquences de mots, lemmes, etc. ou de rechercher des cooccurrences. Cette exploitation linguistique du corpus construit par le projet n'était pas prévue à l'origine et peut se révéler très féconde.

C.6 DISCUSSION

Discussion sur le degré de réalisation des objectifs initiaux, les verrous restant à franchir, les ruptures, les élargissements possibles, les perspectives ouvertes par le projet, l'impact scientifique, industriel ou sociétal des résultats.

Si l'essentiel des objectifs décrits dans le projet initial financé par l'ANR a été rempli et si les résultats dépassent le plus souvent les espérances, il est quelques rares secteurs où ils restent en deçà. En particulier, les tentatives menées pour mettre en place un partenariat avec la bibliothèque municipale de Rouen ont échoué. Les images diffusées sur le site d'édition des dossiers documentaires de *Bouvard et Pécuchet* ne sont donc pas des images couleur haute définition, mais seulement les images numérisées en noir et blanc du microfilm de sauvegarde des documents. Si la qualité de ces images est suffisante pour l'exploitation scientifique littéraire des contenus, elle n'a pas permis la réalisation de tous les traitements qui avaient été envisagés dans le cadre de la thèse d'un doctorant en informatique travaillant sur l'extraction automatisée des lignes et des régions textuelles dans les images. Si les documents venaient un jour à être restaurés par leur institution de conservation, une campagne de numérisation intégrale du corpus pourrait alors être menée - et si un accord dans ce sens intervenait avec la bibliothèque municipale de Rouen, les images HD pourraient alors remplacer sur le site d'édition les actuelles images en noir et blanc.

D'autre part, si la quasi-totalité du corpus manuscrit est accessible au format PDF, c'est-à-dire sous la forme de transcriptions ultra-diplomatiques, tous ces fichiers n'ont pas pu encore être encodés en XML-TEI. Le balisage, dans lequel seules quelques opérations mineures peuvent être automatisées, est en effet un processus particulièrement complexe et chronophage qui demande des opérateurs à la fois

formés aux logiques scientifiques du projet et maîtrisant la grammaire de la TEI ainsi que l'utilisation d'un éditeur XML. Cette double compétence n'a malheureusement pas pu être développée dans l'ensemble de l'équipe scientifique, ce qui aurait été évidemment souhaitable. L'encodage du corpus est donc toujours en cours et va se poursuivre dans les mois à venir.

Outre les collaborations en cours évoquées en C.5, qui vont permettre d'exploiter les résultats produits par le projet, il est envisagé de repousser les limites du corpus. En circonscrivant les contours du « second volume » aux dossiers documentaires conservés à la bibliothèque municipale de Rouen, le projet avait fait le choix – pour le dossier soumis à l'ANR – d'éditer un ensemble patrimonial cohérent et hautement représentatif du dessein flaubertien. Néanmoins, il n'y a pas une exacte concordance entre ces dossiers et la matière assignable au « second volume » de *Bouvard et Pécuchet*. L'écart se lit en termes d'excédents : des documents ont été serrés dans les dossiers qui n'y ont pas leur place légitime, comme le rappelle cette mention inscrite sur un sous-dossier par la nièce de Flaubert : « Papiers trouvés ça et là sur la table de travail » (c'est-à-dire trouvés après la mort de l'écrivain, lorsque la légataire universelle a rangé, comme elle a pu, les affaires du défunt). Mais les dossiers pèchent principalement par manque : plusieurs ensembles de notes de lecture prises par Flaubert pour la rédaction du premier volume, et donc potentiellement réutilisables pour la composition du second, n'ont pas été donnés par Caroline Franklin Grout à la Ville de Rouen. Ils ont été dispersés lors des deux grandes ventes successorales qui ont suivi son décès en 1931 et ont depuis disparu dans des collections privées. Or, il serait éminemment souhaitable que la cohérence intellectuelle du corpus des dossiers documentaires de *Bouvard et Pécuchet* puisse être rétablie de manière virtuelle grâce à la publication sur le site du projet Bouvard de la totalité des dossiers ayant servi à la préparation du dernier roman de Flaubert, quel que soit le lieu de leur conservation matérielle actuelle. C'est dans cette direction que le site d'édition pourra évoluer.

C.7 CONCLUSIONS

Grâce à la caution scientifique qu'a représentée sa sélection au titre de l'appel à projets CORPUS 2007 et grâce au financement induit, le projet Bouvard a pu réaliser sur un [site web](#) les deux objectifs qu'il s'était initialement donnés : mettre à la disposition du public et des chercheurs un corpus outillé constitué de documents complexes d'importance culturelle et scientifique reconnue (les dossiers documentaires de *Bouvard et Pécuchet*) ; et expérimenter les possibilités d'extension du champ de l'édition critique à des contenus textuels fragmentaires en permettant de produire des agencements de citations à la demande (dont une partie proposent des « seconds volumes » possibles pour le roman inachevé de Flaubert). Tout en poursuivant ces deux objectifs, l'équipe scientifique et technique du projet, enracinée dans l'UMR lyonnaise LIRE, a tissé des relations interdisciplinaires suivies à l'échelon local avec différents laboratoires, soit directement soit par l'intermédiaire de diverses structures ([ISH](#), [MUTEC](#), [Cluster 13](#) puis [ARC 5](#) de la région Rhône-Alpes), mais aussi aux échelons national ([Consortium CAHIER](#) partie prenante de la Très Grande Infrastructure de Recherche [CORPUS](#)) et international (collaborations suivies avec l'université de Messine en Italie et l'université Rikkyo de Tokyo au Japon). Elle a ainsi pris sa place dans le paysage en pleine évolution des humanités numériques.

C.8 RÉFÉRENCES

- « *Bouvard et Pécuchet* : les "seconds volumes" possibles - Documentation, circulations, édition », Colloque de fin de projet ANR, ENS de Lyon, dir. S. Dord-Crouslé, France (2012) ; actes à paraître dans la *Revue Flaubert* (<http://flaubert.univ-rouen.fr/revue/>), n° 13 (2013)
- Dord-Crouslé S., « Un dossier flaubertien mal connu : les notes pour le chapitre "Littérature" de *Bouvard et Pécuchet* » ; *Histoires littéraires*, n° 24 (2005) 119-135 [[halshs-00199139](#)]
- Dord-Crouslé S., « Notes de lecture et édition du "second volume" de *Bouvard et Pécuchet* : configurations complexes de l'inachèvement » ; à paraître dans les actes de la journée d'études du 1^{er} juin 2012 : « *Bouvard et Pécuchet*. Genèse et interprétation » [[fichier audio en ligne](#)] dans *Flaubert. Revue critique et génétique*, n° 7, 2012 [[halshs-00745159](#)]
- Dord-Crouslé S., Mangiapane S., Palermo Di Stefano R. M., *Éditer le chantier documentaire de Bouvard et Pécuchet. Explorations critiques et premières réalisations numériques*, Messina, Andrea Lippolis Editore, (2010) 261 [[halshs-00549164](#)]
- Dord-Crouslé S., Morlock-Gerstenkorn E., « Le "modèle abstrait" du corpus Bouvard : première approche », Journée d'étude "Constitution et exploitation de corpus issus de manuscrits - Lectures, écritures et nouvelles approches en recherche documentaire" organisée par Cécile Meynard et Thomas Lebarbé, France (2009) [[halshs-00368044](#)]
- Dord-Crouslé S., Morlock-Gerstenkorn E., « Sur le modèle du kaléidoscope : concevoir l'édition électronique du "second volume" de *Bouvard et Pécuchet* », *Nouveaux Cahiers François Mauriac*, 19 (« L'édition critique, de l'imprimé au numérique : François Mauriac et les autres... ») (2011) 169-183 [[halshs-00602644](#)]
- Dord-Crouslé S., Morlock-Gerstenkorn E., « L'édition électronique des dossiers de *Bouvard et Pécuchet* de Flaubert : des fragments textuels en quête de mobilité », *Le patrimoine à l'ère du numérique : structuration et balisage*, sous la dir. C. Bougy, C. Dornier et C. Jacquemard, PU de Caen (2011) 79-89 [[halshs-00441286](#)]
- Dord-Crouslé S., Morlock-Gerstenkorn E., Tournoy R., « Nouveaux objets éditoriaux. Le site d'édition des dossiers documentaires de *Bouvard et Pécuchet* (Flaubert) », *Cahiers du numérique* 7, 3-4/2011 "Empreintes de l'hypertexte. Rétrospective et évolution", sous la dir. de Caroline Angé (2012) 123-145 [[halshs-00736015](#)]
- Flaubert G., *Bouvard et Pécuchet ; avec des fragments du second volume, dont le Dictionnaire des idées reçues*, éd. de S. Dord-Crouslé, Paris, Flammarion, « GF » (2008)
- Flaubert G., *Universalenzyklopädie der Menschlichen Dummheit. Ein Sottisier*, éd. de H.-H. Henschen, Frankfurt am Main, Eichborn-Berlin (2004)
- Flaubert G., *Le Second volume de Bouvard et Pécuchet, le projet du Sottisier, reconstitution conjecturale de la « copie » des deux bonshommes d'après le dossier de Rouen*, éd. d'A. Cento et L. Caminiti Pennarola, Naples, Liguori (1981)
- Flaubert G., *Bouvard et Pécuchet. Œuvre posthume augmentée de la Copie*, dans les *Œuvres complètes*, éd. nouvelle établie par la Société des Études littéraires françaises..., t. 5 et 6, Paris, Club de l'Honnête Homme (1972)
- Flaubert G., *Le Second volume de Bouvard et Pécuchet*, éd. de G. Bollème, Paris, Denoël, « Dossier des Lettres Nouvelles » (1966)
- Flaubert G., *Les dossiers documentaires de Bouvard et Pécuchet*. Transcription intégrale des documents conservés à la bibliothèque municipale de Rouen, accompagnée d'un outil de production de « seconds volumes » possibles, sous la dir. de S. Dord-Crouslé (2012), <http://dossiers-flaubert.ish-lyon.cnrs.fr>
- Flaubert G., *Bouvard et Pécuchet. Transcription intégrale du manuscrit (premier volume : plans et scénarios, brouillons et version finale)*, sous la dir. d'Y. Leclerc et D. Girard (2011-...), http://flaubert.univ-rouen.fr/bouvard_et_pecuchet/chantier/roman_final/BPframe1.html
- Magaudda B., *La genèse du chapitre VI de Bouvard et Pécuchet : les apports des notes de lecture du dossier "Politique"*. (En annexe: images et transcriptions diplomatiques des manuscrits du dossier

- "Politique" de *Bouvard et Pécuchet*), Thèse soutenue devant l'Université de Messine (École doctorale de Philologie), Italie, le 20 juin 2011, dir. Mme Rosa Maria Palermo (2011) [[halshs-00601759](#)]
- Malleron V., « Text lines and snippets extraction for 19th century handwriting documents layout analysis », *10th International Conference on Document Analysis and Recognition*, Espagne (2009) [[halshs-00420060](#)]
 - Nakajima T., *Les figures religieuses dans l'œuvre de Flaubert*, Thèse de doctorat. Directeur de thèse : Mme Gisèle SEGINGER, professeur des universités à l'université Paris Est (2009) [[halshs-00450590](#)]
 - Pierssens M., Compte rendu de « Éditer le chantier documentaire de *Bouvard et Pécuchet*. Explorations critiques et premières réalisations numériques, textes réunis par R. M. Palermo Di Stefano, S. Dord-Crouslé, S. Mangiapane, Messine, Andrea Lippolis Editore, 2010, 261 p. », [Site Flaubert](#) (2011)
 - Seter B., *Il dossier « Socialisme-Politique » nelle note preparatorie per il secondo volume di Bouvard et Pécuchet*, Thèse soutenue devant l'Université de Messine (« Lingue e Letterature galloromanze »), Italie, 1^{er}/06/2012, dir. Mme Rosa Maria Palermo (2012) [[halshs-00744988](#)]
 - TEI Consortium. (2011). TEI P5: Guidelines for Electronic Text Encoding and Interchange. <http://www.tei-c.org/Guidelines/P5/>
 - Sugaya N., « Pour une édition électronique du second volume de *Bouvard et Pécuchet* : le dossier médical », *Bulletin de la Section française*, Faculté des lettres, Université Rikkyo, 38 (2009) 37-47 [[halshs-00371294](#)]
 - Portier P-E., Calabretto S., « [DINAH, a philological platform for the construction of multi-structured documents](#) », Dans *The European Conference on Research and Advanced Technology for Digital Libraries (ECDL)*, Lalmas M., Jose J., Rauber A., Sebastiani F., Frommholz I. ed., ECDL 2010 September 6-10, 2010, Glasgow, Springer, « Research and advanced technology for digital libraries » (2010) 364-375

D LISTE DES LIVRABLES

Quand le projet en comporte, reproduire ici le tableau des livrables fourni au début du projet. Mentionner l'ensemble des livrables, y compris les éventuels livrables abandonnés, et ceux non prévus dans la liste initiale.

Date de livraison	N°	Titre	Nature (rapport, logiciel, prototype, données, ...)	Partenaires (souligner le responsable)	Commentaires
Numérisation et traitement des images du corpus					
2008		Acquisition et traitement numérique du microfilm de sauvegarde	Données	LIRE	
2008-10		Spécifications établies en vue de la numérisation des documents originaux	Rapport	LIRE - LIRIS	Auteur : V. Malleron
2009-09		Acquisition d'un échantillon de documents originaux numérisés	Données	LIRE	
2010		Outil de zonage des images	Prototype	LIRE - LIRIS	Auteur : V. Malleron. Outil abandonné
2012		Outil de zonage des images issu de la Plateforme DINAH	Prototype	LIRIS	Auteur : P.-E. Portier
Base de données					
2007		Structuration	Données	ISH	Auteur : R. Tournoy
2009		Base bibliographique	Données	ISH	Auteur : S. Dord-Crouslé
Transcription					
2007		Manuel à destination des collaborateurs	Manuel technique	LIRE	Auteur : S. Dord-Crouslé
Encodage					
2008-07		Manuel d'encodage V1	Rapport de stage	ISH	Auteur : V. Le Rolle
2012		Manuel d'encodage V2	Manuel technique	ISH	Auteur : E. Morlock
Mise en ligne des données					
2007		Wiki à destination des collaborateurs scientifiques	Manuel technique	ISH	
2008		Site web fermé destiné au travail collaboratif	Logiciel	ISH	
2009		Interface plurilingue pour les pages statiques	Données	ISH - Université de Messine	
2012-01		Site web ouvert	Logiciel	ISH	http://dossiers-flaubert.ish-lyon.cnrs.fr/
Publications et manifestations scientifiques (voir le détail en E2 et E3)					
Autres livrables					
2011		Collection sur HAL-SHS	Données	LIRE	http://halshs.archives-ouvertes.fr/ANR-07-CORP-009
2011		Carnet de recherche	Données	LIRE	http://flaubert.hypotheses.org/
2012		Entrepôt OAI-PMH	Données	LIRE-ISH	http://dossiers-flaubert.ish-lyon.cnrs.fr/oai/

E IMPACT DU PROJET

Ce rapport rassemble des éléments nécessaires au bilan du projet et plus globalement permettant d'apprécier l'impact du programme à différents niveaux.

E.1 INDICATEURS D'IMPACT

Nombre de publications et de communications (à détailler en 0)

Comptabiliser séparément les actions monopartenaire, impliquant un seul partenaire, et les actions multipartenaires résultant d'un travail en commun.

Attention : éviter une inflation artificielle des publications, mentionner uniquement celles qui résultent directement du projet (postérieures à son démarrage, et qui citent le soutien de l'ANR et la référence du projet).

		Publications multipartenaires	Publications monopartenaire
International	Revue à comité de lecture		7
	Ouvrages ou chapitres d'ouvrage		26
	Communications (conférence)		
France	Revue à comité de lecture		13
	Ouvrages ou chapitres d'ouvrage		5 + 2 sous presse + 1 à paraître en 2013
	Communications (conférence)		27
Actions de diffusion	Articles vulgarisation		2
	Conférences vulgarisation		8
	Autres		

Autres valorisations scientifiques (à détailler en [94])

Ce tableau dénombre et liste les brevets nationaux et internationaux, licences, et autres éléments de propriété intellectuelle consécutifs au projet, du savoir faire, des retombées diverses en précisant les partenariats éventuels. Voir en particulier celles annoncées dans l'annexe technique).

	Nombre, années et commentaires (valorisations avérées ou probables)
Brevets internationaux obtenus	Sans objet
Brevet international en cours d'obtention	Sans objet
Brevets nationaux obtenus	Sans objet
Brevet nationaux en cours d'obtention	Sans objet
Licences d'exploitation (obtention / cession)	Sans objet
Créations d'entreprises ou essaimage	Sans objet
Nouveaux projets collaboratifs	En gestation
Colloques scientifiques	5 (2008, 2009 X 2, 2010 et 2012)
Autres (préciser)	1 Mémoire de recherche (2008) 1 Rapport de stage (2008) Insertion dans différents réseaux structurants pour la recherche en SHS

E.2 LISTE DES PUBLICATIONS ET COMMUNICATIONS

Répertorier les publications résultant des travaux effectués dans le cadre du projet. On suivra les catégories du premier tableau de la section E.1 en suivant les normes éditoriales habituelles. En ce qui concerne les conférences, on spécifiera les conférences invitées.

2008

International

Articles dans des revues à comité de lecture

- [1] Sugaya N., « Luxure et Bêtise : les notes de lecture de Flaubert sur le Père Debreyne », *Bulletin de la section française*, Faculté des Lettres (Université Rikkyo - Tokyo), 37 (2008) 19-30 [[halshs-00371053](#)]

France

Articles dans des revues à comité de lecture

- [2] Dord-Crouslé S., « Le beau dans tous ses états - le moment esthétique dans *Bouvard et Pécuchet* », *Gustave Flaubert - La revue des lettres modernes*, 6 "Fiction et philosophie" sous la dir. de Gisèle Séginger (2008) 85-107 [[halshs-00388399](#)]
- [3] Vatan F., « La "puissance de l'image". Bouvard et Pécuchet, disciples de Taine ? », *Gustave Flaubert - La revue des lettres modernes*, 6 "Fiction et philosophie" sous la dir. de Gisèle Séginger (2008) 109-141 [[halshs-00391844](#)]
- [4] Yamazaki A., « Le dossier "Philosophie" de *Bouvard et Pécuchet* : Hegel et Spinoza », *Gustave Flaubert - La revue des lettres modernes*, 6 "Fiction et philosophie" sous la dir. de Gisèle Séginger (2008) 225-253 [[halshs-00395621](#)]

Conférences de vulgarisation

- [5] Dord-Crouslé, S. « Parcours et modes de circulation dans l'espace d'une collection numérique : le projet Bouvard », communication prononcée à l'occasion des « 20 ans de l'Institut des Sciences de l'Homme » le 19 juin 2008 à Lyon [[vidéo en ligne](#) sur le site de l'ISH ; et [support de présentation](#)]

2009

International

Articles dans des revues à comité de lecture

- [6] Sugaya N., « Pour une édition électronique du second volume de *Bouvard et Pécuchet* : le dossier médical », *Bulletin de la Section française*, Faculté des lettres, Université Rikkyo, 38 (2009) 37-47 [[halshs-00371294](#)]
- [7] Mangiapane S., « Le dossier "Agriculture" dans les notes de lecture de *Bouvard et Pécuchet* (premiers éléments) », *Plaisance*, 17, Anno VI (2009) 157-168 [[halshs-00442924](#)]

Ouvrages ou chapitres d'ouvrage

- [8] Sugaya N., « Flaubert et le roman démocrate », *Balzac, Flaubert. La genèse de l'œuvre et la question de l'interprétation* (Proceedings of the Second International Conference, 14-16 dec. 2007, Nagoya, Japan), Nagoya University, Graduate School of Letters (2009) 111-116 [[halshs-00371442](#)]
- [9] Arahara Y., « Des mots et des fossiles - La géologie dans *Bouvard et Pécuchet* », *Balzac, Flaubert. La genèse de l'œuvre et la question de l'interprétation* (Proceedings of the Second International Conference, 14-16 dec. 2007, Nagoya, Japan), Nagoya University, Graduate School of Letters (2009) 95-102 [[halshs-00371692](#)]
- [10] Wada M., « Comment interpréter l'ambiguïté épistémologique dans *Bouvard et Pécuchet* de Flaubert », *Balzac, Flaubert. La genèse de l'œuvre et la question de l'interprétation* (Proceedings of the Second International Conference, 14-16 dec. 2007, Nagoya, Japan), Nagoya University, Graduate School of Letters (2009) 103-110 [[halshs-00391405](#)]
- [11] Malleron V., « Hierarchical decomposition of handwritten manuscripts layouts », *Lecture Notes in Computer Science - Computer Analysis of Images and Patterns*, Springer Berlin / Heidelberg (2009) 221-228 [[halshs-00420059](#)]

- [12] Malleron V., « Text lines and snippets extraction for 19th century handwriting documents layout analysis », *10th International Conference on Document Analysis and Recognition*, Espagne (2009) [[halshs-00420060](#)]

Communications ou conférences

- [13] Magaouda B., « Le dossier "Politique" dans les notes de lecture de *Bouvard et Pécuchet* », Journée Nationale des Recherches doctorales en Langue et Littérature françaises en Italie (Université de Pescara "Gabriele D'Annunzio"), Italie (2009) [[halshs-00442928](#)]

France

Articles dans des revues à comité de lecture

- [14] Dord-Crouslé S., « Flaubert libre lecteur. À propos de l'*Abrégé du catéchisme de persévérance* de l'abbé Gaume », *Flaubert. Revue critique et génétique* (<http://flaubert.revues.org>), 1 « Nouvelles recherches critiques et génétiques » (2009) 17 [[halshs-00354507](#)]
- [15] Sugaya N., « Bouvard et Pécuchet et le savoir médical », *Flaubert. Revue critique et génétique* (<http://flaubert.revues.org>), 1 « Nouvelles recherches critiques et génétiques » (2009) 20 [[halshs-00358225](#)]

Ouvrages ou chapitres d'ouvrage

- [16] Nakajima T., *Les figures religieuses dans l'œuvre de Flaubert*, Thèse de doctorat. Directeur de thèse : Mme Gisèle SEGINGER, professeur des universités à l'université Paris Est (2009) [[halshs-00450590](#)]

Communications ou conférences

- [17] Dord-Crouslé S., Morlock-Gerstenkorn E., « Le "modèle abstrait" du corpus Bouvard : première approche », Journée d'étude "Constitution et exploitation de corpus issus de manuscrits - Lectures, écritures et nouvelles approches en recherche documentaire" organisée par Cécile Meynard et Thomas Lebarbé, France (2009) [[halshs-00368044](#)]
- [18] Malleron V., « Corpus de Manuscrits Numériques : Apports de l'analyse d'image au travers de l'exemple des "Dossiers de *Bouvard et Pécuchet*" », Journée d'étude "Constitution et exploitation de corpus issus de manuscrits - Lectures, écritures et nouvelles approches en recherche documentaire" organisée par Cécile Meynard et Thomas Lebarbé, France (2009) [[halshs-00377387](#)]
- [19] Morlock-Gerstenkorn E., « The files of *Bouvard et Pécuchet* », Public session of the TEI Technical Council, France (2009) [[halshs-00419300](#)]
- [20] Malleron V., « Extraction automatisée de lignes et de fragments textuels dans les images de manuscrits d'auteur du XIX^e siècle », MANifestation des JEunes Chercheurs en Sciences et Technologies de l'Information et de la Communication, France (2009) [[halshs-00443548](#)]

Conférences de vulgarisation

- [21] Dord-Crouslé S., « L'édition électronique des dossiers de *Bouvard et Pécuchet* de Flaubert : des fragments textuels en quête de mobilité », Séminaire de bilan et de prospective du Cluster 13 « Culture, patrimoine, création » mis en place et soutenu par la Région Rhône-Alpes, vendredi 23 octobre 2009, Château de Montchat, Lyon 3e, France (2009) [[halshs-00426295](#)]
- [22] Malleron V., « Le numérique et l'interdisciplinarité au service des dossiers de *Bouvard et Pécuchet* : Vers une mobilité retrouvée », Séminaire de bilan et de prospective du Cluster 13 « Culture, patrimoine, création » mis en place et soutenu par la Région Rhône-Alpes, vendredi 23 octobre 2009, Château de Montchat, France (2009) [[halshs-00426391](#)]

2010

International

Articles dans des revues à comité de lecture

- [23] Mangiapane S., « Dialogismes transtextuels dans la genèse de *Bouvard et Pécuchet* de Flaubert », *Plaisance VII*, 19 - Actes du 6^e colloque "Testo, Metodo, Elaborazione elettronica. Dialogismi", Messina, (2010) 21-32 [[halshs-00443073](#)]

- [24] Palermo Di Stefano R. M., « Les notes de lecture de Flaubert sur l'Histoire de l'Art chez les Anciens », *Plaisance* VII, 19 - Actes du 6^e colloque "Testo, Metodo, Elaborazione elettronica. Dialogismi", Messina, (2010) 11-20 [[halshs-00504761](#)]
- [25] Micale S., « Éléments dialogiques dans le "second volume" de *Bouvard et Pécuchet* », *Plaisance* VII, 19 - Actes du 6^e colloque "Testo, Metodo, Elaborazione elettronica. Dialogismi", Messina, (2010) 33-44 [[halshs-00505454](#)]

Ouvrages ou chapitres d'ouvrage

- [26] Sugaya N., *Flaubert épistémologue. Autour du dossier médical de Bouvard et Pécuchet*, Rodopi, (2010) 276 [[halshs-00480255](#)]
- [27] Dord-Crouslé S., Mangiapane S., Palermo Di Stefano R. M., *Éditer le chantier documentaire de Bouvard et Pécuchet. Explorations critiques et premières réalisations numériques*, Messina, Andrea Lippolis Editore, (2010) 261 [[halshs-00549164](#)] ; 25 contributions dont :
- [28] Dord-Crouslé S., « Vers une édition électronique des dossiers de *Bouvard et Pécuchet* », 15-20 [[halshs-00549160](#)]
- [29] Mangiapane S., Patané S., Panarello S., Stelitano S., « Une interface multilingue pour l'édition électronique des dossiers de *Bouvard et Pécuchet* », 21-29 [[halshs-00550567](#)]
- [30] Bara O., « Autour de "Rocaille. Théâtre". Petite enquête sur un dossier préparatoire de *Bouvard et Pécuchet* », 49-59 [[halshs-00550671](#)]
- [31] Barel-Moisan C., « Quelques pistes sur le dossier "Socialisme": organisation, destination, problématiques », 61-79 [[halshs-00550673](#)]
- [32] Dord-Crouslé S., « Entre notes de lecture et fragments préparés pour le second volume, les transferts de citations à l'épreuve du dossier Religion », 81-96 [[halshs-00426296](#)]
- [33] Gleizes D., « Modalités de la prise de notes dans les dossiers préparatoires de *Bouvard et Pécuchet* : le cas du dossier "Socialisme" », 97-104 [[halshs-00550675](#)]
- [34] Le Calvez E., « République de 1848 », 105-120 [[halshs-00370973](#)]
- [35] Leclerc Y., « La section "Théâtre" dans les dossiers de *Bouvard et Pécuchet* », 121-126 [[halshs-00550676](#)]
- [36] Magauda B., « Le dossier "Politique" dans la documentation préparatoire de *Bouvard et Pécuchet* : description du corpus », 127-140 [[halshs-00550678](#)]
- [37] Mangiapane S., « De la citation à la paraphrase. Réécritures du savoir encyclopédique dans les dossiers de *Bouvard et Pécuchet* », 141-155 [[halshs-00550679](#)]
- [38] Micale S., « Le dossier "Styles (Spécimen de) - Périphrases", de loin et de près », 157-170 [[halshs-00550680](#)]
- [39] Mombert S., « Le jugement critique. Étude génétique du dossier "Critique" des dossiers de *Bouvard et Pécuchet* », 171-185 [[halshs-00550682](#)]
- [40] Petit N., « Le dossier "Journaux" : de *L'Éducation sentimentale* au "second volume" de *Bouvard et Pécuchet* », 187-205 [[halshs-00550683](#)]
- [41] Seter B., « Le dossier "Socialisme-Politique" dans les notes préparatoires pour le second volume de *Bouvard et Pécuchet* », 207-214 [[halshs-00550684](#)]
- [42] Sugaya N., « Régularités et distorsions : les transferts d'extraits dans le dossier médical de *Bouvard et Pécuchet* », 215-228 [[halshs-00550686](#)]
- [43] Wada M., « Éduquer et écrire : le dossier pédagogique de *Bouvard et Pécuchet* », 229-236 [[halshs-00550687](#)]
- [44] Yamazaki A., « La destination des notes de lecture du dossier "Philosophie" », 237-251 [[halshs-00550689](#)]

France

Articles dans des revues à comité de lecture

- [45] Dord-Crouslé S., Mercier F., « Les notes prises par Flaubert sur la revue *L'Artiste* », *Gustave Flaubert - La revue des lettres modernes*, 7 "Flaubert et la peinture" sous la dir. de Gisèle Séginger (2010) 255-309 [[halshs-00522648](#)]

Ouvrages ou chapitres d'ouvrage

- [46] Dord-Crouslé S., « De l'esclavage selon Flaubert », *Littérature et esclavage, XVIIIe-XIXe siècles*, sous la dir. de S. Moussa, Desjonquères (2010) 371-382 [[halshs-00397326](#)]

Conférences de vulgarisation

- [47] Dord-Crouslé S., Tournoy R., « Présentation du projet d'édition en ligne des dossiers de *Bouvard et Pécuchet* (Flaubert) », École d'été ALHIS (Approches Littéraires et Historiques des Sources), 7-9/09/2010, Univ. de Saint-Étienne.

2011

International

Articles dans des revues à comité de lecture

- [48] Magaudda B., « Le transfert des notes de lecture du Dossier "Politique" dans le chapitre VI de *Bouvard et Pécuchet* : illustration d'un exemple significatif », *Publif@rum*, 16 (2011) 1-10 [[halshs-00648315](#)]

Ouvrages ou chapitres d'ouvrage

- [49] Magaudda B., *La genèse du chapitre VI de Bouvard et Pécuchet : les apports des notes de lecture du dossier "Politique"*. (En annexe: images et transcriptions diplomatiques des manuscrits du dossier "Politique" de Bouvard et Pécuchet), Thèse soutenue devant l'Université de Messine (École doctorale de Philologie), Italie, le 20 juin 2011, dir. Mme Rosa Maria Palermo (2011) [[halshs-00601759](#)]
- [50] Mangiapane S., « Flaubert et la tentation du dictionnaire », *Genèse du Dictionnaire. L'aventure des synonymes*, Actes des Septièmes Journées Italiennes des Dictionnaires, Université de Messine 2-4 décembre 2010, Schena Editore - Alain Baudry et Cie (2011) 231-246 [[halshs-00596747](#)]

France

Articles dans des revues à comité de lecture

- [51] Dord-Crouslé S., Morlock-Gerstenkorn E., « Sur le modèle du kaléidoscope : concevoir l'édition électronique du "second volume" de *Bouvard et Pécuchet* », *Nouveaux Cahiers François Mauriac*, 19 ("L'édition critique, de l'imprimé au numérique : François Mauriac et les autres...") (2011) 169-183 [[halshs-00602644](#)]
- [52] Magaudda B., « Le savoir politique dans le chapitre VI de *Bouvard et Pécuchet* : apports livresques et fiction », *Revue Flaubert*, 11 - Fictions du savoir, savoirs de la fiction dans *Bouvard et Pécuchet* (2011) 1-10 [[halshs-00677952](#)]
- [53] Leclerc Y., « Notes de cours sur *Bouvard et Pécuchet*, fictions du savoir et savoirs de la fiction », *Revue Flaubert*, 11- Fictions du savoir, savoirs de la fiction dans *Bouvard et Pécuchet* (2011) 1-28 [[halshs-00679092](#)]

Ouvrages ou chapitres d'ouvrage

- [54] Wada M., « Une leçon de choses inattendue ou comment Victor a appris la notion de la propriété dans *Bouvard et Pécuchet* », *Comment naît une œuvre littéraire* (Actes du Colloque à l'Université de Kyoto, 7-9 décembre 2007), Champion (2011) 175-189 [[halshs-00597746](#)]
- [55] Dord-Crouslé S., Morlock-Gerstenkorn E., « L'édition électronique des dossiers de *Bouvard et Pécuchet* de Flaubert : des fragments textuels en quête de mobilité », *Le patrimoine à l'ère du numérique : structuration et balisage*, sous la dir. C. Bougy, C. Dornier et C. Jacquemard, PU de Caen (2011) 79-89 [[halshs-00441286](#)]

Articles de vulgarisation

- [56] Barel-Moisan C., « Fictions du savoir, savoirs de la fiction : Flaubert, *Bouvard et Pécuchet* », dans Barel-Moisan C., Giboux A., Mcintosh-Varjabédian F., Robineau-Weber A.-G., *Fictions du savoir, savoirs de la fiction : Goethe, "Les affinités électives", Melville, "Mardi", Flaubert, "Bouvard et Pécuchet"*, Atlande (2011) 1-100 [[halshs-00679103](#)]

Conférences de vulgarisation

- [57] Dord-Crouslé S., Morlock-Gerstenkorn E., « Le projet Bouvard » ; Atelier de lancement du Consortium CAHIER (Corpus d'Auteurs pour les Humanités: Informatisation, Édition, Recherche), Paris, Sorbonne, 14/15-12-2011.

2012

International

Ouvrages ou chapitres d'ouvrage

- [58] Seter B., *Il dossier « Socialisme-Politique » nelle note preparatorie per il secondo volume di Bouvard et Pécuchet*, Thèse soutenue devant l'Université de Messine (« Lingue e Letterature galloromanze »), Italie, 1^{er}/06/2012, dir. Mme Rosa Maria Palermo (2012) [[halshs-00744988](#)]

France

Articles dans des revues à comité de lecture

- [59] Dord-Crouslé S., « La place de la fiction dans le second volume de *Bouvard et Pécuchet* », *Revue en ligne Arts et savoirs* (ISSN 2258-093X), 1- *Bouvard et Pécuchet : la fiction des savoirs* (2012) 1-21 [[halshs-00674134](#)]
- [60] Dord-Crouslé S., Morlock-Gerstenkorn E., Tournoy R., « Nouveaux objets éditoriaux. Le site d'édition des dossiers documentaires de *Bouvard et Pécuchet* (Flaubert) », *Cahiers du numérique* 7, 3-4/2011 "Empreintes de l'hypertexte. Rétrospective et évolution", sous la dir. de Caroline Angé (2012) 123-145 [[halshs-00736015](#)]
- [61] Nakajima T., « Notes de lecture et généalogie des idées. Le discours apologétique de Jeufroy et la pensée maistrienne », *Revue en ligne Arts et savoirs* (ISSN 2258-093X), 1- *Bouvard et Pécuchet : la fiction des savoirs* (2012) 1-10 [[halshs-00677954](#)]
- [62] Sugaya N., « La poétique documentaire de *Bouvard et Pécuchet*. Visibilité des savoirs et arrangement scénarique », *Revue en ligne Arts et savoirs* (ISSN 2258-093X), 1- *Bouvard et Pécuchet : la fiction des savoirs* (2012) 1-12 [[halshs-00677953](#)]

Ouvrages ou chapitres d'ouvrage

- [63] Dord-Crouslé S., « "Ce Louis-Philippe en littérature" : Flaubert juge de Casimir Delavigne », *Casimir Delavigne en son temps (Vie culturelle - Théâtre - Réception)*. Actes du colloque de Rouen, 24-25 octobre 2011 (2012) 327-339 [[halshs-00715987](#)]
- [64] Dord-Crouslé S., « Bouvard, Pécuchet et la "vache désespérée" : le magnétisme entre savoir et farce » ; à paraître dans : *La science au miroir de la littérature : triomphe et contestations de la Révolution à nos jours*, sous la dir. de Lise Dumasy, Droz, sous presse.

Communications ou conférences

- [65] Colloque « *Bouvard et Pécuchet* : les "seconds volumes" possibles - Documentation, circulations, édition », ENS de Lyon, dir. Stéphanie Dord-Crouslé, France (2012) ; à paraître dans la *Revue Flaubert*, n° 13, 2013 :
- [66] Angé C., « Édition de fragments : les enjeux de la mise en forme numérique » [[halshs-00678861](#)]
- [67] Heiden S., Lavrentev A., « Exploration textométrique du corpus BOUVARD » [[halshs-00678874](#)]
- [68] Portier P.-E., « Manipulations multimodales pour la construction de documents multistructurés » [[halshs-00678876](#)]
- [69] Marchal H., « La poésie scientifique dans les pages préparées pour le second volume de *Bouvard et Pécuchet* » [[halshs-00678879](#)]
- [70] Palermo R. M., « La "Littérature" de *Bouvard et Pécuchet* d'après les notes sur les "Grands écrivains" » [[halshs-00678882](#)]
- [71] Rey P.-L., « Les goûts littéraires de Bouvard et Pécuchet » [[halshs-00678883](#)]
- [72] Zékian S., « Faut-il corriger les classiques ? L'"ineptie des critiques" dans les dossiers de *Bouvard et Pécuchet* » [[halshs-00678886](#)]
- [73] Neefs J., « Paroles en l'air : l'espace des dialogues dans *Bouvard et Pécuchet* » [[halshs-00678887](#)]

- [74] Séginger G., « La réécriture de Cuvier : la création du monde entre savoir et féerie » [\[halshs-00678891\]](#)
- [75] Pellegrini F., « "Pour savoir la chimie, ils se procurèrent le Cours de Regnault" : modélisation discursive et savoir disciplinaire » [\[halshs-00678897\]](#)
- [76] Mangiapane S., « Des mots du savoir aux mots de la fiction. Le lexique de l'agriculture dans le chapitre II de *Bouvard et Pécuchet* » [\[halshs-00678900\]](#)
- [77] Vatan F., « Vertige et quête de maîtrise : Flaubert et le Dr Le Fèvre » [\[halshs-00678901\]](#)
- [78] Yamazaki A., « Le magnétisme animal dans les notes de lecture » [\[halshs-00678902\]](#)
- [79] Sugaya N., « Entre le 1er et le second volume : Flaubert et Raspail » [\[halshs-00678909\]](#)
- [80] Marquer B., « "Masticat ridendo mores" - Flaubert et le comique physiologique » [\[halshs-00678913\]](#)
- [81] Barel-Moisan C., « "Des séminaristes en goguette ou des caissiers en délire" : religion, économie et esthétique dans le dossier "Socialisme" » [\[halshs-00678914\]](#)
- [82] Magaouda B., « La construction de l'épisode du droit divin dans le chapitre VI de *Bouvard et Pécuchet* : analyse de la séquence à partir des notes de lecture du dossier "Politique" et des brouillons » [\[halshs-00678916\]](#)
- [83] Nakajima T., « Présence du catholicisme intransigeant de Joseph de Maistre » [\[halshs-00678920\]](#)
- [84] Herschberg Pierrot A., « Le christianisme et l'esclavage : effets de voix dans *Bouvard et Pécuchet* » [\[halshs-00678922\]](#)
- [85] Ogane A., « Le pèlerinage à la Délivrande, une vision ironique du catholicisme » [\[halshs-00678927\]](#)
- [86] Mombert S., « La vie des journaux dans les dossiers de *Bouvard et Pécuchet* » [\[halshs-00678929\]](#)
- [87] Petit N., « Le dossier "Journaux" : de la note documentaire à *L'Éducation sentimentale* » [\[halshs-00678930\]](#)
- [88] Wada M., « Le statut de l'enfant à travers les notes de lecture sur l'éducation » [\[halshs-00678934\]](#)
- [89] Demanze L., « Bouvard et Cie : les écrivains contemporains à l'ombre de Flaubert » [\[halshs-00678938\]](#)
- [90] Dord-Crouslé S., « Notes de lecture et édition du "second volume" de *Bouvard et Pécuchet* : configurations complexes de l'inachèvement » ; à paraître dans les actes de la journée d'études du 1^{er} juin 2012 : « *Bouvard et Pécuchet*. Genèse et interprétation » [\[fichier audio en ligne\]](#) dans *Flaubert. Revue critique et génétique*, n° 8, 2012 [\[halshs-00745159\]](#)

Articles de vulgarisation

- [91] Giguet C., « *Bouvard et Pécuchet* : les "seconds volumes" possibles. Documentation, circulation, édition », *La Lettre de l'ISH*, n° 2, septembre 2012, p. 7.

Conférences de vulgarisation

- [92] Dord-Crouslé S., Tournoy T., « L'édition des dossiers documentaires de *Bouvard et Pécuchet* (Flaubert). D'un roman inachevé à la production de "seconds volumes" à la demande », Déjeuners de l'ISH et de l'ENS de Lyon, ISH, 27/04/2012.
- [93] Dord-Crouslé S., Tournoy T., « Le projet Bouvard : la diversité des objets éditoriaux », Journée organisée par l'URFIST de Lyon, avec le CCSD et le laboratoire ELICO « La publication scientifique pendant et après la thèse : Regards croisés entre SHS et STM », 19/10/2012.

2013

France

Ouvrages ou chapitres d'ouvrage

- [94] Malleron V., « Environnement collaboratif pour l'enrichissement, la valorisation et la documentation d'un corpus multi supports en sciences humaines : l'édition électronique des dossiers de *Bouvard et Pécuchet* de Flaubert », thèse de doctorat, co-direction Philippe Régner (Université Lyon 2) et Véronique Eglin (INSA), soutenance prévue début 2013.

E.3 LISTE DES ÉLÉMENTS DE VALORISATION

La liste des éléments de valorisation inventorie les retombées (autres que les publications) décomptées dans le deuxième tableau de la section E.1. On détaillera notamment :

- brevets nationaux et internationaux, licences, et autres éléments de propriété intellectuelle consécutifs au projet.
- logiciels et tout autre prototype
- actions de normalisation
- lancement de produit ou service, nouveau projet, contrat,...
- le développement d'un nouveau partenariat,
- la création d'une plate-forme à la disposition d'une communauté
- création d'entreprise, essaimage, levées de fonds
- autres (ouverture internationale,...)

Elle en précise les partenariats éventuels. Dans le cas où des livrables ont été spécifiés dans l'annexe technique, on présentera ici un bilan de leur fourniture.

1- Site web : édition du corpus et outil de production d'agencements de citations

<http://dossiers-flaubert.ish-lyon.cnrs.fr/>

2- Collection – dédiée aux productions issues du projet – sur l'archive ouverte HAL-SHS

<http://halshs.archives-ouvertes.fr/ANR-07-CORP-009>

3- Carnet de recherche

<http://flaubert.hypotheses.org/>

4- Colloques et journées d'études

- Journées d'études internationales sur l'édition des dossiers documentaires de *Bouvard et Pécuchet*, organisées par S. Dord-Crouslé, Lyon, 11-12/12/2008 [[programme](#)]
- Grâce à un financement PHC Galileo (France-Italie) pour 2009 : « Les dossiers documentaires de *Bouvard et Pécuchet* de Flaubert : vers une édition en ligne à interface plurilingue » ; responsable français : S. Dord-Crouslé, responsable étranger : R.M. Palermo (Université de Messine) :
 - Journées Bouvard, organisées par S. Dord-Crouslé, Lyon, 2-3/07/2009 [[programme](#)]
 - Journées Bouvard, organisées par R.M. Palermo, Messine, 19-20/10/2009 [[programme](#)]
- Journée d'études Bouvard, organisée par N. Sugaya, Tokyo, 13/03/2010
- Colloque de fin de programme ANR : « *Bouvard et Pécuchet* : les "seconds volumes" possibles - Documentation, circulations, édition », organisé par S. Dord-Crouslé, ENS de Lyon, 7-9/03/2012 [[programme](#)]

5- Mémoires de recherche et rapports de stage

- Rapport de stage (3 mois) de Master 2 professionnel « Patrimoine écrit : histoire et pratiques de l'édition » de Vanessa Le Rolle, « Les Dossiers préparatoires de *Bouvard et Pécuchet* de Flaubert : Approche critique d'une structuration du corpus avec la TEI (*Text Encoding Initiative*) », Université François Rabelais de Tours (Centre d'Études Supérieures de la Renaissance) et Institut des Sciences de l'Homme (Service d'ingénierie documentaire), soutenu le 11 juillet 2008 (co-direction : R. Tournoy, M.-L. Demonet et Ch. Bénévent)
- Mémoire de Master 1 de Félicie Mercier, « Flaubert et la revue *L'Artiste* : édition critique des notes contenues dans les dossiers documentaires de *Bouvard et Pécuchet* », soutenu à l'Université Jean Moulin – Lyon III, le 29 septembre 2008 (co-direction : S. Dord-Crouslé et R. Jomand-Baudry)

6- Insertion dans différents réseaux structurants pour la recherche en SHS

- au niveau local : [ISH](#), [MUTEC](#), [Cluster 13](#) puis [ARC 5](#) de la région Rhône-Alpes
- au niveau national : [Consortium CAHIER](#) partie prenante de la Très Grande Infrastructure de Recherche [CORPUS](#)
- et au niveau international : collaborations suivies avec l'université de Messine en Italie et l'université Rikkyo de Tokyo au Japon

E.4 BILAN ET SUIVI DES PERSONNELS RECRUTÉS EN CDD (HORS STAGIAIRES)

Ce tableau dresse le bilan du projet en termes de recrutement de personnels non permanents sur CDD ou assimilé. Renseigner une ligne par personne embauchée sur le projet quand l'embauche a été financée partiellement ou en totalité par l'aide de l'ANR et quand la contribution au projet a été d'une durée au moins égale à 3 mois, tous contrats confondus, l'aide de l'ANR pouvant ne représenter qu'une partie de la rémunération de la personne sur la durée de sa participation au projet.

Les stagiaires bénéficiant d'une convention de stage avec un établissement d'enseignement ne doivent pas être mentionnés.

Les données recueillies pourront faire l'objet d'une demande de mise à jour par l'ANR jusqu'à 5 ans après la fin du projet.

Identification				Avant le recrutement sur le projet			Recrutement sur le projet				Après le projet				
Nom et prénom	Sexe H/F	Adresse email (1)	Date des dernières nouvelles	Dernier diplôme obtenu au moment du recrutement	Lieu d'études (France, UE, hors UE)	Expérience prof. antérieure, y compris post-docs (ans)	Partenaire ayant embauché la personne	Poste dans le projet (2)	Durée missions (mois) (3)	Date de fin de mission sur le projet	Devenir professionnel (4)	Type d'employeur (5)	Type d'emploi (6)	Lien au projet ANR (7)	Valorisation expérience (8)
LAGRAVIERE Jérémie	H														
FAURE Laetitia	F														
SCHULER Hugo	H														
ARLIN FOURNIE Nathalie	F														

Aide pour le remplissage

(1) **Adresse email** : indiquer une adresse email la plus pérenne possible

(2) **Poste dans le projet** : post-doc, doctorant, ingénieur ou niveau ingénieur, technicien, vacataire, autre (préciser)

(3) **Durée missions** : indiquer en mois la durée totale des missions (y compris celles non financées par l'ANR) effectuées sur le projet

(4) **Devenir professionnel** : CDI, CDD, chef d'entreprise, encore sur le projet, post-doc France, post-doc étranger, étudiant, recherche d'emploi, sans nouvelles

(5) **Type d'employeur** : enseignement et recherche publique, EPIC de recherche, grande entreprise, PME/TPE, création d'entreprise, autre public, autre privé, libéral, autre (préciser)

(6) **Type d'emploi** : ingénieur, chercheur, enseignant-chercheur, cadre, technicien, autre (préciser)

(7) **Lien au projet ANR** : préciser si l'employeur est ou non un partenaire du projet

(8) **Valorisation expérience** : préciser si le poste occupé valorise l'expérience acquise pendant le projet.

Les informations personnelles recueillies feront l'objet d'un traitement de données informatisées pour les seuls besoins de l'étude anonymisée sur le devenir professionnel des personnes recrutées sur les projets ANR. Elles ne feront l'objet d'aucune cession et seront conservées par l'ANR pendant une durée maximale de 5 ans après la fin du projet concerné. Conformément à la loi n° 78-17 du 6 janvier 1978 modifiée, relative à l'Informatique, aux Fichiers et aux Libertés, les personnes concernées disposent d'un droit d'accès, de rectification et de suppression des données personnelles les concernant. Les personnes concernées seront informées directement de ce droit lorsque leurs coordonnées sont renseignées. Elles peuvent exercer ce droit en s'adressant l'ANR (<http://www.agence-nationale-recherche.fr/Contact>).