Zeitschrift: Geschichte und Informatik = Histoire et informatique
Herausgeber: Verein Geschichte und Informatik
Band: 9 (1998)

Artikel: Le SIDOS, une archive de données pour les sciences sociales
Autor: Hadorn, Reto
DOI: https://doi.org/10.5169/seals-7231

Nutzungsbedingungen
Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften auf E-Periodica. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. Das Veröffentlichen von Bildern in Print- und Online-Publikationen sowie auf Social Media-Kanälen oder Webseiten ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Mehr erfahren

Conditions d'utilisation
L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. La reproduction d'images dans des publications imprimées ou en ligne ainsi que sur des canaux de médias sociaux ou des sites web n'est autorisée qu'avec l'accord préalable des détenteurs des droits. En savoir plus

Terms of use
The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. Publishing images in print and online publications, as well as on social media channels or websites, is only permitted with the prior consent of the rights holders. Find out more

Download PDF: 11.07.2025

ETH-Bibliothek Zürich, E-Periodica, https://www.e-periodica.ch

Le SIDOS, une archive de données pour les sciences sociales

Reto Hadorn

1. Le SIDOS

Le SIDOS, Service suisse d'information et d'archivage de données pour les sciences sociales, est une *fondation de l'Académie suisse des sciences humaines et sociales (ASSH)*. Cette académie fédère une quarantaine d'associations professionnelles, parmi lesquelles figurent certaines associations d'historiens. Financée par la Confédération par le biais de l'Office fédéral de l'éducation et de la science, elle a pour fonction principale de soutenir les associations dans leurs activités, notamment en subventionnant des congrès et des publications.

Le SIDOS est conçu comme un service *à disposition de l'ensemble des sciences sociales*, prises ici dans leur extension la plus grande. Il serait hors de propos de débattre de la question de savoir si l'histoire fait partie des sciences humaines ou sociales ... ou si ces deux catégories sont mutuellement exclusives. L'histoire de l'homme étant, par la force des choses, également l'histoire des sociétés, les efforts accomplis par les historiens pour comprendre les sociétés du passé ou le passé de notre société intéressent nécessairement les sciences sociales.

En activité depuis 1993, le SIDOS remplit essentiellement *deux tâches*: il réalise un inventaire de la recherche en sciences sociales en Suisse et procède à l'archivage de données de recherches dans le but de les rendre plus facilement et plus largement accessibles.

L'inventaire de la recherche en sciences sociales.
Le SIDOS réalise *un inventaire de la recherche en sciences sociales*, basé sur une enquête annuelle auprès d'instituts et de services de recherche universitaires, privés et publics ainsi que sur des échanges d'informations avec des institutions qui réalisent des relevés analogues dans des secteurs spécifiques (sciences de l'éducation et de l'environnement par exemple). Cet inventaire comprend aujourd'hui près de 4000 descriptions de projets et est accessible sous différentes formes: un catalogue online (http://www-sidos.unine.ch), un CD-ROM et un fascicule qui peuvent être commandés au SIDOS. Le fascicule ne comprend toutefois que les projets nouveaux annoncés dans l'année.

L'archive de données
La deuxième tâche centrale du SIDOS consiste à solliciter d'instituts universitaires, d'instituts privés ou de services administratifs *le dépôt dans ses archives de données susceptibles d'être réutilisées*, que ce soit à des fins de formation, pour en approfondir l'analyse ou pour procéder à des analyses secondaires.
L'archivage des jeux de données obtenus comporte plusieurs aspects:
❶ un contrôle de la cohérence entre les données proprement dites et la documentation qui les accompagne (validation du jeu de données, promotion de standards de documentation)
❷ l'entretien d'un dispositif de conservation à long terme:
  – conservation de copies multiples en des lieux distincts
  – standardisation des formats et des supports propres à faciliter l'utilisation des données dans les conditions les plus diverses et surtout à opérer la transition vers des générations ultérieures de systèmes informatiques.
❸ la constitution d'un dossier structuré aisément transmissible
❹ la diffusion d'informations sur les données disponibles par le biais d'un catalogue imprimé et d'un catalogue online qui peut être consulté à l'adresse ci-dessus.
Dans le cadre de cette journée de travail, c'est essentiellement des activités d'archivage du SIDOS qu'il sera question.

Quelle contribution le SIDOS peut-il apporter aux activités de recherche des historiens utilisant ou produisant des données sur support informatique ?
Deux directions doivent être explorées:
❶ la mise à disposition de données archivées, au SIDOS ou dans des archives de données étrangères
❷ l'archivage de données produites par les historiens.

2. Mise à disposition des données archivées au SIDOS

Les données archivées au SIDOS sont *à disposition de tous les chercheurs* intéressés à en approfondir l'exploitation ou à les intégrer dans une analyse secondaire. Nous appliquons naturellement les *restrictions* parfois demandées par les auteurs des données: les données peuvent être déposées sous embargo en attendant que les premières publications aient paru ou n'être transmises à l'utilisateur que sur l'autorisation expresse de l'auteur. L'utilisateur doit pour sa part signer une *déclaration* par laquelle il

s'engage à ne pas utiliser les données à d'autres fins que d'analyse, à ne pas les transmettre à des tiers en-dehors du groupe de projet concerné et à citer ses sources dans les règles – nous proposons d'ailleurs un format de citation approprié aux jeux de données.

L'information sur les données disponibles peut être trouvée dans le *catalogue de données* du SIDOS. Une seconde édition a été imprimée ce printemps; l'information est également présentée sur le serveur web du SIDOS sous la forme d'une base de données facile à interroger.

Les activités d'archivage n'ont véritablement démarré au SIDOS qu'il y a deux ans – il a d'abord fallu mettre sur pieds et stabiliser l'enquête sur la recherche; l'éventail des données disponibles est donc nécessairement restreint. Une *centaine de jeux de données* sont actuellement disponibles, pour la plupart issus d'enquêtes relativement récentes, du moins à l'échelle des temps historiques. Quelques exemples:[1]

- Atlas socio-économique de la région insubrienne 1970–1990 (comprend les données disponibles pour toute la Suisse)
- Auswirkungen der Flexibilisierung von Arbeitszeiten und der Technisierung von Haushalten auf Freizeitsnutzung, Alltagsorganisation und Sozialbeziehung – 1991
- Attitudes politiques en Suisse – 1975
- Politische Aktivierungsereignisse in der Schweiz in den Jahren zwischen 1945 und 1978
- Eidgenössische Abstimmungen: Gemeindeergebnisse ab 1981.

Cette liste d'exemples est biaisée en ce sens qu'ont été choisis des jeux de données où la *dimension temporelle* est présente, ou tout au moins l'idée d'un recul temporel. Il faut pourtant bien dire que le recul d'une archive aussi jeune ne peut pas être bien grand. Lors d'un prochain workshop de l'Association *Histoire et informatique* sur les sources de données, disons dans vingt ans, le SIDOS sera en mesure d'alimenter beaucoup plus richement ceux d'entre les historiens qui s'intéressent à l'histoire récente.

3. Données archivées dans des archives de données étrangères

Le SIDOS fait partie de plusieurs *réseaux internationaux d'archives de données pour les sciences sociales*. Le Conseil européen des archives de données pour les sciences sociales (connu sous le nom de CESSDA,

1 Le mélange des langues dans cette liste de titres vient rappeler l'attitude particulière de la Suisse à l'égard du plurilinguisme: comme tout ne peut être traduit, il faut parfois pouvoir compter avec la capacité des uns à comprendre la langue des autres ...

Council for European Social Science Data Archives) se montre particulièrement actif dans les domaines de l'échange de données et la définition de standards d'archivage et de documentation. Il a notamment réalisé un serveur web qui assure un accès aisé aux serveurs de toutes les archives dans le monde qui offrent de l'information sur Internet (http://www.nsd.uib.no/cessda/europe.html; voir figure 1).

Un bref coup d'oeil suffit pour constater que la distribution des archives reproduit de manière très sèche la division nord-sud: s'il y a 11 archives de données en Europe, avec une dominante nord-européenne, 13 en Amérique du Nord, il n'y en a que 4 dans le reste du monde: Israël, l'Afrique du Sud, l'Australie et la Nouvelle-Zélande, liste à laquelle il faudrait encore ajouter l'Inde, sans accès au réseau.

Le chercheur intéressé par l'histoire récente (pour ne pas dire immédiate ...) trouvera peut-être des données utiles en consultant les catalogues de ces archives. A sa demande, le SIDOS peut lui apporter un soutien logistique, obtenir auprès des archives des informations complémentaires, voire décharger le demandeur de la commande et du transfert des données.

Les archives de données historiques

Il existe également des services spécialisés dans l'archivage de données historiques. Voici trois exemples européens:

❶ Zentrum für historische Sozialforschung, Köln, D
❷ Netherlands Historical Data Archive, Leiden, NL
❸ Historical Data Service, Colchester, UK

A vrai dire, il s'agit dans les trois cas de départements faisant partie des archives de données pour les sciences sociales: le Zentralarchiv de Cologne, le Steinmetz Archiev d'Amsterdam et The Data Archive à Colchester.

Chacun de ces dispositifs a ses particularités. Le *Centre de Cologne* se présente sous les traits d'un *centre de recherche* rattaché à l'archive de données. C'est une manière de souligner que l'archivage de données historiques informatisées demande un investissement particulier qui s'apparente à un travail de recherche historique. *L'archive de données historiques néderlandaise*, créée au sein du Département d'histoire de l'Université de Leiden n'a été rattachée à l'archive de données en sciences sociales (Steinmetz, Amsterdam) qu'en 1997; elle va d'ailleurs rester à Leiden, pour bénéficier pleinement du dynamisme du département d'histoire qui l'a créée.

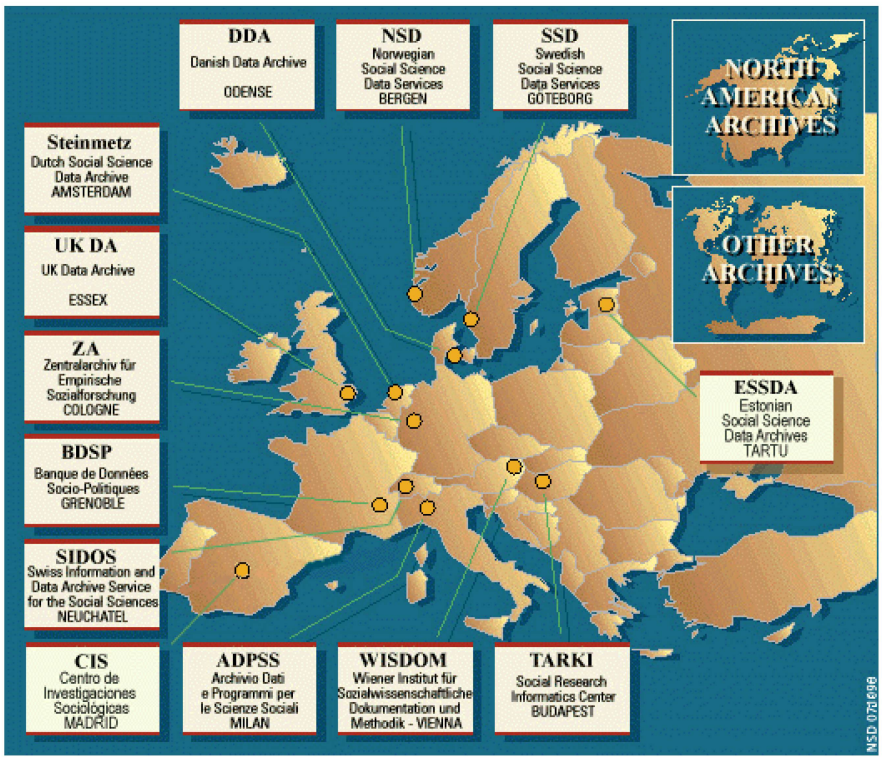

Fig. 1: Archives de données pour les sciences sociales dans le monde. Accès par le Web: http://www.nsd.uib.no/cessda/europe.html.

Pour sa part, le *Service de données historiques de l'archive anglaise* est porté à la fois par l'«Economic and Social Science Research Council» (ESRC), soutien principal de l'archive de données, et par l'«Arts and Humanities Data Service», un réseau d'institutions qui se préoccupe par ailleurs de la conservation de données concernant l'archéologie, les arts visuels et les arts du spectacle. Les trois solutions expriment le souci de traiter les données historiques de manière adéquate, en portant toute l'attention voulue à leur spécificité.

Les *catalogues* de toutes ces archives de données historiques sont accessibles par le biais du réseau Internet. En Angleterre et en Allemagne, les données historiques sont intégrées dans le catalogue général de l'archive. Dans l'interface d'interrogation de l'archive anglaise, il est cependant possible de sélectionner les «Historical Studies» (410) et le serveur de Cologne permet de télédécharger la liste des jeux de données disponibles – liste dans laquelle il suffit de sélectionner les numéros supérieurs à 8000 (~130). *L'archive de données historiques hollandaise*, moins intégrée encore à l'archive de données des sciences sociales, propose son propre catalogue. A la différence des deux autres archives de données historiques, il inclut des jeux de données accessibles qui ne sont pas déposées dans ses fonds.

Voici quelques exemples de jeux de données trouvés dans les divers catalogues. Les titres ne décrivent bien sûr pas le contenu précis des données; ils ont surtout pour but de mettre en évidence l'étendue des périodes couvertes et la diversité des thématiques.

Zentrum für historische Sozialforschung, Köln, D:
http://www.za.uni-koeln.de/data/data-collections/dbvsrch.htm[2]
– Kirchenbücher von Gross-Lüben 1648–1800
– Exiltheater und Exilliteratur
– Wissenschaftsindikatoren ab 1850
Netherlands Historical Data Archive, Leiden, NL:
http://oasis.leidenuniv.nl/nhda/data/dat_orig.htm[3]
– Dutch-Asiatic Shipping, 1595–1795
– Budget and nutrition during the 1930s
– Trade in the port city of Thessaloniki, 1880–1939

2 Télédécharger le fichier dbv(date).zip et ne conserver que les numéros >8000.
«dbv»: Datenbestandverzeichnis – il ne s'agit que d'une table des matières; pour les descriptions complètes, il faut consulter la page du catalogue (Datenbestandskatalog).

3 Pour consulter les détails, prendre un cours accéléré de hollandais. Mais on peut au moins consulter deux listes de titres en anglais, l'une consacrée aux jeux de données répertoriés mais non archivés, l'autre aux jeux de données effectivement archivés.

Historical Data Service, Colchester, UK:
http://hds.essex.ac.uk/
- Mathematical Women in the British Isles, 1878–1940
- Agricultural Rent in England, 1690–1914
- Political Power in Boston, Massachusetts and Charleston, South Carolina, 1818–1843

Si les corpus de données disponibles dans les différents pays accordent une large place à l'histoire nationale, il arrive aussi que d'autres pays soient représentés, comme le montrent certains exemples ci-dessus.

Un rapide sondage dans les collections de données des trois archives révèle que les données résultent pour une part de la saisie informatique de *données numériques disponibles* ou de la codification d'informations relativement standardisées au départ (registres commerciaux ou paroissiaux par exemple). Des *données statistiques agrégées* relevées sur de longues périodes sont également disponibles. Les données *textuelles brutes ou codifiées* sont fréquentes aussi, les *bases de données* complexes plus rares.

L'archive hollandaise de données historiques a développé un centre de compétence en matière de scanning et de reconnaissance des caractères. Aujourd'hui, une bonne moitié de son personnel travaille sur mandat d'autres institutions pour saisir du matériel historique, qui pourra d'une part être conservé plus sûrement sous cette forme et d'autre part être exploité à l'aide des instruments que l'informatique met aujourd'hui à la disposition du chercheur.

4. L'archivage de données historiques en Suisse

Le chemin le plus sûr, si l'on souhaite que le SIDOS puisse mettre à disposition plus de données intéressant les historiens, c'est encore d'y déposer des données historiques! Les historiens sont donc invités à envisager le dépôt à Neuchâtel des données informatisées qu'ils produisent, afin de les rendre plus aisément accessibles à un plus large public de chercheurs.

Le dépôt de données au SIDOS
Le SIDOS a été créé dans la perspective d'archiver en priorité des données quantitatives produites par ces recherches en sciences sociales, dont le prototype est l'enquête représentative. Les recommandations faites aux

chercheurs pour la préparation de leur *dossier* sont naturellement marquées par ce choix:[4]

❶ un ou plusieurs fichiers de données (formats divers acceptés, à discuter)
❷ les informations sur la structure des fichiers de données et les valeurs que peuvent prendre les variables (peuvent être intégrées au fichier de données)
❸ les instruments de collecte de l'information et les conventions de codification s'il y a lieu
❹ la description générale du projet sous une forme plus ou moins détaillée selon les besoins (grille de référence: le questionnaire utilisé par le SIDOS pour son enquête annuelle sur la recherche en cours)
❺ de manière générale, tout document de travail susceptible d'éclairer le nouvel utilisateur des données.

Le SIDOS attache une importance toute particulière à la *documentation*, qui seule peut garantir une bonne compréhension des données et, partant, une interprétation fondée. Le chercheur qui dépose des données aux archives doit se mettre dans la position d'un utilisateur qui n'a pas participé à la construction des données pour se mettre en situation de transmettre toute l'information utile. Il y a généralement un *travail d'édition* à réaliser sur le matériel disponible: mise en forme, contrôles croisés entre documentation et données, explicitation de l'information latente (celle que le chercheur tire de sa familiarité avec la matière première). Ce véritable travail de publication prend naturellement du temps. Mais lorsqu'un dépôt aux archives est envisagé à un stade précoce du projet, il est le plus souvent possible de le maintenir dans des limites raisonnables.

Et l'archivage de données historiques ?

Selon la nature des données, *des solutions appropriées doivent être trouvées* de concert par l'auteur des données et le SIDOS. C'est toujours le cas pour des données textuelles et plus généralement les données qualitatives (voir la brochure citée plus haut). C'est intentionnellement que les pièces du dossier à déposer au SIDOS ont été nommées ci-dessus dans des termes très généraux: étant donnée la variété des sources utilisées par les historiens et les différentes formes et procédures de codification et de digitalisation possibles, ces principes généraux doivent être adaptés de cas en cas. Le

4 Voir la brochure *Archivage de données. Préparation des données et de la documentation* (en allemand: *Daten-Archivierung: Aufbereitung der Daten und der Dokumentation*), SIDOS, Neuchâtel, mai 1995. L'information est également disponible sur le serveur Web du SIDOS.

dossier déposé aux archives doit donc notamment comporter les informations complémentaires suivantes:

❻ la description complète des procédures de digitalisation des documents originaux et des transformations éventuellement subies par les informations relevées

❼ l'inventaire complet des sources utilisées et de la sélection éventuellement opérée pour constituer le corpus étudié.

Des archives spécialisées dans les données historiques apporteraient certainement les solutions les plus appropriées aux données historiques présentées sur support informatique. A l'échelle de la Suisse, une telle spécialisation d'un organisme d'archivage est difficilement réalisable. Par contre, des collaborations ad hoc peuvent être mises en place pour assurer aux données historiques les meilleures conditions d'archivage possibles dans des archives de données généralistes telles que le SIDOS. Des financements complémentaires peuvent, dans certains cas, être trouvés. Le SIDOS peut également réaliser certaines tâches sur mandat et recourir à des historiens pour les réaliser.

Leere Seite
Blank page
Page vide