

TGIR HUMA-NUM

Rapport d'activité du Consortium archiPolis

Novembre 2014

Résumé des objectifs du Consortium

Le Consortium archiPolis a pour objectif de fédérer des laboratoires de sociologie et de science politique afin d'inventorier et de sauvegarder, en format numérique, les enquêtes réalisées à l'aide de méthodes qualitatives et portant sur l'objet politique, des plus spécifiques telles que les élections, les partis, les syndicats, l'action publique, les institutions, la citoyenneté, mais aussi plus largement sur toutes formes collectives d'appartenances et de comportements mettant en jeu l'action publique ou les rapports au pouvoir politique.

Ces enquêtes se présentent sous deux formes matérielles :

- Des enquêtes conservées sur des supports matériels, qu'il faut numériser ; ce sont les plus anciennes qui constituent le fond patrimonial de la recherche empirique des laboratoires partenaires ;
- Des enquêtes nativement numériques ; ce sont les plus récentes et les enquêtes à venir.

Par rapport à ces gisements de données, archiPolis fonctionne comme un réseau coopératif de mise en commun des ressources et des façons de faire, portant sur une série limitée d'opérations :

- inventaire, collecte et préservation physique d'archives scientifiques
- numérisation d'archives scientifiques
- documentation d'archives scientifiques (le terme de « documentation » désigne ici le renseignement de métadonnées, les transcriptions ou toutes autres activités qui permettent de documenter le contexte de production des données et ainsi d'améliorer l'utilisation scientifique des corpus).

En l'absence de savoir-faire antérieur dans la recherche en SHS, la mise au point de standards, la création d'outils communs d'inventaires de numérisation et d'archivage et leur utilisation constituent, avec le repérage préalable des enquêtes au sein des laboratoires, les principales activités du Consortium. Ce travail nécessite des réunions entre laboratoires

partenaires, une formation aux procédures standardisées d'inventaire, de documentation et numérisation.

La constitution d'une archive numérique des enquêtes dans le réseau des laboratoires d'archiPolis doit se faire en étroite collaboration avec beQuali du CDSP qui a vocation à diffuser certaines de ces enquêtes. La compatibilité des standards de classement, d'archivage et de numérisation des deux structures doit être totale.

Il faut enfin souligner que le Consortium ne peut atteindre son objectif de préservation raisonnée et standardisée du patrimoine de la recherche empirique qualitative des laboratoires qu'en s'appuyant sur la confiance des chercheurs et sur leur motivation à déposer à archiPolis leurs données. Parvenir à susciter l'une et l'autre est l'un des enjeux majeurs de la valorisation des activités du Consortium au sein de la communauté scientifique.

Modifications dans l'organisation du pilotage du Consortium

(Cf. les CR des trois derniers comités de pilotage en annexe)

Les difficultés liées à l'organisation et la priorisation des activités du consortium sont apparues et ont été discutées lors du comité du pilotage du 14 mai 2014.

Lors du comité de pilotage du 10 juillet 2014, un changement de portage et de coordination du Consortium a été proposé. Sophie Duchesne, jusque là coordinatrice, a souhaité mettre fin aux fonctions qu'elle occupait depuis l'origine. Son départ a entraîné le remplacement de l'ISP comme unité porteuse du Consortium par le CDSP.

Le comité du pilotage du 10 octobre 2014 a validé ce changement et désigné Bernard Denni, de l'UMR PACTE de Grenoble, comme coordinateur scientifique par un vote à l'unanimité.

La liste des laboratoires membres a été actualisée. L'ISP est sorti du Consortium. Il reste une incertitude sur la participation du Centre d'études européennes, incertitude qui devra être prochainement levée par une consultation directe du CEE. L'OSC est maintenant membre du Consortium.

Les modalités de représentation des laboratoires membres du Consortium ont été précisées ainsi que l'organisation des réunions et du travail collectif.

Il a également été voté la constitution d'un conseil scientifique du Consortium en 2015.

Les décisions concernant l'organisation institutionnelle des relations entre archiPolis et beQuali sont abordées dans la partie du rapport consacrée aux relations entre ces deux structures.

Activités du Consortium en 2014

Avant de présenter le bilan des réalisations du Consortium pour l'année 2014, il faut rappeler les difficultés d'organisation auxquelles nous avons été confrontés, et leur impact sur l'activité du Consortium.

D'abord les désaccords survenus en début d'année au sein du Consortium autour des objectifs et des manières de procéder ont considérablement retardé l'avancement du projet. Les partenaires n'ont pas été en mesure d'avancer de leur côté avant qu'une clarification des rôles, de la coordination et des objectifs soit effectuée. Cette clarification a été largement entreprise depuis lors.

Par ailleurs, il faut souligner un certain nombre de difficultés structurelles, déjà mentionnées en 2013, concernant les possibilités d'implication des personnels locaux au sein des laboratoires. Comme cela a déjà été dit, les corpus ne préexistent pas à l'intervention du Consortium. Dès lors avant de pouvoir numériser les enquêtes, il faut pouvoir les repérer, convaincre les chercheurs (qui le plus souvent disposent des documents), rassembler les matériaux et les traiter. Or dans les faits, il s'avère impossible de confier ces tâches à des personnels temporaires, comme notre expérience préalable l'a montré. Des solutions ont pu être mises en place sur certains sites, comme à Triangle, Pacte, au Centre Émile Durkheim, ou même à l'OSC et au CEVIPOF, avec du personnel permanent. Cette solution consiste à assortir la participation des laboratoires au Consortium à la recherche de ressources locales, du côté des centres de documentation et bibliothèques, qui puissent consacrer une partie de leur temps au projet. Mais cela reste compliqué, du fait des contraintes propres aux configurations locales. En effet, certains partenaires – à Bordeaux, Grenoble et Lyon - ont été confrontés à des déménagements - ou vont l'être - qui mobilisent massivement les personnels, et notamment les documentalistes et archivistes, qui n'ont alors plus le temps de s'occuper d'autres projets.

Le programme pour l'année 2014, tel qu'élaboré dans le dernier rapport pour le conseil scientifique fin 2013, prévoyait les actions suivantes :

- 1- Inventorier et collecter des enquêtes dans les labos

- 2- Numériser des enquêtes, d'après une pré-liste
- 3- Gérer les métadonnées / notices d'enquêtes
- 4- Créer une plate-forme d'auto-archivage

Les réalisations effectives pour l'année 2014 seront présentées en reprenant ces quatre postes.

1 - Avancement sur l'inventaire, la collecte et la documentation d'enquêtes

L'année 2014 a commencé par une visite de Sophie Duchesne à PACTE. Cette visite a permis une prise de contact avec Céline Gaufreteau qui a rejoint l'équipe d'archiPolis ainsi que d'organiser le travail à PACTE et avec le CDSP. L'objectif était de reprendre et de compléter l'inventaire déjà entrepris par Laetitia Rouvière (en tant que vacataire) l'année précédente. Il était également prévu de procéder à l'archivage de l'enquête de thèse de Céline Belot (*Jeunes Français et Britanniques dans le processus de légitimation de l'Union européenne, 2000*), c'est-à-dire de réaliser l'inventaire, la collecte et le classement de l'enquête. Céline Gaufreteau n'a cependant pas pu aller plus loin, s'étant vu confier des responsabilités supplémentaires dans le cadre du déménagement de l'IEP. Néanmoins, l'inventaire déjà réalisé a permis de localiser quelques enquêtes pour lesquelles leurs auteurs ont donné un accord de principe pour un archivage : notamment une enquête de Jacques Barrou sur l'immigration, et une enquête de Dominique Labbé et Dominique Andolfatto sur le syndicalisme. Ce travail sera repris en 2015.

L'inventaire réalisé à Triangle à partir de 2013 et actualisé en 2014 a été déposé sur l'Intranet mais aucune collecte n'a encore été amorcée pour des raisons précipitées liées aux difficultés rencontrées au sein du Consortium. Néanmoins, un travail de fond a été mené autour du référencement des publications et plus largement sur les métadonnées.

Au Centre Émile Durkheim à Bordeaux, trois enquêtes ont été collectées et sont conservées à la bibliothèque. Il s'agit d'une enquête de Thierry Berthet, une de Jacques Palard et une troisième enquête d'Andy Smith. Néanmoins, comme nous l'avons évoqué plus haut, la bibliothèque étant en plein déménagement, ces activités ont été suspendues et reprendront lorsque la nouvelle bibliothèque sera installée.

Au Ceraps de Lille, nous sommes clairement en manque de personnel depuis le départ de Julie Testi (en tant que vacataire) qui a réalisé un inventaire des enquêtes très complet à partir des rapports d'activités du laboratoire. Aucune collecte n'a pu être faite.

L'inventaire déjà réalisé au CEVIPOF a permis d'identifier au moins deux enquêtes pouvant être prises en charge. Il s'agit d'abord de l'enquête *Représentations du champ social, attitudes politiques et changements socio-économiques* de J.-M. Donegani, G. Michelat et M. Simon. En partenariat avec le CEVIPOF et le CDSP, cette enquête a été collectée, inventoriée, classée et envoyée à la numérisation. Un travail de documentation a été également entrepris. Une deuxième enquête, sur les *Femmes en politique* de Mariette Sineau, a été collectée et inventoriée ; à la date de remise du présent rapport, elle est en cours de classement et de préparation pour la numérisation.

Enfin, l'inventaire réalisé à l'OSC a permis d'identifier au moins une enquête, celle d'Agnès Van Zanten sur le *Choix de l'école*. Via un partenariat entre l'OSC et le CDSP, cette enquête a été collectée, inventoriée, classée et préparée pour la numérisation. Un travail de documentation a été également entrepris.

En ce qui concerne l'ISP, Catherine Gaspare a rencontré une quinzaine de chercheurs et, munie de l'enregistrement des entretiens, devait renseigner la grille d'inventaire. Entre temps, elle a quitté l'ISP pour une UMI à Los Angeles, et devait poursuivre ce travail là bas. L'ISP ayant quitté le Consortium, cet inventaire ne sera finalement pas intégré à la liste compilée des inventaires archiPolis. Toutefois, il devrait être valorisé sur le nouveau site web de l'ISP, lancé à la rentrée 2014.

2- Avancement sur la numérisation d'enquêtes

La numérisation d'enquêtes a été rendue possible par l'élaboration d'un cahier de numérisation par l'équipe du CDSP en lien avec la bibliothèque de Sciences Po Paris. La rédaction d'un cahier de numérisation a été entreprise à la fin de 2013 et s'est prolongée pendant toute l'année 2014 en étroit lien avec une série de tests menés avec un prestataire à partir du printemps 2014, en se basant sur l'enquête *Représentations du champ social, attitudes politiques et changements socio-économiques* de J.-M. Donegani, G. Michelat et M. Simon. À la date de remise du présent rapport le prestataire est en train de finir le travail, et il s'agira dans la foulée de procéder au contrôle de numérisation.

Par ailleurs, l'enquête d'Agnès Van Zanten sur le *choix de l'école* a permis d'approfondir cette phase de test, d'affiner les procédures et l'évaluation des coûts de numérisation. Ce travail de préparation a été mené à partir de la fin du printemps, et l'enquête a été envoyée fin octobre au prestataire.

On remarquera que ces deux enquêtes ont été identifiées, préparées et numérisées en partie grâce au réseau archiPolis ; cependant le financement de la prestation a été pris en charge sur le budget propre de la bibliothèque de Sciences Po Paris.

Une troisième enquête est en cours de préparation en vue de sa numérisation, celle de de Mariette Sineau, sur *les femmes en politique*. La dépense liée à l'opération sera engagée d'ici le 30 novembre pour une prestation devant être réalisée en décembre.

3- Avancement sur la gestion des métadonnées et la diffusion des notices d'enquêtes

Le travail de normalisation des métadonnées, engagé en 2013, a été poursuivi en 2014.

Cela s'est traduit par :

- une mise à plat de l'ensemble des métadonnées utilisées
- une mise en conformité avec les divers standards impliqués (DDI, EAD, DC, etc.)
- une rationalisation du schéma de données et la mise en place d'un format « maison » combinant les éléments les plus pertinents de chaque standard utilisé et servant de format pivot pour les conversions d'un standard à l'autre.

En remplacement de la grille Excel, un prototype d'application a été conçu et commencé à être implémenté : il consiste en un masque de saisie des notices d'enquêtes avec l'utilisation de listes de valeurs contrôlées, basé sur le modèle de données décrit ci-dessus.

Tout ceci a été mis en place dans le but de faciliter le travail réalisé à archiPolis, à plusieurs niveaux pour :

- homogénéiser les saisies de notices d'un laboratoire à un autre
- permettre l'interopérabilité entre les différents dispositifs partenaires auxquels archiPolis sera lié : beQuali, CINES, logiciels d'archives... .

L'objectif est qu'avec ce modèle de données, ce système puisse importer et exporter des informations qui suivent ces différents standards.

La partie proprement informatique de ce travail a été réalisée au sein du CDSP par un développeur recruté sur la subvention DIME-SHS .

4- La plate-forme d'auto-archivage : un projet prématuré, mais riche d'enseignements

Le souci de faciliter la conservation des enquêtes qualitatives récentes et en cours, nativement numériques, a conduit en 2013 le Comité de pilotage à s'orienter vers la création d'une plate-forme d'auto-archivage mise à la disposition des chercheurs.

Suite à cette décision, un travail de réflexion et une série de tests ont été conduits au CDSP dès les premiers mois de 2014 en vue de l'élaboration d'un cahier des charges pour la réalisation de cette plate-forme.

Une réunion destinée à approfondir ce cahier des charges a été organisée en janvier 2014 au CDSP, de manière à préparer l'ensemble des points devant être gérés par l'application. Ont été passés en revue les points suivants : utilisateurs et rôles utilisateurs, modalités de dépôt, modalités de saisie des inventaires archiPolis, classement et restitution du plan de classement, modèle de données, fonctionnalités en termes de recherche et de tri, politique en terme de versioning, limitation du volume de dépôt, lien avec l'application de commande du Centre Quetelet, gestion de séries d'enquêtes, gestion des identifiants uniques et liens entre documents, stratégie et déroulement du projet. Cette réunion a servi à rédiger une version plus précise du cahier des charges.

Un séminaire a ensuite été organisé à Paris le 24 mars 2014 (rassemblant des chercheurs membres d'archiPolis ainsi que de l'équipe du CDSP) afin d'engager une réflexion sur la conception des besoins des chercheurs pour permettre de finaliser le cahier des charges, notamment pour des spécifications fonctionnelles et une politique d'archivage. Une série de problèmes ont été abordés et discutés : modalités de dépôt, degré de classement possible pour le chercheur (offert / imposé), nommage des fichiers, périmètre documentaire, modalités de l'exposition des données, sélection des enquêtes versus auto-archivage, sort des enquêtes non mises à disposition via beQuali, etc.

En parallèle, un important travail de veille et de test des outils existants a été mené au CDSP par le développeur DIME-SHS. Selon une procédure de benchmark ont été testées des solutions comme Alfresco, Fedora Commons (interface Hydra Project), Data Verse, E-Prints etc.

Mais ce projet a été confronté à d'importantes difficultés de faisabilité de toute nature – difficultés techniques et coûts du projet, mais aussi incertitude sur l'adhésion des chercheurs à ce projet, compatibilité du principe de l'auto-archivage avec l'objectif initial d'un archivage raisonné et scientifiquement construit, etc. Une réflexion s'est engagée à partir de la fin du

printemps 2014 sur la pertinence et la faisabilité du développement d'une telle plate-forme d'auto-archivage. Finalement, le Comité de pilotage du 10 octobre dernier a estimé que l'auto-archivage était actuellement prématuré compte tenu de l'état de l'art en matière d'archivage des enquêtes et des habitudes de travail des chercheurs et pose trop de difficultés techniques pour pouvoir être traité dans des conditions sereines. Le Consortium, suivant la recommandation du Conseil scientifique d'Huma – Num, a décidé de sursoir à ce projet de repository pour se recentrer sur le projet fondateur d'archiPolis de repérage, d'archivage et de numérisation d'enquêtes qualitatives dans les laboratoires.

La somme importante engagée sur ce projet (20 K€) ne permet certes pas d'atteindre les objectifs initiaux, mais le travail ainsi financé a permis d'identifier et de commencer de traiter un ensemble de problèmes inhérents à tout archivage numérique. En effet, l'essentiel du travail engagé concernait la saisie des métadonnées pour les inventaires et la mise en place de procédures au sein d'une application partagée. S'il a été décidé de sursoir à l'auto-archivage par les chercheurs, l'objectif reste de mettre en commun des descriptions d'enquêtes réalisées par les différents partenaires du réseau. La préparation du cahier des charges a permis de définir les besoins décrits dans les paragraphes précédents (rôles utilisateurs, champs de métadonnées pour la description au niveau de l'enquête, spécifications fonctionnelles du formulaire de saisie etc.). La prestation informatique constitue donc un apport important pour la mise au point des cahiers des charges qui vont être produits pour encadrer l'inventaire, la numérisation et l'archivage des gisements d'enquêtes dans les laboratoires partenaires.

Consolidation du réseau et échanges entre les partenaires

La consolidation du réseau a commencé à être mise en place en 2013. En février 2014, un point à distance, mobilisant l'ensemble des personnels impliqués localement dans les opérations d'inventaire et de collecte, a été organisé sous la forme d'une vidéo conférence. Par la suite, les problèmes d'organisation précités ont considérablement ralenti les échanges entre partenaires. Le dernier comité de pilotage a permis de relancer la dynamique, en organisant des groupes de travail thématiques pour 2015.

Par ailleurs, le Consortium a bien noté que le Conseil scientifique d'Huma-Num recommandait à archiPolis de s'ouvrir à d'autres pôles d'archives de chercheurs. Les

difficultés de l'année 2014 ont également freiné cette démarche. La constitution d'un Conseil scientifique et technique (CST) propre à archiPolis va fournir l'occasion d'initier de telles collaborations en France mais sans doute aussi à l'étranger.

Informations sur le budget 2014

Du fait des retards et difficultés rencontrés, seule une partie du budget 2014 (la dotation d'Etat) a pu être utilisée. Par ailleurs, le financement de la numérisation des enquêtes de Guy Michelat et d'Agnès Van Zanten a été effectué sur le budget de la bibliothèque de Sciences Po Paris, ce qui a limité les coûts pour archiPolis. Le travail réalisé par le développeur du CDSP, concernant aussi bien les activités de veille et de test que les premiers développements du masque de saisie des notices d'enquêtes, a d'abord été financé sur le budget DIME-SHS ; il a ensuite été remboursé à DIME-SHS comme une prestation, à hauteur de 20 000 €.

Du matériel de conservation a été commandé pour les partenaires, de manière à pouvoir archiver physiquement les enquêtes dans de bonnes conditions localement ; il s'est agi ici de suppléer les labos qui ne peuvent pas toujours financer ce type d'opérations.

L'ensemble de la subvention d'Etat 2014 a été dépensé. En revanche, compte tenu des difficultés rencontrées dans l'animation scientifique du Consortium, nous demandons à Huma-Num de pouvoir reporter l'ensemble des fonds propres d'archiPolis de 2014 sur l'année 2015 (environ 33 K€) .

Programme d'action 2015

Confrontés à une situation très difficile de crise, le Comité de pilotage a œuvré de façon à trouver une issue acceptable pour les différentes parties en litige avec la préoccupation constante de préserver et de poursuivre le travail impressionnant conduit sous la responsabilité de S. Duchesne. Cette volonté trouve ses premières concrétisations dans le changement de porteur du Consortium qui revient au CDSP et lors de la réunion du Comité de pilotage du 10 octobre dernier qui a désigné un nouveau coordinateur et remobilisé les acteurs impliqués en posant les grandes orientations de travail pour l'année à venir.

L'enjeu principal pour les prochains mois réside dans la relance des opérations au sein des laboratoires partenaires afin d'aboutir dans un délai raisonnable, compte tenu des moyens

disponibles, à la constitution d'un fond d'archive conséquent et harmonisé des enquêtes qualitatives de ces laboratoires. Pour cela, il faut avancer simultanément dans plusieurs directions toutes interdépendantes :

- clarifier les relations entre archiPolis et beQuali ;
- définir les standards utilisés par archiPolis et beQuali dans toute la chaîne de production d'archives ;
- traduire ces standards dans des applications informatiques ;
- former les utilisateurs de ces outils ;
- impliquer autant que possible les services de documentation des institutions partenaires dans ce processus d'archivage ;
- remotiver les chercheurs des laboratoires partenaires pour qu'ils participent au projet ;
- les sensibiliser aux enjeux de l'archivage des données d'enquête en cours et à venir ce qui signifie aussi diffuser les bonnes pratiques de recherche qualitative.

Tous ces points, qui recourent largement les remarques faites par le CS d'Huma-Num, ont été abordés lors du Comité du 10 octobre. Des groupes de travail thématique se mettent en place pour les traiter autour de quatre chantiers principaux qui seront présentés après avoir fait le point sur la nature des relations entre le Consortium, le CDSP et beQuali.

1 – Relations entre archiPolis et beQuali

Ces deux structures restent bien distinctes, même si maintenant elles sont portées toutes les deux par le CDSP. Le fait que le coordinateur scientifique n'appartienne pas au CDSP traduit l'autonomie des deux projets.

ArchiPolis fédère plusieurs laboratoires travaillant sur l'objet politique afin d'organiser un repérage et un archivage standardisé de leurs enquêtes patrimoniales, en cours et à venir, avec les informations contextuelles nécessaires à leur ré-analyse.

BeQuali, au sein de l'Equipex DIME SHS, a pour fonction de mettre à disposition des chercheurs des enquêtes – et leurs méta-données - préalablement sélectionnées par son CST pour leur intérêt scientifique, à des fins d'analyses secondaires.

Sur le plan des thèmes scientifiques, le champ d'archiPolis est plus étroit que celui de beQuali ; sur le plan des gisements de données, il est plus large puisqu'il couvre toutes les enquêtes qualitatives d'un laboratoire, dès lors qu'elles présentent une documentation suffisante pour nourrir des archives (retranscription d'entretiens ou enregistrement et méta-données de contextualisation).

La diffusion des données n'est pas dans la mission d'archiPolis. Le Consortium travaille à la constitution d'un vaste gisement de données qualitatives, structuré et documenté, dans lequel beQuali viendra puiser pour nourrir son propre catalogue d'enquêtes mises à dispositions de la communauté scientifique. Il est donc essentiel que les deux projets travaillent en étroite synergie pour concevoir et adopter les mêmes standards et normes d'archivages ou du moins que ceux-ci soient totalement compatibles entre eux. Le travail de réflexion sur la constitution de l'archive (son contenu, sa forme, etc.) et les outils informatiques qui en découlent doivent donc être conduits en commun.

Pour tenir compte de cette architecture complexe au plan de l'organisation interne, le Comité de pilotage a demandé la création d'un bureau exécutif d'archiPolis, inscrit dans la nouvelle convention. Il se compose du coordinateur scientifique d'archiPolis, d'un représentant de beQuali et du directeur du CDSP. Il gère les questions administratives, la signature de documents, la production de rapports scientifiques et financiers, la préparation des réunions et la gestion des relations institutionnelles.

Le Comité propose également de désigner en son sein un ingénieur et un chercheur qui seraient invités permanents du CST de beQuali, sans droit de vote. Cette proposition doit toutefois être acceptée par le CST de beQuali qui se réunit le 11 décembre prochain.

2-Chantier inventaire, collecte et documentation

Il est urgent de relancer le travail d'inventaire, de collecte et de documentation des enquêtes dans les laboratoires. Toutefois cette opération, déjà bien avancée, bute actuellement sur l'absence de consignes claires relatives à ce qui doit être inventorié et aux modalités de réalisation de cet inventaire. Le tableau Excel utilisé n'est pas adapté à un travail de cette ampleur, avec l'exigence scientifique qu'il suppose.

Il est donc urgent de travailler dans deux directions :

- mieux définir la liste des documents nécessaires à la contextualisation de l'enquête et leur plan de classement, en en tenant compte à la fois des impératifs d'exigence scientifique et des contraintes matérielles rencontrées sur le terrain
- se doter dans les meilleurs délais d'une plateforme d'inventaire qui permette de réaliser des inventaires standardisés au niveau national, interopérables en termes de formats, pouvant être aisément réactualisés. Dans chaque laboratoire partenaire les sauvegardes des enquêtes doivent être standardisées (contenu et mise en forme). Un cahier des charges, qui s'appuie notamment sur l'expérience de beQuali et sur celle tirée du travail sur le développement d'un auto-archivage, est en cours de rédaction pour la production de cette application.

Ce travail préalable fait, il sera possible de commencer à procéder à l'archivage physique des enquêtes.

Pour cela, il est prévu d'amener chaque site à travailler en lien avec ses bibliothèques et ses centres de documentation et d'archivage. Dans les semaines qui viennent, des contacts vont être pris avec ces services dans chaque institution.

L'objectif global est de professionnaliser au maximum ce travail de « ratissage » des données à archiver en recourant à des outils partagés adéquats et en impliquant, autant que faire se peut, un personnel compétent et stable. Ce personnel devra être formé à ce travail d'inventaire, de collecte et de documentation.

3- Chantier numérisation

Nous souhaitons passer à une phase plus active en matière de numérisation d'enquêtes en testant notamment la solution de gestion en partie délocalisée de la numérisation, en partenariat avec le CDSP, ce qui implique éventuellement d'adapter le cahier des charges, ou du moins de former aux protocoles les personnels chargées de ce travail dans les laboratoires.

4-Chantier gestion des métadonnées / notices d'enquêtes

Nous devons terminer :

- la modélisation et la correspondance des métadonnées ;
- la révision du cahier des charges pour l'outil de saisie des inventaires ;
- le développement informatique de ce dernier et sa mise en production.

Pour cela il nous faut également régler la question de l'hébergement de cette application et des droits d'administrations des serveurs.

5- Chantiers formation et valorisation

Toutes ces opérations techniques appellent un accompagnement des différents acteurs à travers des actions de formation et de valorisation.

- La formation concerne déjà les personnels de recherche du Consortium et des centres de documentation impliqués afin qu'ils s'approprient les différents outils qui viennent d'être décrits : plan de classement, préparation des archives, conservation des archives papiers, numérisation, ...
- Elle concerne aussi les chercheurs. La mission d'archiPolis ne se réduit pas à une seule opération d'inventaire et d'archivage. Le gisement de données créé en 2015 doit être ensuite alimenté de façon continue par de nouvelles enquêtes. Pour cela, il faut que les chercheurs prennent conscience de l'importance de la préservation de leurs matériaux empiriques. Cette évolution doit être aidée par la formations des chercheurs pour les sensibiliser aux dimensions juridiques des enquêtes (protection des enquêtés, sécurisation des données, propriété intellectuelle, propriété des données, archives publiques, etc.); mais aussi par une formation aux bonnes pratiques de recherche qualitative qui facilitera l'inventaire et l'archivage des données produites. Ces opérations de formation des chercheurs, sans être au cœur de l'activité du Consortium, sont en lien direct avec ses objectifs. Elles devront être conçues et conduites en collaboration avec beQuali.
- Enfin la valorisation de l'activité d'archiPolis, par le carnet Hypotheses.org créée en 2014, par des publications scientifiques, des articles dans les supports de communication du CNRS, des kit de démonstration de l'archivage des données et de leur possible réutilisation via beQuali, etc. permettra de donner une visibilité concrète au travail du Consortium et de montrer aux chercheurs l'intérêt scientifique de leur participation à ce projet.

Budget prévisionnel 2015 (Grandes masses) :

Pour mener à bien le projet, il faut s'appuyer sur de la numérisation, de la prestation informatique et des frais divers liés (frais de mission et valorisation).

Il est encore difficile de chiffrer précisément le coût prévisionnel d'une numérisation d'enquête. En effet cela dépend à la fois de la nature du corpus et de son hétérogénéité, plus que de son volume. Néanmoins nous savons que la numérisation d'une enquête dont les entretiens existent uniquement sur support papier et sont très annotés pourra atteindre près de 10 000€ TTC. En effet, toute annotation nécessite un traitement manuel qui entraîne un surcoût de plus de deux euros par page. De la même façon, étant donné que le cahier des charges n'est pas encore finalisé, il est difficile à ce stade de chiffrer très précisément la prestation informatique.

Selon nos estimations nos besoins sur la subvention d'état, sont les suivants :

- 30 000 € pour la numérisation d'enquêtes ; le nombre d'enquêtes à numériser en 2015 dépendra du coût, estimé entre 6 000 € et 10 000 € selon la nature du corpus (son hétérogénéité, matérialité du support, etc.).
- 20 000 € pour la prestation informatique nécessaire au développement d'une application de saisie des inventaires ; cette somme correspond à quatre mois de salaire chargé d'un développeur.
- 10 000 € pour l'animation et le fonctionnement du Consortium et la valorisation de son activité.

Les ressources propres seront affectées en priorité au financement des renforts en personnels.