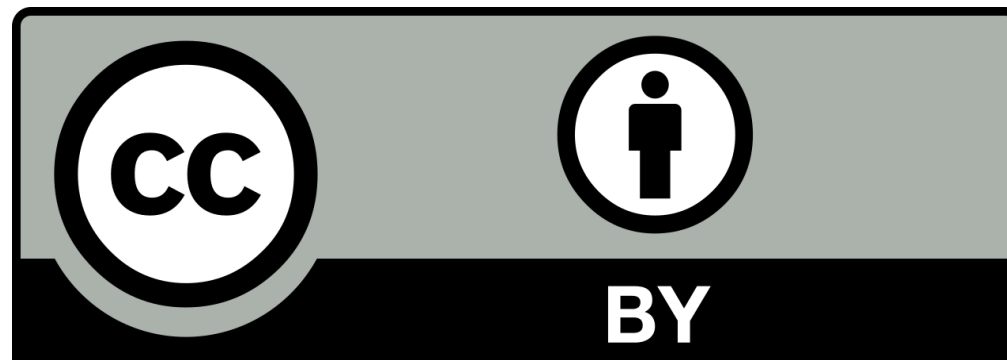


Patrimoine numérique & Science Ouverte

Anne Baillot, Le Mans Université

9 avril 2021, séminaire Covid 19-Museum



Au programme:

- I. Ordonner et partager le patrimoine en contexte numérique
- II. Interopérabilité et reproductibilité
- III. Principes et infrastructures de la Science Ouverte
- IV. Vers les bonnes pratiques de demain

I. Patrimoine et mémoire collective

- Stratégies de patrimonialisation: dimension politique; concept de *Erinnerungskultur*
- Empreinte de l'historiographie du XIXème siècle et nationalisme infrastructurel
- Redéfinition des contours CHI/GLAM dans un contexte international favorisé par le web
- Structuration par projets: mise en réseaux et financements par actions

I. Métadonnées

- Institutions patrimoniales = championnes de la structuration des métadonnées
- Infrastructure pré-existant au numérique
- Systèmes étonnamment peu standardisés
- Fascination rétrospective pour les artefacts de classement non-numériques
- Pas d'imitation dans le passage au numérique (contrairement au paradigme de la « page » du livre encore largement émulée)



I. L'exception qui confirme la règle

- Le Zettelkasten de Luhmann: <https://niklas-luhmann-archiv.de/bestand/zettelkasten/tutorial>

Niklas Luhmann-Archiv

Person und Theorie


Online-Bestand

Nachlass

Über das Projekt

Aktuelles

Suche im Online-Bestand

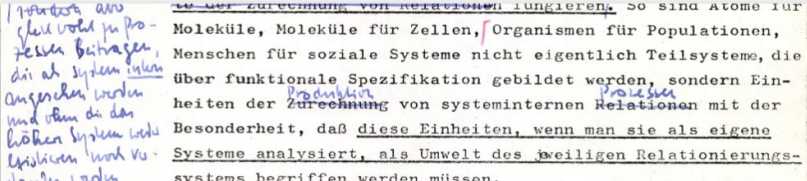


Zettelkasten

Luhmanns zentrales Arbeits- und Denkwerkzeug war ein Zettelkasten, den er über 40 Jahre lang gepflegt hat und der schließlich ca. 90.000 Zettel umfasste. Er besteht aus zwei weitgehend getrennten Sammlungen: Zettelkasten I (ca. 1952-1963) mit 7 Auszügen und Zettelkasten II (ca. 1963-1997) mit 20 Auszügen. Diese Sammlungen sind hier in einer digitalisierten und editierten Version einsehbar. Dabei stehen verschiedene Einstiegsinstrumente (Übersichten, Register) sowie eine Volltextsuche zur Verfügung. (Datenbestände im Aufbau)

ZETTELSTUCHEINHALTSÜBERSICHTAUSZUGSÜBERSICHTSCHLAGWORTEPERSONENBIBLIOGRAPHIE

TUTORIALLESEWEGE




Manuskripte/Typskripte

Luhmanns Nachlass enthält eine große Zahl von Manuskripten, die meistens als maschinenschriftliche Typskripte oder als Ausdrucke von Computerdateien vorliegen. Zum einen handelt es sich um bislang unpublizierte Monographien und Aufsätze sowie Vorlesungs- und Vortragsskripte aus den 1960er bis 1990er Jahren, zum anderen um Vorfassungen bereits veröffentlichter Werke. Die Texte werden hier als digitalisierte Faksimiles zugänglich gemacht. (Seite im Aufbau)

MONOGRAPHIENAUFSÄTZE

VORLESUNGENVORTRÄGETUTORIAL




Audio/Video

Luhmanns umfassende Lehr- und Vortragstätigkeit wurden seit den 1980er Jahren häufiger auf Audio- oder Videomitschnitten dokumentiert; außerdem existieren einige Interviewaufzeichnungen. Diese Aufnahmen sind hier abrufbar. (Datenbestände im Aufbau)

ALLE DOKUMENTEVORLESUNGEN

VORTRÄGEINTERVIEWS




Bestandsregister


Das Bestandsregister bietet einen Zugang zu allen Materialien des Forschungsportals. Zur Zeit kann auf die durchsuchbare Datenbank zugegriffen werden, die alle Literaturangaben aufführt, die im Portalbestand erwähnt werden. Mittelfristig wird das Bestandsregister durch ein Schlagwort- und ein Personenregister ergänzt. (Datenbestände im Aufbau)

LITERATURSCHLAGWORTEPERSONEN


TUTORIAL




PROFESSUR FÜR
KULTUR- UND
Kommunikationswissenschaft
UNIVERSITÄT BIELEFELD



UNIVERSITÄT
BIELEFELD



BERGISCHE
UNIVERSITÄT
WUPPERTAL



CC
DH

Die Datenbestände des Forschungsportals werden fortlaufend erweitert. Entsprechend unterliegen die Einträge insbesondere in den Bestandsregistern und die Verlinkungen in und zwischen einzelnen Datenbeständen einem ständigen Ergänzungsprozess.

Version 4 (26.01.2021)

[Impressum/Kontakt](#) [Datenschutzhinweise](#) [Einen Fehler melden](#)

I. L'archivage par la métadonnée

- Génération de métadonnées=fonction primordiale de l'archive (mise en ordre)
- Standardisation numérique => mise en place d'arbres normalisés
- Forme d'autorité propre aux institutions patrimoniales
- Approche collaborative des métadonnées issues d'institutions patrimoniales techniquement possible, mais peu mise en œuvre: pourquoi?

I. Exemple de viaf: <http://viaf.org/>

- Contributions issues des fichiers d'autorité des différents pays du monde, portés par les bibliothèques nationales
 - Biais important: surreprésentation des métadonnées bibliothécaires et archivistiques
 - Contributeur allemand: la DNB (dnb.de)
 - Ouverture à l'enrichissement des données par les projets de recherche à forte composante biographique
 - Système complexe de vérification des informations
- => partage d'autorité entre institutions patrimoniales et chercheur.es sur les métadonnées.

I. Exemple de partage des métadonnées biographiques

Link zu diesem Datensatz	http://d-nb.info/gnd/118627813
Person	Voltaire
Geschlecht	männlich
Andere Namen	Arouet, François-Marie (Wirklicher Name) Voltaire, François-Marie Arouet Voltaire, Franciscus de Voltaire, J. F. M. A. de Voltaire, F. M. A. von Voltër, Fränswä Märyä Aruwë da Voltair Völtërs Voltaire (falsche Namensform) Arouet de Voltaire Fu, Ertai Fuertai Fu er tai Vadé, Antoine (Pseudonym) Vadé, Catherine (Pseudonym) Vadé, Guillaume (Pseudonym) Ralph, docteur (Pseudonym)



<https://de.wikipedia.org/wiki/Wikipedia:BEACON>



↳ Mentions dans cette édition

Brief von Adolf von Buch an Louis de Beausobre (Dresden, 16. und 18. Dezember 1763) ([p. 1](#))
Brief von Johann Albrecht Euler an Jean Henri Samuel Formey (Sankt Petersburg, 1. März 1767) ([p. 2](#))
Brief von Johann Albrecht Euler an Jean Henri Samuel Formey (Sankt Petersburg, 22. Februar 1768) ([p. 2](#))

↳ Mentioned as a part of a group

AUTRES SOURCES

↳ PND-BEACON

Personennamendatei (PND)
Deutsche Biographie

CCAB ID
Great Russian Encyclopedia Online ID
Wikipedia-Artikel
NNDB people ID
Encyclopædia Britannica (2)
Cultureel Woordenboek ID
Cultureel Woordenboek ID
Biographie und Bibliographie bei "Klassiker der Geschichtswissenschaft" bei historicum.net [Voltaire]

<https://www.berliner-intellektuelle.eu/entity?p0079+fr>

I. Données pas vraiment données

- Echange d'informations entre institutions la plupart du temps problématique
 - Informations pertinentes pour les institutions patrimoniales et pour recherche différentes
 - Travail pour faire converger les modèles de métadonnées comme point de départ
 - Recherche: attribution d'autorité sur des champs de métadonnées, inusité dans les institutions patrimoniales.
- >métadonnées: situation malgré tout relativement libre car pas de droits
- Données: question des droits (processus de numérisation, images, objets)

I. Données nativement numériques

- Multiplication des initiatives autour du COVID
- Source de prédilection: Twitter (par ex. @coronarchiv)
- Initiatives d'abord de moissonnage, puis de valorisation
- Exemple 2020: Early Corona Twitter Dataset <https://hal.archives-ouvertes.fr/hal-02861167/document>
- Exemple 2021: expérience (« super-experiment ») « bibliothèques et COVID »: <https://twitter101dh.github.io/>
- Question des données numériques nativement privées

II. Quel est le problème?

- Souvent peu de visibilité sur ce qui sépare donnée privée et donnée publique => problème de fond (trop vaste pour mon propos)
- Convergence en F autour de l'Open Data (culture, administration, recherche)=> point de départ ici
- Deux questions:
 - est-ce qu'on veut travailler en commun (patrimoine+recherche) sur des données?
 - est-ce qu'on peut?

II Interopérabilité

- Question de l'interopérabilité
- Syd Bauman, 2011: « Balisage Paper: Interchange vs. Interoperability »
<https://www.balisage.net/Proceedings/vol7/html/Bauman01/BalisageVol7-Bauman01.html>
- > plaide pour le « blind interchange »
- Chercher des solutions pour échanger des informations dans les deux sens: besoin d'un langage stable

II. Choix techniques dans la communauté des Humanités Numériques

- Choix techniques ancrés dans l'histoire de la communauté des humanités numériques textuelles: XML comme langage de base (lisible par l'homme et par la machine)
- Spécification par la Text Encoding Initiative depuis les années 1990
- Développement de modèles de données standardisés correspondant aux besoins des sciences du texte (description de manuscrits, de phénomènes textuels)

II. Exemple de fichier XML-TEI

```
<text>
  <body>
    <div type="transcription">
      <pb facs="PPN73975727X\PHYS_0001" n="1"/>
      <note hand="#pencil_1" place="top-margin-right" type="foliation">5</note>
      <note hand="#pencil_2" place="top-margin-left">34.1823</note>
      <p><unclear cert="medium" reason="illegible"><del rend="strikethrough">1</del></unclear></p>
      <opener> 3<space extent="3w"/><dateline rend="align(right)">Sonntag Abend 10 Uhr</dateline>
        <salute>Mein lieber einziger <hi rend="latin"><persName ref="#p0237">Chamisso</persName></hi>, </salute></opener>
      <p rend="nolb">so nenn ich Dich <del rend="overwritten">in</del><add place="across">ja</add> in<lb/>der <del rend="overwritten">
        >r</del><add place="across">R</add>egel, d<del rend="overwritten">a</del><add place="across">r</add>um will <del
        rend="overwritten"><unclear extent="1w" reason="illegible"/></del><add place="across">ich</add> es auch jetzt thun. Es
        ist<lb/>mir, als könnte ich nicht zu Bett, gehn, ohne Dir eine<lb/>schöne <del rend="overwritten"><unclear extent="1c"
        reason="illegible"/></del><add place="across">g</add>ute Nacht zu wünschen, und daß Du eben so<lb/>schön schlafen möchtest
        als ich, ob<del rend="overwritten">b</del><add place="across">g</add>leich ich diese<lb/>Nacht einen ungenehmen und doch zugleich
        angenehmen<lb/>T<del rend="overwritten"><unclear extent="2c" reason="illegible"/></del><add place="across">raum</add> gehabt
        haben, ich habe nehmlich auf <persName ref="#p0249">Varn<lb/>harens</persName> Schoß gesessen, und <add place="above">er</add> hat
        mich zu meinem größ<lb break="no"/>ten Aerger geliebkost; <add place="above">und in Deiner Gegenwart</add> Dann aber habe ich Dich
        um<lb break="no"/>arm – und wir waren recht froh als er uns verlassen<lb/>hatte.</p>
      <p rend="nolb">Heute, oder diesmal bekommst Du viel zu<lb/>lesen, welchen Du wohl wirst zu erst vernehmen, <rs ref="#p3174"
        >Mullen</rs> habe<del rend="strikethrough">n</del> ich, <del rend="overwritten">s</del><add place="across">w</add>eil <add
        place="above">ich</add> es <del rend="strikethrough">ha</del> vor Neugier nicht habe aushalten<lb/>können, lesen müssen.</p>
      <p>Wenn ich nur erst wieder<lb/>Nachricht hatte, ob, wann, und wie Du angekommen bist.<lb/><rs ref="#p3174">Mullens</rs> Brief sieht
        viel hübscher aus, als meiner<lb/>ist auch sehr zärtlich, ich werde am Ende ganz neidisch<lb/>werden auf sie.</p>
      <p><persName ref="#p1754">Ernst</persName> hat mir gestern recht Kummer gemacht, er sehr neidisch<lb/>und häßlich gegen den kleinen
        <persName ref="#p2219">Johannis Horkel</persName>, es scheint<lb/>wirklich aus einem häßlichen Zug seines Charakters her <hi
        rend="subscript">zu</hi><lb/>komen, der mich sehr betrübt, den ich fürchte er wird<lb/>schwer aus zu rotten sein. <space
        extent="1w"/>
        <persName ref="#p3128">Max</persName> macht uns hingegen allen<lb/>recht viel Freude.</p>
      <p>Dienstag werde ich mit <rs ref="#p0239">Ede</rs> und Umschweif<note resp="#henrike.bobzien">Diese Aussage bezieht sich auf die
        später durchgeführte und beschriebene Reise zur Pfaueninsel. <persName ref="#p0239">Julius Eduard Hitzig</persName> und
```

II. La TEI et les autres formats

- Echange de données possible avec des formats archivistiques comme EAD (exemple dans EHRI: <https://www.ehri-project.eu/ehri-for-institutions>)
- Solution stable et pérenne mais coûteuse en temps et en énergie => multiplication des solutions plus rapides, moins pérennes, moins interopérables dans le contexte du web sémantique

II. Structures de sauvegarde

- Question de la pérennité: implique le format mais aussi le stockage et sa maintenance
- Contexte français favorable avec [huma-num](#)
- Dimension européenne sous-jacente: huma-num= branche française de [DARIAH](#)



DARIAH ERIC: A network to enhance and support digitally enabled research and teaching across the Arts and Humanities

Description

The Digital Research Infrastructure for the Arts and Humanities (DARIAH) aims to enhance and support digitally-enabled research and teaching across the arts and humanities. DARIAH is a network of people, expertise, information, knowledge, content, methods, tools and technologies from its member countries. It develops, maintains and operates an infrastructure in support of ICT-based research practices and sustains researchers in using them to build, analyse and interpret digital resources. By working with communities of practice, DARIAH brings together individual state-of-the-art digital arts and humanities activities and scales their results to a European level. It preserves, provides access to and disseminates research that stems from these collaborations and ensures that best practices, methodological and technical standards are followed.

II. Penser la temporalité

- Penser à des usages futurs de ses données
- Impact sur les formats et sur les modèles de données
- Exigence souvent incompatible avec les injonctions des agences de moyens
- Projection de questions de recherche diverses sur des corpus => modifie les modes d'accès/portes d'entrée sur ces corpus

II. Les exigences de reproductibilité

- Reproductibilité: concept globalement étranger aux approches herméneutiques traditionnelles dans les sciences molles
- Reproductibilité: norme vécue comme étant imposée par les sciences dures
- Nécessite d'anticiper l'accès structuré aux données (sur le plan technique, juridique)
- Nécessite donc de valoriser le travail de constitution de corpus, actuellement quasiment pas intégré aux critères d'évaluation

III. Grands principes de la Science Ouverte

- Concernent données primaires et données secondaires
- Régulent les modes de publication et donc d'accès aux données

-> site web: <https://www.ouvrirlascience.fr/>

- « Fermeture » variable selon les disciplines
- Doctrine des principes FAIR dominante
- Mise en œuvre systématique dans l'ESR



- NB Il existe un petit guide exprès pour les doctorant.es:
<https://www.ouvrirlascience.fr/passeport-pour-la-science-ouverte-guide-pratique-a-lusage-des-doctorants/>

III. Rôle des infrastructures publiques

- Soutien à cette politique sous la forme de plateformes
- Voir plateformes du CCSD: HAL, sciencesconf, épisciences, sciencescall
- Soutien à des dispositifs de publication de revues, blogs et ouvrages numériques comme la plateforme OpenEdition
- Projets européens dans laquelle la France (et les Pays-Bas pour tout ce qui est GLAM) fait partie des bons élèves
- Institutions patrimoniales: infrastructures propres pour leurs fonds numériques
- Favorise la réutilisabilité et la valorisation des données patrimoniales dans la recherche => besoin d'une relation de confiance réciproque

III. Charte de réutilisation des données patrimoniales

- Objectifs: favoriser les bonnes pratiques, l'utilisation des données patrimoniales par la recherche, l'intégration des résultats de la recherche aux institutions patrimoniales
- Heritage Data Reuse Charter: <https://www.dariah.eu/activities/open-science/data-re-use/>
« The basis of the charter is a moral contract, the [Mission Statement](#) all stakeholders must adhere to. It comprises of six core principles: Reciprocity, Interoperability, Citability, Openness, Stewardship, Trustworthiness. »
- Mission Statement: <https://datacharter.hypotheses.org/77>

IV. Les bonnes pratiques pour valoriser les données patrimoniales dans la recherche: step 1, DMP

- Plans de gestion de données vécus comme une contrainte, sont en réalité une chance
 - Première version en amont de la collecte, amendé ensuite
 - Permet d'avoir une vision d'ensemble du type de données, leur archivage, leur traitement, leur(s) utilisation(s) possible(s)
 - Impose de se poser des questions en prise sur l'interprétation et pas juste sur la technique: modèle d'annotation par exemple
- > <https://doranum.fr/plan-gestion-donnees-dmp/>

IV. Step 2: valoriser la recherche sur corpus par les Data Papers

- Type de publication encore peu connu en SHS
- Permet de valoriser le travail sur données
- Publication dans des revues spécialisées (catégorie de plus en plus souvent proposée/ouverte)
- Peut facilement s'appuyer sur le contenu du DMP (donc ne nécessite pas beaucoup de travail, valorise la collecte structurée des données)

IV. Step 3, publier dans des épirevues

- Travers des publications revues papier traditionnelles: opaque, long, invraisemblablement compliqué à trouver (peu d'exemplaires) et à citer...
- Sonnette d'alarme tirée massivement sur les dérives dans le système de publication de revues => déclaration de San Francisco DORA (San Francisco Declaration on Research Assessment: <https://sfdora.org/read/>)
- Palette de solutions comporte les épirevues (*overlay journals*) => plus besoin de maison d'édition
- Plateforme française: épisciences: <https://www.episciences.org/>

IV. Step 4, penser à notre empreinte carbone

- Sauvegarde et échange des données: empreinte environnementale importante
- Arbitrages politiques encore peu clairs; tendance au greenwashing et à la responsabilisation individuelle

⇒ besoin d'une prise de conscience et d'une action collectives

- Voir les travaux du collectif labos1point5: <https://labos1point5.org/>
- Important: prendre en compte impact carbone, impact biodiversité et impact H2O dans la consommation d'énergie, notamment numérique.

Merci beaucoup pour votre attention!

anne.baillot@univ-lemans.fr

Twitter: @AnneBaillot

<https://cv.archives-ouvertes.fr/annebaillot>