

Annexe 1 : Bilan scientifique (format libre, 2 à 3 pages maximum)

L'école a été construite comme une alternance de moments de réflexions générales sur la question de l'exploration en SHS et de temps plus techniques ou pratiques de partage d'expériences et de formations appliquées à des méthodes ou des outils. La communauté scientifique éclectique réunie à cette occasion était invitée à réfléchir à la question générale de l'exploration, à ses propres pratiques et fortement incitée à en découvrir d'autres.

De fait, notre semaine a donné lieu à des échanges scientifiques tout à la fois riches et féconds. Ceux-ci ont associé intervenant.es et participant.es non seulement *via* les présentations et les réactions et questions suscitées, mais aussi par des discussions collectives programmées ou informelles stimulantes. La confrontation des expériences concrètes, des approches et des disciplines a confirmé le constat à l'origine de notre projet : l'exploration ne se définit pas comme une étape clairement identifiée dans un processus de recherche, elle prend au contraire de multiples formes selon les objets et les approches. Elle intervient à plusieurs moments de la recherche, suscite des allers-retours, mobilise des savoir-faire, ce qui explique sans doute qu'elle reste une pratique encore floue et mal documentée en SHS.

Un premier sentiment général a convenu de la nécessité de formaliser cette étape clef de la recherche, et nous avons pu observer et discuter quelques pistes en la matière (rédaction et partage de carnets d'enquête personnels ou collectifs, production de "projets informatiques" associant données et programmes documentés, utilisation d'outils de collaboration ou de versionnement adaptés, utilisation des plans de gestion de données...). De même, il semble acquis que l'exploration mériterait une meilleure valorisation dans le cadre des publications scientifiques - les *data paper* sont par exemple une voie possible - mais cette semaine a permis de mesurer la difficulté qu'il y a encore de parler et formaliser l'exploration "essai-erreur", "l'invisible" - i.e le travail de conception, le travail de terrain, le travail qui conduit à définir la question de recherche...

Il nous semble que de nombreux autres aspects abordés au cours des échanges de la semaine ont résonné parmi les participant.es. Nous en présentons ici quelques-uns, que nous appréhendons dans ce bilan scientifique comme autant de portes désormais ouvertes à une réflexion collective.

A/ Réflexions épistémologiques sur l'exploration : liens entre les sources/l'enquête, les données et la connaissance, le rapport à la preuve.

Les séances plénières ont permis de nous interroger sur la validité épistémologique de la notion d'exploration. La question des bénéfices et des risques a été abordée, en axant sur les façons de limiter ces derniers et de répondre à quelques faux procès. A travers une série d'exemples concrets, l'école a permis d'illustrer les principales méthodes explora-

toires – en particulier visuelles et graphiques mais également statistiques et textuelles – qui se sont développées dès le milieu du 20^e s. en les comparant avec le paradigme dominant communément appelé « hypothético-déductif » (cf plénière de J.D. Fekete en particulier).

Ainsi, plusieurs interventions ont convergé quant au constat du caractère bicéphale de l'exploration en SHS : processus qui permet d'évaluer la pertinence éventuelle d'une (ou plusieurs) hypothèse(s) fondée(s) sur des *a priori* généraux ou des intuitions, il est capable aussi de faire émerger de nouvelles interrogations et hypothèses. A cet égard, à l'heure des "Big Data" et de la démocratisation de l'accès à des méthodes d'analyse complexes, il devient fondamental d'interroger les résultats d'une exploration réalisée à l'aide d'outils spécifiques à l'aune des concepts mobilisés par ces outils et des modalités de leur mise en œuvre.

B/ Réflexions autour des démarches collaboratives, de l'interdisciplinarité et des rapports entre les SHS et l'informatique (quid de la décomposition/complémentarité des tâches ?)

Le travail collectif est désormais fréquent, notamment lorsqu'il est nécessaire de recourir à des environnements, des outils informatiques ou des méthodologies complexes. Du point de vue de l'exploration, cela pose la question générale de la transparence de la chaîne qui conduit des données à la connaissance que l'on en extrait. De quelle manière "thématiciens" et "informaticiens", "chercheurs" et "ingénieurs" vont-ils collaborer, se partager le travail, et s'assurer d'une compréhension réciproque, tant des questions de recherches que des "objets" manipulés et des contraintes éventuelles (importance de la "traduction" et de la définition d'un langage commun) ?

A travers le partage de plusieurs expériences pluridisciplinaires entre informaticien.nes, chercheur.es en « visual network analytics » et plus largement chercheur.es et ingénieur.es en sciences humaines et sociales, les stagiaires ont pu bénéficier et discuter de quelques "leçons" tirées au fil de collaborations concrètes. Nous retenons ici deux grandes affirmations : *Les algorithmes ont peu de valeur sans données de qualité* et *Les données à elles seules ne permettent souvent pas de poser de bonnes questions*. Ces acceptions renvoient toutes deux au lien indissociable entre la problématique d'une recherche, le recueil et l'analyse des données, et, dans le cas d'un travail collectif et/ou pluridisciplinaire, à l'importance d'un dialogue serein, ouvert et non dissymétrique entre des profils et des disciplines aux langages différents. Ainsi, le débat très riche qui a agité notre assemblée durant la semaine a surtout été orienté sur les manières d'améliorer le dialogue entre SHS et informatique (au sens large). Chaque partie ayant – chacun en a convenu – à faire évoluer sa position :

- Les SHS ont tout à gagner à s'écarter d'une posture « consumériste » de l'informatique où les informaticiens seraient uniquement associés comme pourvoyeurs d'outils plutôt que comme partenaires. Cela est d'autant plus vrai au moment d'une démarche exploratoire, afin de re-

lativiser l'effet "magique" de certains outils et se donner le moyen de les choisir et de les utiliser à bon escient.

- Du côté des sciences informatiques (plus généralement les sciences formalisatrices), une appropriation des thèmes, concepts et méthodes des SHS peut contribuer à réduire l'utilisation des SHS comme prêtonom d'une pluridisciplinarité de surface, pourvoyeuse de "data", souvent interprétées en dehors de leur contexte de production et de la littérature SHS pré-existante. L'appropriation mutuelle et la discussion des problématiques des un.es et des autres enrichit grandement une collaboration.

Dans cette perspective, la réflexion initiée sur la pratique du "codage" (au sens des SHS ou au sens de l'informatique) a montré qu'elle constitue une dimension importante des démarches exploratoires et collectives.

C/ Réflexions méthodologiques autour de la massification des données (web) et de la nécessité d'améliorer nos capacités d'exploration de ces dernières en intégrant les démarches en usage dans les SHS.

Le Web est le support d'une masse de données, plateformes, documents et matériaux de toutes sortes (textes, vidéo, images,...), souvent structurés et sémantisés. Associé aux méthodes et services utilisant l'Intelligence Artificielle, tel l'apprentissage profond (*deep learning*), il offre un potentiel d'exploration comparative et cumulative qui est incarné et mis en œuvre par ce qu'il est désormais convenu d'appeler la Science des données.

Nous avons consacré une journée spécifique de l'ET à la problématique de l'exploration du web, en nous focalisant notamment sur la présentation des concepts, outils et procédures permettant de traiter les sources sémantiques et textuelles. Afin d'éviter une utilisation « mécanique » des outils d'exploration présentés dans les ateliers, il nous a semblé qu'il fallait y associer une réflexion distanciée sur les aspects méthodologiques et épistémologiques liés à leur exploration. L'intervention de S. Parasie a permis de poser les termes de l'exploration des sources textuelles du web et d'en éclairer les enjeux.

L'essor du web et des réseaux sociaux offre aux sciences sociales un volume considérable de matériaux textuels qui sont le support d'un registre étendu d'activités sociales. Cet élargissement des sources de l'enquête en sciences sociales se heurte toutefois à plusieurs obstacles : (1) ces matériaux textuels sont produits par des plateformes numériques qui sollicitent, encadrent et mettent en forme les expressions et les échanges, mais contrôlent aussi la façon dont les chercheur.es y ont accès. (2) l'essor de nouvelles techniques d'analyse textuelle issues des mondes informatiques ajoute une opacité des algorithmes et leur difficile appropriation par les chercheur.es. (3) l'ancrage social des personnes qui s'expriment sur le web et les réseaux sociaux demeurent en grande partie inconnu : font ainsi souvent défaut des informations aussi cruciales pour la sociologie que le niveau de revenus ou de diplôme, la catégorie sociale,



Écoles thématiques 2020 – BILAN

et même l'âge ou le genre de ceux et celles qui prennent la parole en ligne.

Un ensemble de recherches contemporaines relevant des sciences sociales, qui explorent des méthodes originales pour surmonter ces obstacles ont été présentées et discutées afin d'ouvrir des pistes et redonner une épaisseur sociale aux traces textuelles, et ainsi à mettre ces techniques au service d'un questionnement de sciences sociales.

Plus généralement, il est apparu que les approches issues des sciences sociales ont sans doute aussi beaucoup à apporter à cette "science des données" fondée sur le nombre et l'accessibilité numérique. Les approches historique ou archéologique d'interrogation et de mise en perspective des sources de données et de construction d'objets de recherche à partir de traces documentaires ou matérielles, ou bien encore les approches qualitatives ou ethnographiques peuvent permettre non seulement de caractériser et contextualiser les mondes sociaux virtuels, les dispositifs et les données numériques mais aussi de les aborder dans une démarche plus générale de compréhension du monde social.

- La spécificité des démarches qualitatives

Ayant délibérément fait le choix dans l'ET d'aborder l'exploration par l'exposition d'une panoplie d'outils permettant de manipuler des "données", nous avons pu mesurer la difficulté persistante durant la semaine à aborder la question de l'exploration du point de vue d'une démarche "qualitative", en général pas ou peu outillée et souvent intéressée à l'exploration d'un terrain de recherche à l'aide de sources très hétérogènes ou de protocoles d'enquête spécifiques (entretiens libres, observation participante,...). Quelques interventions ont rappelé la richesse que l'on pouvait tirer de telles approches.

Il nous a semblé que la présentation d'outils et des formes d'interrogations et de résultats qu'ils proposaient étaient à même de donner des "idées" d'exploration aux personnes qui ne les pratiquaient pas, ce qui a semble-t-il été apprécié. Par ailleurs, certains outils plus particulièrement utilisés dans les approches qualitatives ont aussi été évoqués, tels les CA-QDAS (*Computer-assisted qualitative data analysis software*). Une séance spécifique sur des retours d'expériences de ces pratiques de codage qualitatif en présence d'informaticien.ne.s et de "travailleurs de la donnée", aurait sans doute pu permettre, malgré la grande différence de sources, d'expériences et de vocabulaires, d'établir des analogies dans les façons de faire ainsi que de réfléchir aux processus cognitifs généralement à l'œuvre dans les démarches exploratoires (classer, compter, rapprocher,...) et à la valeur heuristique des "données" à l'aune de la diversité des situations et représentations sociales.