

OPERATION DE RECHERCHE N°6 : questionnaire		
Questions	Réponses	Explications sur les champs à remplir
1) Questions d'ouverture : acteurs et objectifs du projet?		
Présentation globale (énoncé du projet)		
	L'informatique administrative en France : instruments, acteurs, savoirs (années 1950-1980)	
Quelle est la composition de l'équipe du projet ?		décrire
personnes	D. Dompnier, P.Y. Baudot, H. Mazoyer, A. Healy	
profils : chercheurs, ingénieurs...		
équipes	Triangle	
disciplines		
établissements		
sites	Lyon, Paris?	
pays		
rôles de chacun		
Apparat Critique :		
Objectifs ? Quels types d'informations voulez-vous extraire des documents sources ?	Objectif : comprendre comment l'informatique s'insère dans le eu administratif. Comment il en reproduit les luttes classiques et les rapports de force qui le structure. Comment les techniques représentent une force d'inertie spécifique ? Quelles compétences , quels savoirs, quelles expertises sont nécessaires pour maîtriser ces techniques? 1) Analyse de la création et de l'institutionnalisation de lieux d'expertise entre savoirs et pouvoirs (exple : SEM Icare fondée en 1970 qui tiendra une opération d'informatisation concurrente à celle du Ministère de l'Intérieur). 2) Analyse des nouveaux rapports induits par l'informatique entre centre et périphérie (exple : informatisation des procédures électorales qui entraine une lutte autour de la constitution du fichier électoral entre INSEE/Ministère de l'Intérieur/Préfectures VS les Communes qui souhaitent protéger la vie privée des électeurs).	
	Où trouver les sources : 1) Archives de la SEM Icare conservées aux Archives municipales cote 2023W et Entretiens. 2) Archives départementales nationales, Versements au centre des archives contemporaines de Fontainebleau relatifs à l'informatisation du fichier électoral (versements 20050222, 19970063, 19920649), listes bibliographiques de travaux universitaires et de travaux d'experts.	
	Types de matériaux à analyser : travaux universitaires, travaux d'experts, archives, entretiens, base de données	
	Types de productions de l'équipe (résultats attendus) : Ouvrage collectif? Ré-édition des archives critique papier? Ré-édition critique numérique? Base de données? Encodages des archives? Mise disposition des archives en ligne?	

souci de conservation du document sous sa forme originelle? (mode image? mode texte?)			
besoin d'analyser le texte ?			
fonctionnalités de recherche ?			
quel mode de consultation ? unique ou plusieurs parcours de lecture...			
-- Quel mode d'accès ? ouvert/restreint?			
Quel rythme de publication ?			
-- publication régulière ou publication d'un produit fini?		planifier	
-- dates de début et de fin du projet ?	2009-2013?		
Organisation du travail envisagée ?			
-- modalités d'intervention et de collaboration sur les documents ? Travail collaboratif ?			
-- mémoire, traçabilité des modifications ?			
Existe-t-il des influences extérieures à la tenue du projet?			
devez-vous respecter un standard défini pour la communauté scientifique à laquelle vous appartenez ou à laquelle appartiennent les ingénieurs du projet?	Utilisation antérieure du format XML-TEI pour l'encodage de l'Almanach national (1750-1850).		
Outils, supports informatiques envisagés ?			
-- outil de travail (back office) / outil de publication (front office)	utilisation antérieure de 4D. Utilisation antérieure par une autre équipe (LAHRHA) de la plateforme Millefeuille (Environnement de développement Eclipse) pour l'encodage de l'Almanach national (1750-1850). Encodage TEI-XML dont le but était d'analyser les documents historiques pour étudier la structure de l'administration française au XVIII-XIXe (mise en valeur des déplacements de la sphère gouvernementales à la sphère administrative).		
Ressources disponibles ?			
-- matériel			

-- ressources humaines (collaborateurs technologiques...)			
-- financements			
-- interlocuteurs principaux ?			
Quel(s) est (sont) le(s) public(s) visé(s) ?			
enseignement ? universitaire ? scolaire ?			
grand public ?			
Documents, sites web relatifs ou inspireurs au/du projet ?			
2) Constitution et Analyse de l'ensemble des documents (du corpus) : choix, types et structure des doc, documentation et conservation du corpus source			
Description des document(s) source(s)			décrire le corpus
Le corpus est-il déjà constitué ?	cf. liste des archives sélectionnées pour l'étude.		
quels critères de choix ont présidé à la constitution du corpus?			délimiter le corpus
un ou plusieurs documents ? (volumétrie)	plusieurs		
titre(s)			
un ou plusieurs TYPES de documents/données : Textes, images, sons...	Documents papier / texte : archives (SEM, fichiers électoraux, etc.) ET Document audio/video? : Entretiens avec des agents		recenser les typologies de document
un ou plusieurs FORMATS/SUPPORTS à disposition, qualité, état des documents...:			
--docs papier : texte en écriture manuscrite, caract. d'imprimerie non standard	Docs papiers : textes en écriture manuscrites, textes en caractère d'imprimerie standard		
--docs papier : texte en caractères d'imprimerie standard, livres, "tapuscrit" ;			
--fichiers numériques images : tif, gif, png, jpg, pdf (texte non sélectionnable)			
--fichiers numériques texte : word, odt, pdf avec texte sélectionnable			
Histoire du/des documents			
période(s)			Appréhender l'histoire d'un document pour mieux :
thématique(s)			en comprendre l'évolution
genre(s)			le replacer dans un contexte
inédit(s) ou réédition ?			
à partir d'un manuscrit ou d'un imprimé ?			
une ou plusieurs versions ?			

si plusieurs version(s) / traduction(s) : édition d'une seule version ou plusieurs ?			
Droits			
		connaître et/ou vérifier les droits de manipulation, de modification et de diffusion d'un document	
Quelle est la provenance des documents ? (fonds d'archives, bibliothèques...)	1) Archives de la SEM Icare conservées aux Archives municipales cote 2023W et Entretiens. 2) Archives départementales nationales, Versements au centre des archives contemporaines de Fontainebleau relatifs à l'informatisation du fichier électoral (versements 20050222, 19970063, 19920649), listes bibliographiques de travaux universitaires et de travaux d'experts.		
La question des droits de propriété intellectuelle a-t-elle été réglée ?			
Stockage			
Quels sont les lieux et modes de conservation actuels ?			
Existe-t-il des sauvegardes du corpus source ?			
3) Description, indexation du corpus			
Le corpus a-t-il déjà été décrit ?		catalogage : métadonnées, inventaire : EAD	
si oui, sous quel format ?		base de données, fiches papier...	
si oui, avec quel norme ?		EAD, UNIMARC...	
est-il possible de récupérer les notices existantes?			
Si votre corpus est à décrire :			
Comment décririez-vous la structure de vos documents ?		champs	
Quels éléments clefs de cette structure souhaitez-vous garder ?			
Quels types d'infos clefs relevez-vous dans vos documents ?			
4) Numérisation du corpus			
Description des tâches de la numérisation			
Le corpus a-t-il déjà été numérisé ?	non?	numérisation : formats, matériel	
est-il ou sera-t-il numérisé en interne ou en externe ?			
sous quel format ?			
avec quels outils : matériel+logiciel ?			
stockage des images ?			
existe-t-il des notices décrivant ces images ? Si oui est-il possible de les récupérer ?			

Exploitation et manipulation des images obtenues ou à obtenir lors de cette numérisation		
à quoi serviront les images par la suite?		exploitation : transformation vers du texte? illustrations imagées? objets de commentaire?
nettoyage		traitement des d'images, reconnaissance de forme
recadrage ?		
changement de taille ?		
enregistrement à différents formats ?		
OCRisation ?		
5) Transcription du corpus		
le corpus a-t-il déjà été transcrit ?		saisie? OCRisation? reconnaissance de forme...,
si oui, par qui ?		interne/prestataire externe
si oui, sous quel format ?		*.doc, *.txt, etc...avec encodage ?
si oui, avec quel codage de caractère ?		
si oui, avec quelles règles de transcription ? comment sont-elles signalées ?		(accents, ponctuation, orthographe, mise en page...) annotations de forme (texte souligné, biffé...)
Non, donc transcription à faire :		
Comment sera-t-elle effectuée ?		saisie? OCRisation? reconnaissance de forme...,
Par qui ?		interne/prestataire externe
Sous quel format ?		*.doc, *.txt, etc...avec encodage ?
codage des caractères de la transcription		iso-**-*****-**.
mises en forme éditoriales de la transcription :		
-- quel(s) type(s) de transcription est(sont) prévu(s) ? copie exacte... ?		diplo, etc.
-- quelles règles de transcription seront adoptées ?		(accents, ponctuation, orthographe, mise en page...)
-- comment seront-elles signalées ?		annotations de forme (texte souligné, biffé...)
6) Enrichissement : marquage, puis apparat critique du corpus / de(s) document(s)		
Enrichissement grâce à un marquage/descriptif des documents (textes, images, etc) sources		utilisation de descripteurs standards
-- quel encodage ? norme ? standard ?		tei, dc, iptc

Apparat Critique :			
enrichissement de type annotation ? quel type d'annotation ?		Définition de types d'annotations récurrents : annotations « de forme ou de style », contextuelles, historiques, biographiques, bibliographiques, sémantiques...	
enrichissement sous forme d'annexe ?			
enrichissement sous forme d'analyse quantitative...			
enrichissement lexical (lien vers des ontologies [thésaurus?]?)		Liens vers des documents de référence pour l'annotation (dictionnaire de langue, dictionnaire biographique...)	
par qui ? lecteurs ?		Définition de domaines, de points de vue, réseaux de description ? Porphyry, Arcane...	
enrichissement visuel ?		Enrichissement de l'édition critique avec intégration de documents « annexes » : comptes rendus, illustrations...	
choix des liens entre les différents documents du corpus		ex : entre une image du texte et le texte transcript	
Mise en évidence et exploitation des variantes (plusieurs versions, traductions)			
comment seront-elles présentées, comparées pour l'élaboration de l'édition ?		choix d'un texte pivot, découpage de texte, alignement, lecture en parallèle ou arborescence, collation, multifenêtrage, présentation des variantes sous forme d'annotations, de « marqueurs » (couleurs différentes par ex), représenter l'historique du document, dresser le tableau des relations de dépendance entre variantes, liens vers d'autres documents...	
comment seront-elles utilisées ?			
ajout, modification, suppression de texte (du mot au chapitre : quelle granularité ?), dans quelles versions ou éditions ?			
7) Publication en ligne du corpus (images, texte, appareil critique, etc.)			
Type de documents présents dans la publication en ligne :			
les images seront-elles présentes dans la publication en ligne?			
le(s) textes transcrits seront-ils présents dans la publication en ligne?		styles, mise en page, principes éditoriaux, choix d'une structure éditoriale	

existera-t-il des variantes d'un même texte dans la publication en ligne?			
des notes critiques seront-elles présentes (visible) dans la publication en ligne?			
des ontologies seront-elles présentes dans la publication en ligne?		Si oui laquelle?	
Choix éditoriaux et recherche dans la publication :			
comment seront présentées et/ou exploitées les images?		outils de visualisation (zoom, vignettes...)	
comment seront présentés et/ou exploités les textes?		choix éditoriaux	
comment seront présentées et/ou exploitées les notes		moteur de recherche, index, etc.	
comment seront présentés et/ou exploitées ontologies?			
quel type d'accès ? ouvert / restreint ?			
comment sera publié ce projet ? En continu, en une seule fois ?		classement arborescent, publication différée...	
Liens entre les éléments de la publication :			
comment seront présentés les liens entre les images et le texte?			
comment seront présentés les liens entre les textes (variantes par ex)			
comment seront présentés les liens entre les notes et le texte? les notes et les images?			
comment seront présentés les liens entre les ontologies et le texte?			
8) Stockage / pérennisation du corpus			
quels sont les modes de stockage et de sauvegarde prévus à chaque étape du projet ?		stockage et conservation (formats, évolutions, modèle OAIS)	
9) Valorisation de l'édition et de la gestion de projet			
Quels moyens seront utilisés pour faire connaître et valoriser cette édition ?		Référencement	
		Communication sur des sites institutionnels	
		RSS	
10) Evaluation			
Comment seront évalués les usages de cette édition ?			