

## **TGIR Huma-Num**

# **Bilan annuel consortium CAHIER**

**Consortium CAHIER**

**Corpus d'Auteurs pour les Humanités : Informatisation, Édition, Recherche**

Adresse Web : <https://cahier.hypotheses.org/>



Document rédigé par T. Lebarbé, F. Idmhand et L. Rageot

Relu et corrigé par les membres du Comité de Pilotage du Consortium CAHIER

Transmis pour avis au Conseil Scientifique du Consortium CAHIER

ERRATUM :

**Page 10 - 2.6 Les formations 2018**

b. ANF Aussois – participation au financement de la formation, conjointement avec l’Institut des Données, IDEX Grenoble, dans le cadre d’une formation soutenue par le service de la formation permanente du CNRS **au lieu de** ANF Aussois – participation au financement de la formation, conjointement avec l’Institut des Données, IDEX Grenoble, dans le cadre d’une formation de l’Agence Nationale de Formation du CNRS

# Table des matières

1	Les objectifs du consortium.....	3
1.1	Nouveaux objectifs : réalisations du consortium.....	4
1.2	Pilotage du consortium : .....	4
2	Activités du consortium en 2018.....	5
2.1	L'accueil des nouveaux membres.....	5
2.1.1	SATELLITES.....	5
2.1.2	Édition Numérique des Cahiers d'Henri de Régner (ENCHRE) .....	5
2.1.3	Les Archives d'Augustin Thierry (ArchAT).....	6
2.2	L'accueil de membres internationaux.....	6
2.3	Le soutien aux projets membres .....	6
2.4	Les activités de coordination .....	7
2.4.1	Les réunions du comité de pilotage .....	7
2.4.2	La réunion du conseil scientifique.....	8
2.4.3	L'Assemblée générale 2018 .....	8
2.5	L'atelier Annuel 2018.....	9
2.6	Les formations 2018 .....	10
2.7	Le soutien à la participation à des conférences internationales.....	11
2.8	Le soutien à des séminaires et journées d'étude.....	11
3	Le soutien aux groupes de travail .....	12
3.1	Groupe de Travail [Questions Juridiques].....	12
3.2	Groupe de Travail [EVENT].....	13
3.3	Groupe de Travail [Data Cahier].....	13
3.4	Groupe de Travail [R2Cahier].....	14
3.5	Groupe de Travail [Typologie].....	15
3.6	Groupe de Travail [Crowdsourcing].....	16
3.7	Mutualisation inter-consortium des activités des groupes de travail.....	16
4	Présentations des activités lors de réunions internationales .....	17
4.1	DARIAH Annual Event 2018, Paris.....	17
4.2	[DH2018] à Mexico.....	17
5	Description de l'apport aux communautés scientifiques .....	18
5.1	Plus-value apportée par l'organisation en consortium.....	19
5.2	Etat d'avancement .....	19
5.2.1	Objectif réseau : .....	19
5.2.2	Objectif formation : .....	20
5.2.3	Objectif réflexion et bonnes pratiques (GDT) : .....	20
5.2.4	Objectif soutien à la recherche : .....	20
5.3	Difficultés rencontrées.....	20
6	Informations sur le budget 2018.....	22
7	Programme 2019.....	24
7.1	Objectif formation : .....	24
7.2	Objectif soutien aux projets : .....	24
7.3	Objectif bonnes pratiques : .....	24
7.4	Objectif réseau : .....	24
7.5	Objectif avenir du consortium : .....	25
7.6	Budget Prévisionnel 2019.....	25

# 1 Les objectifs du consortium

Le consortium CAHIER s'est constitué en juillet 2011 sous l'impulsion de Marie-Luce Demonet (CESR, Tours). Initialement, il rassemblait 12 projets de recherche partageant des problématiques communes de constitution, d'exploitation et de valorisation numériques de patrimoines textuels dans un contexte qui était alors marqué par l'absence de structures de référence (institutions, sociétés savantes, etc.). CAHIER s'est alors engagé dans le soutien et l'accompagnement des projets numériques sur des corpus d'écrivains dans le but de promouvoir le renouvellement des méthodes et des pratiques scientifiques dans le domaine des sciences du texte, traditionnellement peu informatisé. Pour cela, CAHIER a soutenu, et soutient toujours, le développement de solutions techniques originales: il a formé et développé des outils numériques qui ont pu être testés et éprouvés par la "communauté" de CAHIER et par des communautés impliquées dans les Humanités Numériques en Europe et dans le monde (grâce aux différents réseaux de recherches dans lesquels sont impliqués ses membres).

Depuis sa création, CAHIER n'a cessé de croître. Face à l'augmentation du nombre d'adhésion, CAHIER s'est restructuré en 2017 de façon à simplifier son organisation avec un comité de pilotage élu par l'assemblée générale des membres et une coordination tricéphale composée de deux responsables et d'une gestionnaire.

L'adhésion à CAHIER est une adhésion sur projets: l'existence d'une « organisation » autour de projets est assez atypique dans le paysage de la recherche : elle met la proposition scientifique au cœur du réseau. Dans le cas des Humanités Numériques, une telle structure souligne l'inhérente trans-disciplinarité induite par ces types de projets et permet à CAHIER de transcender les ancrages disciplinaires, institutionnels et professionnels.

Pour être membre de CAHIER, les porteurs de projets adressent une demande d'adhésion aux coordinateurs du consortium et s'engagent à respecter quelques principes énoncés sur la page du site internet (<https://cahier.hypotheses.org/contacts>): avoir un projet d'édition de sources, utiliser des formats interopérables et s'engager à rendre les données ouvertes et accessibles. Actuellement 65 projets, soit plus de 300 membres, composent le consortium. Au sein d'un tel réseau, les équipes-projets trouvent un petit soutien financier destiné à les accompagner dans les premières étapes de leur "conversion numérique". Surtout, ils trouvent des espaces de réflexions et d'échange leur permettant de faire émerger des idées, des recommandations et de nouvelles épistémologies.

CAHIER est un consortium dynamique qui organise différents moments de rencontres tout au long de l'année permettant à ses membres de tirer profit de l'existence de ce réseau scientifique et des compétences de ses chercheurs et ingénieurs :

- une assemblée générale annuelle,
- un atelier de formation et de réflexion méthodologique annuel,
- des formations ponctuelles autour de techniques ou de technologies des humanités numériques des textes,
- et des rencontres plus spécifiques autour de groupes de travail que les membres constituent eux-mêmes à partir des questions techniques et/ou théoriques qui les préoccupent, et qui sont susceptibles d'intéresser l'ensemble des membres (voir ci-après).

Pour résumer, CAHIER est un réseau d'intérêt scientifique facilitant les collaborations nationales et internationales entre chercheurs et équipes projets: cette ouverture est l'une des grandes plus-values du consortium.

## 1.1 Nouveaux objectifs : réalisations du consortium

Depuis sa création en 2011, le consortium a visé la constitution d'un réseau de chercheurs et de bonnes pratiques en Humanités Numériques. Il poursuit cet engagement.

En sept ans (2011-2018), CAHIER est devenu un interlocuteur reconnu des discussions sur la construction et/ou le traitement des données informatiques en sciences humaines. Après une première étape dédiée à la mise en place des bonnes pratiques de l'informatisation des données (I de CAHIER) et de l'édition (E de cahier), c'est dorénavant le potentiel de recherche (R de CAHIER) et de création de nouveaux savoirs à partir des données constituées qui prédomine dans les activités du consortium comme le confirment les programmes de l'Assemblée Générale 2018 et de l'atelier thématique 2019 programmé en juin/juillet à Poitiers.

Par ailleurs, le réseau lui-même évolue, dans sa géographie, dans sa couverture thématique et en fonction des projets qui rejoignent le consortium. Le Consortium CAHIER évolue au plus près des tendances et besoins de sa communauté, avec les progrès des humanités numériques des sciences des textes.

## 1.2 Pilotage du consortium :

Le consortium a souhaité faire évoluer son organisation en 2017. Le fonctionnement initial disposait d'un comité de pilotage constitué d'un représentant de chaque projet membre. Lorsque le consortium a atteint les 70 projets, le comité de pilotage ne pouvait plus se réunir dans son ensemble. Un nouveau fonctionnement a donc été proposé aux membres :

- un Bureau constitué du coordinateur, de la coordinatrice adjointe et de la gestionnaire
- un Comité de Pilotage constitué de 20 membres élus parmi les projets membres du consortium
- une Assemblée Générale, réunie annuellement, constituée de représentants de tous les projets du consortium
- un Conseil Scientifique international, représentatif des domaines d'activité du consortium, constitué de quatre personnalités externes, désignées pour leur compétence par le Comité de Pilotage

La mise en place de ce nouveau mode de fonctionnement a été réalisée fin 2017 (après le précédent rapport).

La première intervention du Conseil scientifique a pu avoir lieu durant l'Assemblée Générale de décembre 2017. Le CS a porté un regard critique et intéressé. Le présent rapport lui est également remis en vue d'une intervention lors de la prochaine Assemblée Générale programmée le 23 novembre, à Paris.

## 2 Activités du consortium en 2018

### 2.1 L'accueil des nouveaux membres

En 2018 (jusqu'à la date de rédaction de ce bilan), quatre projets ont soumis une demande d'adhésion. Trois ont été acceptés par le bureau et un projet a été refusé. La liste des projets membres est tenue à jour sur notre carnet d'hypothèses :

<https://cahier.hypotheses.org/membres>

Les trois projets acceptés sont respectivement :

#### 2.1.1 SATELLITES

**Responsable:** Fatiha Idmhand (Institut des Textes et Manuscrits Modernes - Archivos)

**Résumé synthétique :** Satellites – Les intellectuels « satellites ». Un autre regard sur la circulation des idées

Le projet "Satellites" étudie les échanges et la circulation des objets culturels durant les principaux conflits du monde hispanique (1923-1980). Dans une perspective transatlantique (entre l'Europe et les Amériques), "Satellites" examine tout particulièrement le rôle des "médiateurs" des transferts culturels (Michel Espagne), et plus particulièrement, des intellectuels qui ont agi à l'ombre des figures les plus connues : les intellectuels "Satellites". Moins médiatiques et donc mal connues, « Satellites » propose de les redécouvrir à travers l'exploitation de leurs archives, pour mesurer de quelles façons ils ont contribué à la diffusion des connaissances d'une zone culturelle à l'autre.

#### 2.1.2 Édition Numérique des Cahiers d'Henri de Régnier (ENCHRE)

**Responsable:** Bernard Roukhomovsky (Litt&Arts – Arts et pratiques du texte, de l'image, de l'écran et de la scène)

**Résumé synthétique :** Les milliers de notes consignées par Henri de Régnier, durant un demi-siècle (1887-1936), sur ses cahiers intimes – et, de façon plus ponctuelle, sur ses carnets de voyage et calepins divers – n'étaient pas destinées à la publication. C'est à bien des titres cependant qu'elles se recommandent à l'attention. Témoignage éminemment précieux sur la vie de celui qui fut l'une des figures les plus en vue de la génération symboliste et de la Belle Époque, mais aussi sur la vie littéraire, artistique et mondaine de son temps, elles forment un objet textuel au statut singulier.

Le projet ENCHRE (Édition numérique des Cahiers d'Henri de Régnier) a pour ambition de mettre à la libre disposition du public – et spécifiquement des chercheurs (historiens, généticiens, poéticiens, spécialistes de l'auteur et/ou de la période, de l'écriture pour soi et/ou de soi, etc.) –, un texte entièrement révisé (corrigé et complété), une ressource conçue en fonction des multiples enjeux (biographique, historique, génétique, générique, etc.) qui en font l'intérêt.

### 2.1.3 Les Archives d'Augustin Thierry (ArchAT)

**Responsable** : Aude Déruelle (Pouvoir Lettres Normes - POLEN)

**Résumé synthétique** : Augustin Thierry (né à Blois en 1795, mort à Paris en 1856) est une figure marquante dans la constitution de la discipline historique en France. Ses archives représente un ensemble qui comporte des cahiers mêlant des objets divers (brouillons, projets, sources, notes de lectures, correspondances, voire ordonnances médicales).

Le projet ArchAT vise à explorer la méthode de Thierry et les sources de ses œuvres, en les replaçant dans le double contexte, biographique et épistémologique, de sa cécité et du renouveau des études historiques en France à cette époque ; et, d'autre part, à mettre en valeur un patrimoine archivistique conservé à Blois, mais très peu connu.

## 2.2 L'accueil de membres internationaux

Le comité du pilotage de CAHIER s'est interrogé sur la façon d'internationaliser le réseau pour lui donner une plus grande visibilité et mieux promouvoir ses résultats. Deux rencontres internationales, facilitées par des invitations de la TGIR-HumaNum en 2018, ont accéléré cette réflexion: les journées DARIAH organisées à Paris, et la participation de CAHIER, aux côtés de la TGIR, à [DH2018] à Mexico. L'une des pistes actuellement envisagée par CAHIER est la création d'un statut de "Membre associé international" du consortium. Cette proposition sera soumise à l'Assemblée générale de novembre 2018, une communication spécifique sera ensuite préparée par les coordinateurs du consortium et diffusée à l'international.

Ce statut, qui ne permettra aucune prise en charge financière, permettra toutefois d'étendre le réseau et de croiser les pratiques de CAHIER avec celles des autres communautés. Pour l'avenir, cette ouverture de CAHIER à l'international pourrait déboucher sur un groupe de travail soumis à DARIAH ou sur la proposition d'une action COST, en lien avec un sujet précis.

## 2.3 Le soutien aux projets membres

Lors du Comité de pilotage du 25 juin 2018, il a été décidé de soutenir financièrement les 2 projets qui avaient fait une demande.

Projet	Objet de la demande
Marc-Michel Rey	Une aide pour la transcription (commencée), soutien pour l'encodage XML-TEI, valorisation du fonds en partenariat avec la médiathèque de Tonnerre (numérisation, lien avec le site Rey...),  L'édition critique en ligne du fonds pour septembre 2019 au plus tard.
Édition génétique numérique du «Robinson» de Paul Valéry	Mise en ligne des transcriptions et plugin Omeka pour liens avec ontologies

Le premier appel à soutien ayant été lancé à une période qui a nui à sa visibilité, les membres du comité de pilotage ont relayé les demandes de certains projets membres et sollicité la publication d'un nouvel appel. Celui-ci a été lancé en octobre. A la date de remise du rapport, les demandes n'ont pas encore été étudiées.

Il convient de rappeler ici les règles qui permettent aux projets de bénéficier d'un soutien financier de CAHIER :

- les subventions doivent être utilisées pour des tâches de :
  - numérisation (image ou texte)
  - d'enrichissement
  - de développement d'outils pour l'informatisation, l'édition et la recherche
- les données acquises doivent être mises à la disposition de la communauté et les développements d'outils doivent être mutualisés/mutualisables,
- les projets ne disposant d'aucun autre financement (ANR, Labex, Equipex, Région...) sont prioritaires,
- les projets ayant déjà reçu 2 subventions du consortium ne sont pas prioritaires,
- les projets subventionnés doivent afficher le soutien apporté par le consortium CAHIER et la TGIR Huma-Num sur les supports relatifs au projet (site de projet, communications, publications, codes sources),
- pour solliciter une deuxième subvention, la présentation préalable des résultats obtenus grâce à la première subvention sera nécessaire,
- un devis sera à joindre à la demande pour toute prestation externe,
- la présentation des résultats obtenus par les projets membres de CAHIER doit être réalisée lors de la réunion annuelle de l'assemblée générale (soit, au moins à l'année n+1).

## 2.4 Les activités de coordination

### 2.4.1 Les réunions du comité de pilotage

Le comité de pilotage se réunit au moins quatre fois par an, en visio, et deux fois en présentiel lors de l'atelier annuel (organisé à Montpellier en 2018) et avant l'Assemblée générale ordinaire de fin d'année (en général à Paris). Durant l'année 2018, les réunions ont eu lieu :

- le 16 mars 2018 (visioconférence)
- le 25 juin 2018 (atelier à Montpellier)
- le 10 octobre 2018 (visioconférence)
- le 22 novembre 2018 (assemblée générale ordinaire à Paris)

Le comité de pilotage est composé de membres élus par l'Assemblée Générale, il assure le bon fonctionnement du consortium, la répartition et l'attribution des fonds dont dispose le consortium et définit les orientations scientifiques du consortium (groupes de travail à mettre en place, soutiens à des projets fédérateurs, etc.).

Afin de compléter ces réunions, des fils de discussion sont organisés sur la liste de diffusion du comité de pilotage, et, le cas échéant, des décisions collectives sont assurées par



la biais de la liste, quitte à être synthétisées par une décision formelle lors du Comité de Pilotage suivant.

## 2.4.2 La réunion du conseil scientifique

Né en 2017, le conseil scientifique est composé de :

- Lou Burnard (Université d'Oxford)
- Elizabeth Burr (Université de Leipzig)
- Susan Schreibman (Université de Maynooth)
- Bertrand Jouve (CNRS)

La formation d'un conseil scientifique international au sein de CAHIER, sa forte motivation et son regard externe, constituent une forte plus-value pour le fonctionnement du consortium. Issus de domaines des humanités numériques au niveau international, il s'agit de personnalités reconnues qui apportent un regard critique et des suggestions d'évolutions, tant sur les thématiques du consortium que sur son avenir.

Le bilan de 2017 a été transmis et lu par le Conseil scientifique qui, par les voix de deux de ses membres, a émis un avis public durant l'AG du 15 décembre 2017. Leur rapport complet a été ré-examiné ensuite durant la réunion du Copil du 16 mars 2018.

La prochaine réunion du Conseil scientifique est prévue le 22 novembre 2018. Les membres du conseil scientifique reçoivent une copie du présent rapport d'activité, ils sont invités à le lire, à l'examiner durant leur réunion du 22 novembre et à faire un retour auprès de l'Assemblée Générale du 23 novembre.

## 2.4.3 L'Assemblée générale 2018

L'assemblée générale 2018 aura lieu le Vendredi 23 novembre 2018 à Paris. Elle sera précédée, la veille, d'une réunion du Comité de Pilotage et d'une réunion du Conseil scientifique.

L'ordre du jour de l'AGO est le suivant :

### **Vendredi 23 novembre 2018**

**à l'Université Sorbonne Nouvelle – Paris 3 – Centre Censier – 13, rue de Santeuil –  
75005 Paris**

#### **Matin : Salle Las Vergnas, 3ème étage.**

8h45 – 9h00                      Ouverture

9h00 – 11h00                    **Assemblée Générale**

- Rapport d'activité 2018 et projet 2019
- Discussion sur les groupes de travail
- Discussion sur les orientations scientifiques à prendre
- Discussion sur l'avenir et le rôle du consortium
- Discussion sur les activités à mettre en commun avec les autres consortiums

11h00 – 11h15                    Pause

- 11h15 – 11h45      **Intervention du Conseil Scientifique**
- 11h45 – 12h30      **Projets Cahier : états des lieux/résultats intermédiaires et présentation de projets entrés dans le consortium en 2017 – 2018**

**Après-midi : Salle 307, 3ème étage.**

- 14h00 – 16h00      **Conférences invitées**

Emilie Masson (CNRS, Service de Protection des Données du CNRS)  
*Des données personnelles dans vos CAHIERS*

Maurizio Lana (Università del Piemonte Orientale, Italy)  
*Questions méthodologiques dans l'étude des textes: le cas de l'attribution des textes anonymes gramsciens.*

- 16h00 – 16h15      **Conclusions**

## 2.5 L'atelier Annuel 2018

L'atelier de formation annuel de CAHIER s'est tenu à Montpellier, dans les locaux de l'Université Paul Valéry durant trois jours et demi, du 26 au 29 juin 2018.

L'édition de cette année a été dédiée à la rétro-numérisation de documents historiques, et plus particulièrement à la constitution de corpus lexicographiques (programme complet : <https://cahier.hypotheses.org/3640>). Dans le cadre de l'implication de la TGIR Huma-Num dans diverses activités d'infrastructure au niveau européen, la participation à l'atelier a été élargie à des personnes externes au consortium, avec le soutien de CLARIN et DARIAH.

La journée du 26 a été introduite par Laurent Romary (Inria), qui a présenté la TEI du point de vue de la description des ressources lexicales et des recommandations TEI-Lex0. L'après-midi a été consacrée à la personnalisation de la TEI avec le langage ODD (One Document Does it all) par Lou Burnard, co-fondateur de la TEI et membre du Conseil Scientifique de CAHIER.

Deux projets de lexicographie, membres de CAHIER, ont ensuite été présentés : Nénufar (sur le Petit Larousse Illustré, par Hervé Bohbot, Praxiling) et Basnage (sur les dictionnaires français du 17e siècle, par Geoffrey Williams, LiCoRN); ainsi que trois projets de l'Université Paul Valéry, en lien avec le domaine de la lexicographie: VÉgA (Vocabulaire de l'Égyptien Ancien) par Frédéric Servajean (Labex ARCHIMEDE), LexArt (terminologie des arts) par Michèle-Caroline Heck (CRISES) et le Petit Thalamus par Gilda Caiti Russo (LLACS). D'autres projets montpellierains, moins avancés que les précédents, ont également été présentés pendant l'atelier. Enfin, Francesca Frontini (Praxiling) a présenté les fonctionnalités offertes aux projets d'humanités computationnelles par le Web des données (données liées), illustrées notamment par les activités du projet européen PARTHENOS et les initiatives de l'infrastructure ELEXIS.

Plusieurs séances de formation ont été tenues autour d'outils utilisés par les projets membres de CAHIER : GROBID (GeneRation Of Bibliographic Data) par Mohamed Khemakhem (Inria), TXM par Alexey Lavrentiev (IHRIM) et divers outils liés aux traitements d'OCR dans la numérisation de documents par Hervé Bohbot (Praxiling). Le groupe de travail [Data\_Cahier] a profité de l'atelier pour assurer une permanence et

accompagner les projets dans la constitution d'un entrepôt OAI (grâce à l'outil WebOai: <http://weboai.cahier.huma-num.fr/>) en vue du signalement de leurs données sur Isidore.

L'atelier 2018 a permis des rencontres fructueuses entre personnes déjà impliquées dans les projets d'humanités numériques de CAHIER et d'autres motivées par ce domaine.

## 2.6 Les formations 2018

Le consortium participe de trois manières à la formation de ses membres. D'une part, en organisant et en prenant intégralement en charge les formations demandées par ses membres ou leurs demandes de participations à des formations existantes et ce, après validation du comité de pilotage. Et, d'autre part, en participant activement au financement de formations, ce qui revient à les labelliser.

1. Formations organisées par le consortium CAHIER :
  - a. Formation à l'outil OpenTheso et au format SKOS - 30 et 31 janvier 2018, ENS Lyon  
  
Organisée à la demande du groupe de travail "Typologie", la formation a été organisée à Lyon et donnée par Miled Rousset (Maison de l'Orient et de la Méditerranée et membre du consortium MASA). 13 personnes formées.  
Programme : <https://cahier.hypotheses.org/3421>
  - b. Rétro-numérisation de documents historiques et partage dans le Web sémantique : l'exemple de la lexicographie  
  
Cette formation a été organisée durant l'atelier de formation annuel du consortium Cahier – Montpellier – 26-29 juin 2018. 25 personnes formées.  
Programme : <https://cahier.hypotheses.org/3640>
2. Soutien aux membres de CAHIER désirant se former à certaines technologies des Humanités Numériques.
  - a. Initiation à l'encodage XML-TEI des textes patrimoniaux, Tours, 22/10/2018 au 24/10/2018 – 2 personnes prises en charge
  - b. ANF Aussois – participation au financement de la formation, conjointement avec l'Institut des Données, IDEX Grenoble, dans le cadre d'une formation de l'Agence Nationale de Formation du CNRS
3. Formations labellisées par CAHIER
  - a. École d'Été Éditions Numériques / Humanités Numériques – Grenoble – 28 mai – 2 juin 2018. <https://eedeen.sciencesconf.org>
  - b. ANF Concevoir et exploiter les sources numériques de la recherche en SHS – Action Nationale de Formation – 24 – 28 septembre 2018 – Aussois (73) - <http://anf.msh-vdl.fr>

## 2.7 Le soutien à la participation à des conférences internationales

Le consortium alloue une enveloppe de 600€ aux membres intervenant dans les conférences internationales en humanités numériques sur des problématiques liées à celles de CAHIER. En participant à ces conférences, les activités de CAHIER sont promues au niveau international. Le communicant s'engage en contrepartie à indiquer clairement la contribution du consortium dans sa communication orale et, le cas échéant, écrite. 4 soutiens ont été octroyés en 2018 :

- Conférence on Historical Lexicology and Lexicography, ICHLL - Santa Margherita Ligure, Italie- 20-22 juin 2018 - Hervé Bohbot (projet NENUFAR)
- Conférence on Historical Lexicology and Lexicography, ICHLL - Santa Margherita Ligure, Italie - 20-22 juin 2018 - Geoffrey Williams (projets Basnage et Édition électronique des œuvres de Louis de Boissy)
- TEI2018 - Tokyo - 9-13 septembre 2018 - Toshinori Uetani (projet Bibliothèques Virtuelles Humanistes (BVH))
- TEI2018 - Tokyo - 9-13 septembre 2018 - Alexey Lavrentev (projet Base de français médiéval (BFM))

## 2.8 Le soutien à des séminaires et journées d'étude

Suite à une demande faite auprès du bureau du consortium, la question de la labellisation de séminaires de recherche et de journées d'étude par le consortium CAHIER, a été posée en septembre 2018 . Dans le cas qui se présentait, le comité de pilotage s'est montré unanime. Néanmoins, il est apparu nécessaire de cadrer ce principe de labellisation d'événements scientifiques. Une discussion est en cours afin de définir clairement ces modalités.

Une journée d'étude à par conséquent été labellisée pour l'année 2018 :

- Indexer les corpus numériques, Atelier organisé par Emmanuelle Perrin (UJM–IHRIM) et Fabienne Vial-Bonacci (CNRS-IHRIM), 1er février et 29 mars 2019 IHRIM - Université Jean Monnet-Saint-Étienne

## 3 Le soutien aux groupes de travail

CAHIER accompagne les travaux de ses groupes de travail: ce sont eux qui produisent la documentation scientifique et qui transmettent les “bonnes pratiques” des Humanités Numériques. Les groupes sont constitués à l’initiative des membres du consortium, pour répondre à une question technique et/ou scientifique des humanités numériques. Le consortium compte actuellement 8 groupes actifs.

### 3.1 Groupe de Travail [Questions Juridiques]

#### Les membres du groupe :

- Elisabeth Greslou, coordinatrice (UGA, UMR Litt&Arts, Grenoble)
- Stéphanie Dord-Crouslé (CNRS, UMR IHRIM, Lyon)
- Elysabeth Hue-Gay (Université Lumière Lyon 2, UMR HiSoMA, Lyon)
- Laurence Rageot (CNRS, MSH Val de Loire, Tours)
- Cécile Beauchamps (Presses universitaires de Caen)
- Anne Garcia Fernandez (CNRS, UMR Litt&Arts, Grenoble)
- Émilie Masson (SPD – Service de protection des données, CNRS, Nancy)
- Nadine Pontal (UMR IHRIM et Centre Max Weber, Lyon)

#### Bilan 2018

Le groupe s’est encore étoffé en 2018 avec l’arrivée de Nadine Pontal (UMR IHRIM et Centre Max Weber, ENS de Lyon) et de Émilie Masson (juriste au SPD du CNRS). Les compétences d’Émilie Masson, seule juriste du groupe, constituent un apport certain.

Les réunions, souvent organisées en visioconférence, ont permis de terminer les fiches correspondant à la première phase du guide. Elles sont disponibles sur le site du consortium à l’adresse <https://cahier.hypotheses.org/corpus-auteurs-aspects-juridiques> et concernent les droits liés à l’acquisition des sources et à la constitution des corpus. Cependant, de nouveaux besoins sont remontés à travers les questions des membres du consortium et il semblerait souhaitable de rajouter des entrées « Dictionnaire » et « Anthologie ». La partie « Données personnelles » sera prochainement remplacée par une page sur le RGPD. Un travail sur des exemples de documents (conventions, contrats) est en cours et le site devrait bientôt proposer des exemples de pratiques et des jurisprudences.

En 2018, le groupe a entrepris le travail sur la seconde phase du guide: il est consacré à l’exploitation et à l’exposition des sources. Il s’agit d’aborder tous les aspects juridiques liés à un affichage web : crédits, mentions légales et obligatoires, responsabilités, contributeurs bénévoles, etc. En établissant la liste des acteurs impliqués, nous avons pu mesurer l’ampleur de la tâche compte tenu du peu de documentation et de recul sur le sujet mais également des différentes lois qui, quelquefois, s’appliquent de façon plus ou moins contradictoire.

#### Objectifs 2019

- Rédiger des recommandations pour les porteurs et acteurs de projets concernant cette seconde phase
- Rencontrer d’autres communautés travaillant sur des problématiques similaires ou liées afin d’élargir notre réflexion, notamment au sein des consortiums de

HumaNum (MASA, CORLI, ArchEthno par exemple): l'idée serait d'organiser une journée/atelier/échange consacrée aux aspects juridiques ;

- Organiser une réunion commune avec le groupe travaillant sur la correspondance pour approfondir les questions juridiques liées à ce type de corpus.

## 3.2 Groupe de Travail [EVENT]

### Les membres du groupe :

- Pierre-Yves Buard
- Marie-Luce Demonet
- Ioana Galleron
- Fatiha Idmhand
- Cécile Meynard
- Elena Pierazzo
- Julia Roger
- Geoffrey Williams

### Bilan 2018

Les travaux du groupe de travail EVENT ont débouché, durant l'année 2018, sur la rédaction d'un guide et d'une grille d'analyse des « Publications numériques de corpus d'auteurs ». En cours de finalisation, ce document sera présenté à l'occasion de l'Assemblée générale de novembre 2018. Le groupe de travail conclue ses travaux avec la publication de ce guide.

### Objectifs 2019

Le groupe pourra reprendre le guide durant l'année 2019 pour le corriger ou apporter des modifications à la marge. L'enjeu du groupe sera de communiquer et de lui assurer la plus large diffusion possible (via des traductions par exemple).

Le groupe de travail poursuivra dorénavant ses travaux au sein de deux autres groupes dont les enjeux sont liés au guide : le groupe [TEI] et le groupe [Data\_Cahier].

## 3.3 Groupe de Travail [Data\_Cahier]

### Les membres du groupe :

- Nathalie Arlin
- Fatiha Idmhand
- Maud Ingarao
- Michael Nauge

### Bilan 2018

Durant l'année 2018, le groupe de travail a finalisé le développement de l'outil WebOai, résolu les bugs techniques et organisé la mise en place d'entrepôts OAI et de sitemaps pour les projets membres de CAHIER (<http://weboai.cahier.huma-num.fr/>). Dorénavant, tous les projets de CAHIER peuvent disposer d'un entrepôt OAI pour signaler leurs données via Isidore. Le travail en étroite collaboration avec HumaNum a largement facilité la levée des

principaux verrous et accéléré les processus notamment lors des rencontres HumaNum organisées à Lyon où le groupe [Data\_Cahier] s'est rendu.

Parallèlement aux développements techniques, le groupe a organisé un consulting des projets pour leur apporter une expertise personnalisée sur leurs fichiers XML-TEI fondée sur une analyse qualitative des données. Des conseils ont ensuite été proposés pour la réalisation des teiHeaders. Un accompagnement a également été proposé, selon les cas.

Au terme de l'année 2018, les projets suivants ont été expertisés et accompagnés :

- Corpus14 : remodelage des fichiers - 1543 xml sont affichés, et après d'ultimes corrections, on pourra atteindre 1670 fichiers. Un sitemap a été créé et un teiHeader mis en ligne.
- Dossiers de Bouvard et Pécuchet & Correspondance de Bayle : actuellement en attente de la mise en ligne du sitemap et du teiHeader
- Basnage : Après quelques conseils et modifications, le teiHeader est conforme et peut bénéficier d'un sitemap
- Horace : Le teiHeader fourni a été légèrement adapté

D'autres projets de CAHIER ont été sollicités par ailleurs : D'Alembert, Condorcet, Projets eMan, IG-Louvre, Testaments de Poilus, Nenufar, Digiflor, Archives e-LV (mais les données du projet sont déjà sur Nakala, il suffit de les signaler à Isidore), Manuscrits de Stendhal et ITHAC.

L'Assemblée générale de novembre permettra de présenter une nouvelle fois le groupe et les missions qu'il s'est données.

### **Objectifs 2019**

Pour l'année 2019, le groupe va poursuivre le travail de terrain sur les données de Cahier. L'objectif est, à terme, d'offrir à tous les projets ne disposant pas d'entrepôt OAI, un accès à WebOai afin que toutes ces données soient moissonnées par Isidore.

L'autre chantier concernera le stockage des données sur Nakala. Le groupe de travail va organiser l'accompagnement des projets pour le stockage de leurs données sur la grille de services. Le travail sera réalisé en collaboration avec HumaNum (qui révise Nakala en ce moment) pour faciliter ces dépôts et en collaboration avec les autres consortiums. Une journée de travail inter-consortium sur Nakala sera proposée pour 2019.

## **3.4 Groupe de Travail [R2Cahier]**

### **Les membres du groupe :**

- Pierre-Yves Buard
- Marc Douguet
- Ioana Galleron
- Anne Goloubkoff
- Fatiha Idmhand
- Cécile Meynard
- Julia Roger
- Antoine Widlocher

## **Bilan 2018**

En 2018, et toujours sur la question du traitement informatique du « personnage » dans les récits de fiction, le groupe a poursuivi ses travaux sur l'encodage sémantique du personnage. Le développement d'une interface d'encodage XML (en XXE) a permis de poursuivre la réflexion sur les possibilités et limites du modèle arborescent de XML. Les premiers résultats de ces travaux ont donné lieu à une publication scientifique parue en 2018 dans la revue DHQ (<http://www.digitalhumanities.org/dhq/vol/12/1/000363/000363.html>) et à une présentation collective lors du colloque [DH] de Mexico (<https://dh2018.adho.org/spotting-the-character-how-to-collect-elements-of-characterisation-in-literary-texts/>).

## **Objectifs 2019**

A l'issue de cette première phase exploratoire et poussée sur le modèle XML, le groupe veut maintenant tester d'autres modèles que l'approche arborescente de la structuration XML : notamment des modèles qui reposent sur des graphes, avec des modèles ou vues adaptées au traitement spécifique des co-références. Ici, le travail sera mené sur les relations entre les personnages ou entre les idées.

## **3.5 Groupe de Travail [Typologie]**

Les membres du groupe :

- Marie-Luce Demonet
- Pierre-Yves Buard
- Sacha Diwersy
- Ioana Galleron
- Marie-Hélène Lay
- Alexei Lavrentev
- Thomas Lebarbé
- Cécile Meynard
- Anne Réach-Ngô
- Julia Roger
- François Vignale
- Geoffrey Williams

## **Bilan 2018**

La réunion du 14 décembre 2017, précédant l'Assemblée Générale, avait été l'occasion de réfléchir à une refonte complète du thésaurus « OpenTheso Cahier » en cours d'élaboration, à partir d'une classification ontologique différente, plus proche des usages du monde documentaire (Fiction/ non-Fiction), ce qui impliquerait le dédoublement du thésaurus (Forme/ Genre).

Le groupe a avancé techniquement en organisant en janvier 2018 un atelier « SKOS-Open Theso » à l'ENS de Lyon, animé par le concepteur de l'outil. Les stagiaires (13, venant de CAHIER et du labex OBVIL principalement) ont pu s'initier au langage SKOS, installer et manipuler la dernière version de l'outil, d'abord en version bac à sable, en attendant l'hébergement par Huma-Num (réalisé en mai 2018 à l'adresse <https://opentheso.humanum.fr/opentheso/>). Les exercices se sont avérés fort utiles pour exploiter les possibilités de paramétrage, déplacement, renommage, alignements, importation/ exportation, etc.



À la réunion du 29 mars les membres du groupe se sont réparti les genres ou groupes de genres pour essayer de les placer dans l'arborescence en fonction des nouvelles classifications, ce qui a été rediscuté le 5 juin, notamment pour la catégorie des dictionnaires. Deux membres du groupe ont pris rendez-vous le 11 juin avec le groupe de travail sur le référentiel « Rameau » de la Bibliothèque nationale, en cours de refonte pour tenir compte non seulement de la catégorie « sujet », mais aussi de « genre-forme ». Une présentation par la BnF lors de l'atelier CAHIER à Poitiers en 2019 est envisagée.

Les travaux du GT ont été rapidement présentés lors de l'atelier CAHIER de Montpellier en juin.

### **Objectifs 2019**

La prochaine étape (réunion prévue avant l'AG de novembre) est la stabilisation du ou des thésaurus, la limitation à cinq niveaux, afin de pouvoir attribuer aux « concepts » des identifiants stables.

## **3.6 Groupe de Travail [Crowdsourcing]**

### **Les membres du groupe :**

- Thomas Lebarbé
- Richard Walter
- Emmanuelle de Champs

### **Bilan 2018**

Le groupe de travail Crowdsourcing, après avoir établi une grille d'analyse de correspondance entre outils et besoins des chercheurs, fin 2017, a connu un temps de latence en 2018.

### **Objectifs 2019**

L'objectif de 2019 est de repartir de la grille définie fin 2017 pour rédiger collectivement un guide de bonnes pratiques à l'attention de la communauté CAHIER.

## **3.7 Mutualisation inter-consortium des activités des groupes de travail**

A la demande de certains groupes de travail, et plus particulièrement des groupes de travail [Questions juridiques] et [Data\_Cahier], les coordinateurs du consortium ont été missionnés pour prendre contact avec les coordinateurs des autres consortiums actifs (MASA, CORLI, Archives des Ethnologues, COSME, en particulier) afin d'identifier avec eux les points d'achoppement et les actions à mener collectivement, pour chaque groupe de travail.

Côté [Questions juridiques], le consortium MASA a d'ores et déjà mis à disposition un outil de gestion et publication des guides de bonnes pratiques sur lequel CAHIER a déposé ses données ; côté [Data\_Cahier], des membres ont été invités à une première réunion de travail sur Nakala organisée fin novembre par des membres du Consortium Archives des ethnologues.

## 4 Présentations des activités lors de réunions internationales

Sur invitation de la TGIR Huma-Num, le consortium a pu s'investir dans deux événements majeurs de la vie des humanités numériques à l'échelle internationale. Au niveau européen, lors du DARIAH Annual Event, organisé à Paris du 22 au 24 mai 2018, au niveau international, lors de la conférence DH'18, organisée à Mexico du 26 au 29 juin 2018. Ces deux événements ont permis de promouvoir le consortium et de donner une visibilité internationale au réseau et aux activités du consortium. Nous avons pu constater l'intérêt croissant des chercheurs européens pour CAHIER. Suite à ces interactions notamment, nous avons entamé les réflexions sur un statut de membre associé international (voir supra).

### 4.1 DARIAH Annual Event 2018, Paris

Plusieurs membres du consortium, dont les deux coordinateurs, ont pu se rendre à l'événement annuel de DARIAH afin de participer aux groupes de travail et assister aux présentations. Outre une meilleure compréhension de l'infrastructure de recherche européenne, les moments de rencontre informelle nous ont permis de tisser des liens avec une communauté internationale qui partage nos préoccupations et nos objectifs.

Par ailleurs, la TGIR nous a permis de présenter le consortium lors d'une session de poster, donnant ainsi de la visibilité à une structure et une organisation parfois difficile à comprendre depuis l'étranger.

### 4.2 [DH2018] à Mexico

Pour la conférence internationale DH'2018, HumaNum a proposé une table ronde destinée à présenter les consortiums et leurs activités. Pour CAHIER, cette rencontre a été l'occasion de confronter nos orientations avec celles d'équipes étrangères, notamment autour de la création, du développement ou de la maintenance d'outils, des actions réalisées en matière d'ouverture des données, de publication de bases de données, corpus en ligne, etc.), de l'élaboration et de la diffusion des bonnes pratiques, des collaborations avec d'autres consortiums, des collaborations interdisciplinaires, des collaborations européennes et/ou internationales (DARIAH, CLARIN, actions COST, colloques et/ou projets internationaux, etc.).

De ce point de vue, le consortium considère que le travail, entre les consortiums et HumaNum pour DH2018 a été une réussite et se réjouit de la poursuite de la réflexion sur les formes d'une nouvelle participation de HumaNum et de ses consortiums à DH2019.

CAHIER apprécie l'implication de la TGIR Huma-Num pour faire connaître, partager son expérience et confronter ses pratiques et objectifs avec les activités de la communauté internationale.

## 5 Description de l'apport aux communautés scientifiques

Le consortium CAHIER transcende les disciplines et les frontières établies par les sections disciplinaires (CNRS, CNU, etc.). CAHIER s'est construit en fédération de chercheurs et d'ingénieurs autour de problématiques scientifiques au-delà des silos habituels de manière à répondre aux problématiques communes des équipes de recherche impliquées dans les sciences des textes. Le consortium a construit une communauté scientifique des Humanités Numériques des textes qui compte aujourd'hui 300 membres qui partagent et confrontent leurs approches techniques et méthodologiques des textes en lien avec l'informatisation de ceux-ci. Ils considèrent que de la rencontre entre les humanités et l'informatique émergent des interrogations communes qu'ils peuvent partager au sein du consortium et auxquelles ils peuvent tenter de répondre grâce au partage des approches et aux rencontres, aux formations et groupes de travail.

Dès sa mise en place, le consortium a souhaité organiser des formations aux outils et normes en vigueur : les formations garantissent le socle des "communautés de pratiques", fondement des approches informatiques des humanités et favorisent l'intérêt pour l'échange des données. La prise de conscience de la communauté quant à l'intérêt du respect des normes pour faire évoluer les formats de l'édition critique, génétique, encyclopédique (cf. Accord de labellisation 2011) passe par cette information et par des formations qui assurent le terrain commun (au sens psycholinguistique du terme): les "bonnes pratiques" en matière d'encodage des textes (XML, TEI), de manipulation à des fins éditoriales (XSLT), d'exploitation scientifique (TXM) ou de préparation de données brutes. Contribuer à préparer l'avenir de la recherche en SHS est au cœur des réflexions de CAHIER.

Les groupes de travail, les formations et l'atelier annuel lui donnent l'occasion de diffuser et de valoriser les outils développés ou utilisés par ses membres et les corpus qu'ils éditent. Leur diversité, la variété des enrichissements disponibles et la multiplicité des outils apportent aux membres des regards différents et des perspectives originales qui leur permettent d'enrichir leurs travaux, leurs propres outils et corpus. CAHIER ne promet aucune solution éditoriale mais en propose une palette, certains se recouvrant partiellement, permettant aux chercheurs de s'en saisir en fonction de leurs problématiques spécifiques (une rubrique dédiée a été créée sur notre site). Les groupes de travail qui ont été constitués au sein du consortium, analysent, eux, les approches de façon à produire des documentations ou des recommandations (comme au sein du groupe « questions juridiques »). Chaque groupe de travail s'est donné comme objectif la publication et la diffusion large de prescriptions et de conseils dans le but d'éviter les écueils traditionnels des projets d'éditions de sources textuelles, et de diffuser ses réalisations en informant la communauté de ses avancées.

Les formations, ateliers et groupes de travail privilégient, bien entendu, les membres du consortiums mais sont, par essence, ouverts à la communauté scientifique la plus large, afin, d'une part, de contribuer à la controverse et au débat scientifiques, et d'autre part, de promouvoir des pratiques bien au delà des équipes projets représentées dans le consortium.

## 5.1 Plus-value apportée par l'organisation en consortium

La plus-value apportée par le consortium (par rapport à d'autres formes et à d'autres activités des partenaires) réside dans son organisation en réseau de projets, ce qui facilite à la fois la rencontre transdisciplinaire entre les métiers de la recherche mais également entre secteurs et hors des cloisonnements disciplinaires et institutionnels.

CAHIER tient à diffuser les travaux des groupes de travail et du consortium, pour cela, il s'appuie sur les outils de communication habituels afin d'assurer la visibilité de son activité grâce à une liste de diffusion, [cahier@groupe.renater.fr](mailto:cahier@groupe.renater.fr), à son site web <http://cahier.hypotheses.org/>, son compte twitter @ccahier et sa page Facebook [consortium.cahier](https://www.facebook.com/consortium.cahier). CAHIER a par ailleurs construit une collection sur HAL <https://hal.archives-ouvertes.fr/CONSORTIUM-CAHIER> de façon à y diffuser les supports de communication, de conférence et de formation, mais également les articles scientifiques produits par ses membres et les guidelines susceptibles d'être conçus. Plus récemment, CAHIER s'est investi dans le projet d'un autre consortium afin de contribuer à la plateforme d'exposition des guides de bonnes pratiques de MASA.

Après plus de six années d'existence, CAHIER organise maintenant le signalement massif et l'exposition des données des membres. Pour cela, il s'est doté de l'outil WebOai (<http://weboai.cahier.huma-num.fr/>) pour permettre à tous les projets de bénéficier d'un entrepôt OAI et de présenter ses données sur Isidore. En lien avec l'UMR IHRIM, le consortium a missionné Nathalie Arlin afin qu'elle accompagne les équipes du consortium dans cette démarche. Les données de CAHIER déjà moissonnées par Isidore (plus de 14.000 entrées à ce jour) sont consultables ici :

[https://www.rechercheisidore.fr/search/?source\\_tree=10670/3.06ri7u](https://www.rechercheisidore.fr/search/?source_tree=10670/3.06ri7u) .

Enfin, CAHIER dispose également de plaquettes de promotion. La transmission scientifique dans les canaux « traditionnels » est assurée, elle, par les membres des projets (communications dans le cadre de journées d'études, colloques, publications scientifiques, etc.).

## 5.2 Etat d'avancement

(par rapport aux objectifs fixés pour l'année et par rapport à la programmation initiale)

Le Consortium CAHIER rassemble maintenant une communauté conséquente d'experts (« praticiens ») des Humanités Numériques des textes partageant des problématiques, des réflexions, des obstacles et des enjeux communs et couvrant à la fois le territoire national, les différentes disciplines, les différentes langues, les différentes périodes.

### 5.2.1 Objectif réseau :

Le réseau de CAHIER ne cesse de s'étendre, il a connu une croissance moindre ces deux dernières années par rapport aux années précédentes et ceci peut être imputé aux nouvelles modalités d'adhésion. Toutefois, nous pouvons dire que le besoin de soutien à la recherche était nécessaire et que CAHIER était/est une réponse bienvenue : le consortium joue le rôle

d'appui lorsque les projets ont besoin d'exemples, d'études de cas pour aller vers d'autres sources de financement.

### 5.2.2 Objectif formation :

L'un des objectifs du consortium était d'apporter de nouvelles compétences aux chercheurs et ingénieurs membres. Cette année, peu de formations ont été demandées mais nous avons constaté une évolution dans la nature des demandes. En effet, la communauté commence à être familiarisée avec les notions d'outils et de normes tandis que certaines institutions ont durablement installé leurs formations dans le paysage national (TEI, TXM, etc.). Nous pensons qu'il faut dorénavant passer au cap supérieur et mieux organiser la préservation des données produites (à l'aide de Nakala par exemple) et, surtout, leur exploitation de façon à susciter de nouvelles recherches scientifiques sur les objets/données. Des formations à OpenTheso, XSLT, BaseX, aux langages de requêtes tels que XQuery (ou autres) et même à la gestion de bases de données en lien avec le traitement d'une grande quantité de données nous semblent fondamentales. Il s'agit maintenant pour nous de préparer le terrain de la « curation » des données dans la recherche scientifique sur les corpus d'auteurs : il est l'enjeu des prochaines années. Les expérimentations du groupe de travail R2CAHIER, par exemple, emprunte déjà ce chemin.

### 5.2.3 Objectif réflexion et bonnes pratiques (GDT) :

Les groupes de travail ont été particulièrement actifs : les trois nouveaux groupes initiés en 2017 ont été très productifs et un nouveau groupe se crée en fin d'année, revenant sur les fondements de la constitution du consortium, le groupe TEI. C'est une dynamique efficace, à notre sens, signe de vitalité du réseau. Les activités de ces groupes de travail contribuent à poser les fondements des nouvelles épistémologies des Humanités Numériques. Le consortium est ouvert à toute nouvelle suggestion de création de groupes de travail et nous continuerons à encourager ces initiatives. Actuellement, les demandes de groupe de travail remontent des membres et sont étudiées et validées par le CoPil.

### 5.2.4 Objectif soutien à la recherche :

Le soutien financier aux projets a fait l'objet d'un questionnement au sein de l'assemblée générale et du comité de pilotage du consortium. Nous avons collectivement souhaité maintenir cette pratique à l'heure où les appels à projets attendent des démonstrations, voire des « preuves de concept ». Cette année, 3 projets ont été soutenus, l'appel à soutien n'ayant pas été fait dans une temporalité adaptée – un nouvel appel a été lancé pour la fin d'année. L'aide de CAHIER, certes modeste (3000€), leur permet de mettre le pied à l'étrier et de démarrer.

## 5.3 Difficultés rencontrées

CAHIER est un consortium très actif. Ses membres sont particulièrement impliqués dans son fonctionnement et dans la dynamique des groupes de travail qui tentent de répondre à des problématiques communes. Cette dynamique participative, particulièrement productive et source de débats scientifiques, est néanmoins marquée par une inquiétude grandissante quant à l'avenir du consortium. En effet, CAHIER, en tant que consortium de la TGIR HumNum, a une espérance de vie limitée, alors qu'en tant que réseau de chercheurs, il a besoin de se projeter dans la continuité et la durabilité.

Le modèle imposé par l'Institut des Sciences Humaines et Sociales pour les consortiums est limité à dix ans dans le meilleurs des cas : une première labellisation de quatre ans pour se constituer, une seconde de la même durée pour se réaliser, une troisième de deux ans pour clore ses activités.

CAHIER en est à la troisième année de sa deuxième labellisation et se préoccupe depuis plus d'un an de son avenir au delà du modèle 4-4-2. Dans le meilleur des cas, sous réserve bien entendu d'une troisième labellisation, le consortium n'existera plus sous sa forme actuelle dans trois ans.

En attendant de nouvelles discussions avec HumaNum, plusieurs pistes ont été examinées à ce jour :

- un Groupement de Recherche (GDR<sup>1</sup>), c'est à dire un regroupement d'unités de recherche du CNRS . Les GDR sont créés pour une durée de deux à quatre ans renouvelables. Ils disposent cependant de fonds beaucoup plus limités que ceux dont dispose le consortium actuellement (20 à 30% du budget actuel).
- une Structure Fédérative de Recherche, c'est à dire un consortium d'unités de recherche. En écartant la difficulté non négligeable du financement – chaque unité apportant sa contribution financière – le problème réside dans la dimension administrative et juridique du consortium. Chaque tutelle (universités en l'occurrence) de chaque unité de recherche de chaque projet membre du consortium serait signataire de la convention de consortium, ce qui représente un frein dans le cadre de CAHIER où l'on peut compter plusieurs dizaines d'unités de recherche concernées.
- Une Société Savante, sous la forme d'une association à but non lucratif, loi de 1901. Les sociétés savantes sont pléthores dans le domaine de la recherche, mais ne portent pas les mêmes objectifs ni les mêmes modes de financement. Le plus souvent, elles portent également une revue et une ou plusieurs conférences d'envergure nationale voire internationale. Il semble difficile de placer CAHIER en « concurrence » avec les sociétés savantes existantes, notamment celles dont le consortium couvre une part importante du domaine scientifique comme c'est le cas pour Humanistica (dont CAHIER est membre institutionnel).

Enfin, et indépendamment de ces questionnements, comme le consortium ne dispose pas d'une personnalité juridique propre ou d'une délégation de signature d'une tutelle, contrairement aux unités de recherche (UMR ou EA) ou à la TGIR Huma-Num (UMS 3598), l'absence de personnalité juridique ne lui permet pas de déposer des demandes de financement de projets (ANR, ERC). Une telle existence pourrait lui permettre, à terme, d'assurer sa continuité et donc la mission qu'il s'est donné, au delà des labellisations.

---

<sup>1</sup> Le Groupement de recherche est un regroupement d'unités, en totalité ou en partie, et/ou de FRE et/ou d'ERL, relevant ou non du seul CNRS, autour d'un objectif scientifique

## 6 Informations sur le budget 2018

Afin de mettre en œuvre l'ensemble des actions, nous avons proposé un budget à hauteur de 90.000€ (voir colonne « prévision pour 2018 » du tableau ci-après). Une dotation de 80.000€ nous a été accordée et a été dépensée comme suit :

- Le fonctionnement du consortium reste le même – le travail effectué par Mme Rageot est particulièrement précieux, à la fois comme interlocutrice des projets que comme gestionnaire. Avec l'accord de la MSH-VdL, nous souhaitons continuer ce mode de fonctionnement.
- La gouvernance a certes un coût moindre puisque les réunions du CoPil sont pour moitié effectuées en visio-conférence. Toutefois, sont pris en compte la création du Conseil Scientifique, dont les membres seront pris en charge pour leur venue à l'AG et la prise en charge d'un membre de chaque projet pour l'assemblée générale.
- Nous avons revu à la baisse le coût des formations (certaines d'entre elles étant parfois indissociables des travaux des groupes de travail – voir ci-après). L'enveloppe prévue a été totalement attribuée.
- Les travaux des groupes de travail étaient prévus à hauteur de 20.000€. Afin de rentrer dans le budget alloué et grâce aux économies faites par les participants aux groupes, seuls 6.300€ ont été dépensés en 2018, nous permettant de trouver l'équilibre.
- L'adhésion aux communautés scientifiques et associations savantes nous semble tout aussi essentiel. Il est du ressort de CAHIER de prendre place au sein de la TEI et d'y jouer son rôle représentatif d'une communauté de pratiques française, tout comme au sein d'Humanistica, où le consortium participe activement à la création de la revue de l'association.
- Enfin, comme discuté préalablement, le soutien financier aux projets reste essentiel dans la construction de la communauté. Notre demande de 25.000€ pour cette ligne budgétaire en 2018 a été légèrement dépassée.

DEPENSES	Exécuté 2018		Exécuté 2018
	SE	RP	SE+RP
<b>Dépenses de fonctionnement</b>	<b>55 852,07</b>	<b>23 110,61</b>	<b>78 962,68</b>
<i>Fonctionnement du consortium</i>	10 000,00		10 000,00
<i>Gouvernance du consortium : réunions du comité de pilotage et AG</i>	3 537,07	8 000,00	11 537,07
<i>Travaux des groupes de travail</i>	6 352,15	30,50	6 382,65
<i>Organisations de formations ou soutien à la participation à des formations</i>	14 736,60	4 995,11	19 731,71
<i>Soutien à la participation à des conférences internationales</i>	1 200,00	1 200,00	2 400,00
<i>Cotisations Associations (TEI et Humanistica)</i>	0,00	2 200,00	2 200,00
<i>Soutien à l'organisation d'évènements scientifiques</i>		500,00	500,00
<i>Subventions aux projets</i>	19 964,80	6 185,00	26 149,80
<i>Communication</i>	61,45		61,45



## 7 Programme 2019

Nous poursuivons la promotion des activités de CAHIER et de la TGIR HumaNum sur le territoire.

### 7.1 Objectif formation :

- Atelier annuel de formation
- Formation à Web-OAI in situ. Formation à Nakala.
- Formations SKOS et OpenTheso
- Soutien aux formations « habituelles » - ANF Aussois, TEI-Tours.
- Budget prévisionnel de formations ad-hoc en fonction de l'avancement des groupes de travail.
- Soutien à la participation aux conférences

### 7.2 Objectif soutien aux projets :

Le nombre de demandes d'adhésion et de soutien financier confirme la nécessité d'accompagner des projets « embryonnaires » à se construire dans la perspective des Humanités Numériques afin qu'ils puissent ensuite demander des financements infra-nationaux (région notamment), nationaux (ANR par exemple) et supra- nationaux.

En contrepartie, des règles plus exigeantes ont été mises en place, aussi bien pour l'attribution des soutiens financiers (<3000€) que pour la visibilité des résultats de ces soutiens (présentations publiques aux AG, exposition des données).

### 7.3 Objectif bonnes pratiques :

Le nombre important de membres a fait émerger des groupes de travail autour de besoins et de problématiques partagées. La contrepartie est le coût de ces groupes. Néanmoins, notre première prévision s'étant avérée surestimée, nous la corrigeons : d'un budget initial de 20.000€ nous révisons à 15.000€ pour le fonctionnement des groupes de travail et l'année prochaine s'avérera coûteuse encore : prise en charge et indemnisation d'une ingénieure du consortium, journées d'études, publication numérique et papier de livres blancs, formations.

### 7.4 Objectif réseau :

Notre objectif est de poursuivre la construction d'un réseau cohérent. A cet effet, nous avons choisi de réorganiser le consortium de façon à mieux définir la notion de « projet associé ». Les « projets associés » (i.e., ne pouvant disposer des moyens financiers du consortium) sont ceux qui ne peuvent libérer leurs données (ce qui était le cas depuis la création du consortium), mais également des projets terminés et ne jouant pas, a minima, leur rôle au sein du consortium (défaut de réponse aux demandes d'informations, notamment).

La force du consortium est à la fois son réseau et les liens qu'il permet de construire entre des projets qui n'avaient pas originellement de raison de se retrouver en un lieu commun. Afin d'aider à l'évolution et l'agrandissement de ce réseau, nous avons facilité l'adhésion au fil de l'eau.

Par ailleurs nous souhaitons développer nos objets de communication (posters, diffusion, plaquettes, traductions des rapports et guides), d'où la création d'une ligne budgétaire à cet effet.

## 7.5 Objectif avenir du consortium :

Le consortium a obtenu une labellisation de 2011 à 2015 qui a été renouvelée pour 2015-2019. Même si la TGIR a manifesté le souhait de nous accompagner dans la reconversion de CAHIER, les membres du consortium restent néanmoins préoccupés par ce manque de visibilité à long terme. Il est en effet difficile de planifier l'avenir quand celui-ci est incertain. Des réunions ont d'ores et déjà eu lieu afin de penser « les formes » d'un avenir du consortium (GIS, société savante, etc.) ou plutôt, les façons de pérenniser la communauté scientifique qu'il est en train de construire. Nous comptons évidemment sur le soutien et les conseils de la TGIR Huma-Num, et de l'InSHS plus généralement, afin de co-construire la nouvelle forme structurelle de CAHIER.

## 7.6 Budget Prévisionnel 2019

Le budget prévisionnel 2019 est présenté ci-après, en comparaison du budget prévisionnel 2018 tel que nous l'avons soumis l'an dernier et du budget exécuté cette année.

- **Fonctionnement** : le bon fonctionnement du consortium est assuré par le travail de gestion et d'accompagnement quotidien de Laurence Rageot. Nous remercions la MSH Val de Loire de la décharger d'une partie de son travail.
- **Gouvernance** : les réunions du Comité de Pilotage et l'Assemblée Générale représentent un coût non négligeable du fonctionnement du consortium mais sont un rouage important. L'Assemblée Générale est aussi l'occasion de faire intervenir notre conseil scientifique ainsi que des conférenciers dont nous prenons en charge les frais.
- **Groupes de travail** : comme stipulé précédemment, le consortium souhaite continuer dans la même démarche que l'année précédente. L'une des activités importantes du groupe de travail [Data\_Cahier] est d'accompagner les projets de recherches dans leur mise aux normes et exposition par le WebOAI et nous souhaitons assurer ce service rendu autant aux projets de recherches individuellement qu'à la communauté des humanités numériques. Une grande partie du budget, revu en hausse pour l'an prochain, sera affectée à cet accompagnement.
- **Formations** : les formations assurées ou soutenues par le consortium pallient un manque cruel de formation dans la communauté scientifique, là où l'autoformation n'est pas suffisante, le partage d'expérience est essentiel. Nous souhaitons poursuivre cette entreprise qui donne une visibilité non négligeable au consortium puisque nous associons systématiquement à notre soutien un affichage voire une présentation des activités de CAHIER.
- **Cotisations** : le consortium est adhérent du Consortium TEI et d'Humanisti.ca, deux communautés importantes, au niveau international pour la première, au niveau de la francophonie pour la seconde.
- **Evènements scientifiques** : Il s'agit pour nous d'une nouveauté. Cette forme de labellisation, de reconnaissance de qualité technique et scientifique que le consortium apporte aux évènements scientifiques (colloques, journées d'études, ateliers) est conditionnée par un certain nombre de garanties (portage par un

membre du consortium, adhésion aux valeurs du consortium, adéquation à nos thématiques de travail, ouverture au plus large public).

- **Subventions projets** : ce budget a fait ses preuves pour permettre à des projets de mettre le pied à l'étrier, ou à l'opposé, pour finaliser les données. Le consortium s'accorde à vouloir prolonger ce mode de fonctionnement, sans en augmenter l'enveloppe malgré un léger dépassement cette année.
- **Communication** : nous souhaitons en 2019, dernière année de notre deuxième labellisation, nous appuyer sur des objets de communication et de diffusion afin de laisser une trace tangible des activités du consortium et de son réseau : fascicules, prospectus, traductions des guides, etc.).

DEPENSES	Exécuté 2018	Prévision pour 2018	Prévisions pour 2019
	SE+RP	SE + RP	SE+RP
Dépenses de fonctionnement	78 962,68	90 000,00	100 000,00
<i>Fonctionnement du consortium</i>	10 000,00	10 000,00	10 000,00
<i>Gouvernance du consortium : réunions du comité de pilotage et AG</i>	11 537,07	12 000,00	15 000,00
<i>Travaux des groupes de travail</i>	6 382,65	20 000,00	15 000,00
<i>Organisations de formations ou soutien à la participation à des formations</i>	19 731,71	20 000,00	25 000,00
<i>Soutien à la participation à des conférences internationales</i>	2 400,00		2 500,00
<i>Cotisations Associations (TEI et Humanistica)</i>	2 200,00	3 000,00	2 500,00
<i>Soutien à l'organisation d'événements scientifiques</i>	500,00		2 500,00
<i>Subventions aux projets</i>	26 149,80	25 000,00	25 000,00
<i>Communication</i>	61,45		2 500,00