

Archivage des données de la recherche

Retour d'expérience du laboratoire "Archives Henri Poincaré" & projet DDS

Pierre Willaime

Archives Henri-Poincaré

Love Data Week

Table-ronde

Jeudi 13 mars, Nancy



Au départ il y a (toujours) les sources ;

Et les sources, dans les disciplines historiques, ce sont les archives.

Les archives aux AHP

Des fonds hétérogènes au niveau des documents et des thématiques, exploités avec différentes méthodes et outils pour répondre à des problématiques scientifiques variées.

Un traitement particulier des archives pour la recherche, outillé numériquement

- Un traitement archivistique
 - un investissement important sur le versant humanités numériques,
 - conduisant à la mise en ligne de multiple sites web de valorisation d'archives (26 aujourd'hui),
 - des outils, des visualisations.
- ➔ Un focus qui s'explique par la nature de la structure.



Chaîne de traitement pour les archives

1. Donation/acquisition (convention)
2. Inventaire
3. Établissement plan de classement
4. Description archivistique et organisation du fonds
5. Exploitation et valorisation scientifique

Fonctionnement "théorique". La réalité est plus complexe.

L'exploitation scientifique est en fait présente à chaque étape.

Fonds Nicolas Antoine Hélène Poincaré

FR UMR7117 AP · Fonds · 1825 - 1911 · Ébauche

Nicolas Antoine Hélène Poincaré est né à Nancy, dans la Meurthe-et-Meuse, le 19 juin 1911. Son père est Jacques Jules Nicolas Poincaré (Nicolas Antoine Hélène Poincaré)

Fonds Claude Pair

FR UMR7117 CP · Fonds · 1962 - ?

Claude Pair, né en 1934 à Blâmont (Vosges), est mathématicien, informaticien et physicien. Après des études de mathématiques en classe préparatoire, il s'oriente à partir de 1963 d'abord en physique, puis en mathématiques. Claude Pair

Fonds Ernest Bichat

FR UMR7117 EB · Fonds · 1847 - ?

Ernest Adolphe Bichat naît le 17 septembre 1845 à Lunéville, et meurt le 17 septembre 1911 à Nancy. Par ses professeurs, il poursuit ses études et obtient son baccalauréat en 1863. Ernest Bichat

Fonds Henri Poincaré

FR UMR7117 HP · Fonds · 1854 - 1912

Henri Poincaré

Fonds Isidore Didion

FR UMR7117 ID · Fonds · 1700-1898

Isidore Didion, né à Thionville en 1798 et mort en 1878 à Nancy, est un mathématicien et physicien. Après des études de mathématiques à Pont-à-Mousson, il intègre l'École Polytechnique en 1817. Isidore Didion

Fonds Jean Largeault

FR UMR7117 JL · Fonds · 1930 - 1995

Où sont les données de la recherche là dedans ?

Partout

"Les données de la recherche sont définies comme des enregistrements factuels (chiffres, textes, images et sons), qui sont utilisés comme sources principales pour la recherche scientifique et sont généralement reconnus par la communauté scientifique comme nécessaires pour valider les résultats de la recherche."

Définition de l'OCDE

Dans les documents d'archives

- Conservés aux AHP ou non
- Numérisés ou non



Dans la structuration des fonds

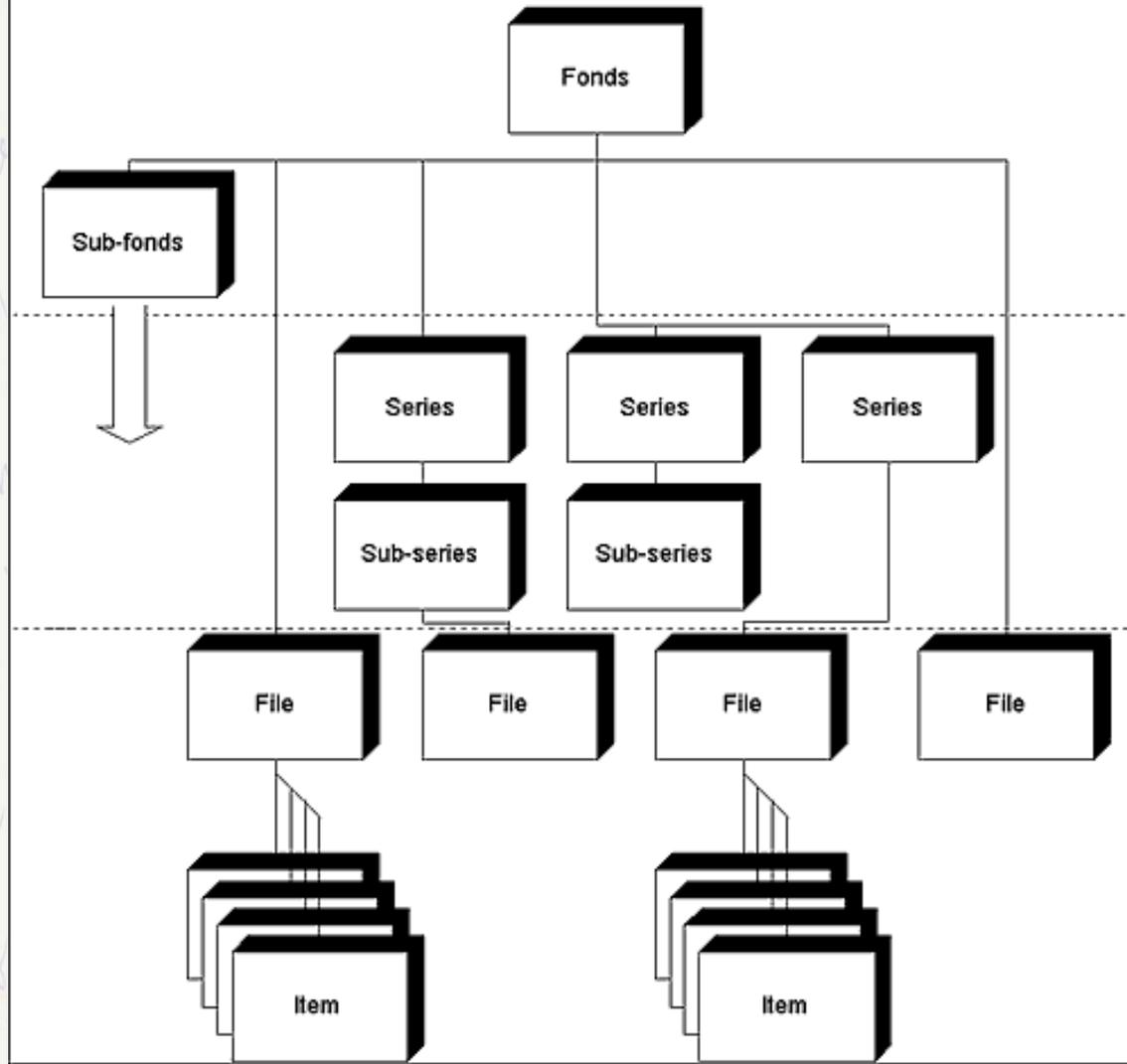
ISAD(G): General International Standard Archival Description – Second edition (1999)



APPENDIX

- A1 The model shows some typical situations and does not include all possible combinations of levels.
- A2 Any number of intermediate levels are possible between any shown in the model.

MODEL OF THE LEVELS OF ARRANGEMENT OF A FONDS



Fonds EB - Fonds Ernest Bichat

- [-] [Fonds] EB - Fonds Ernest Bichat
 - [-] A - Vie personnelle et familiale (1845-1905)
 - + [-] 1 - Photographies
 - + [-] 2 - Correspondance personnelle et politique
 - + [-] 3 - Dossiers de constructions et affaires immobilières
 - + [-] 4 - Éléments généalogiques
 - [-] [Pièce] EB011 - Cahier de coupures de presse
 - [-] [Pièce] EB115 - Reçu de la société de prévoyance et de secours mutuel de Lunéville
 - [-] [Pièce] EB116 - Avertissement d'impôt pour l'année 1909
 - [-] [Pièce] EB124 - Compte de Monsieur Bichat (1900)
 - [-] B - Vie académique et production scientifique
 - + [-] 1 - Articles
 - + [-] 2 - Cours et conférences
 - + [-] 3 - Notes manuscrites et autres documents
 - [-] C - Postérité d'Ernest Bichat et de son entourage

Dans la description des pièces

- Description fine des documents et constitution de bases de données, directement en ligne par les équipes de recherche.
- Choix d'Omeka (Classic puis S) à la fois pour valoriser en ligne mais également pour exploiter.
- Utilisation de schémas de métadonnées riches ("ontologies")
- Construction d'outils d'exploration et de recherche (liés au web sémantique).



LA CORRESPONDANCE D'HENRI POINCARÉ

[Accueil](#)
[Lettres](#)
[Recherche](#)
[Index](#)

LETTRE : Henri Poincaré à Eugénie Poincaré - janvier 1876

Transcription

Métadonnées

Citer ce document

[Janvier 1876]

Ma chère **maman**,

Hier soir j'ai eu trois invitations à diner.

1° Chez **M^{me} Malgaigne**, faite depuis dimanche.

2° Chez **M^{me} Barthélémy** ; j'ai dû refuser, j'ai été chez elle samedi, point ne l'ai trouvée.

3° Chez **M^{me} Brice** chez qui se trouvait l'oncle Charles¹. Envoyé dépêche.

Chez **M^{me} Malgaigne** se trouvaient outre les **Olleris**, les **Lefort**, un **taupin** et sa mère, papillotée² et Pont Lévéquoise³ et enfin un jeune homme très gentil aspirant je crois au conseil d'État. Le diner n'a pas présenté l'aspect lugubre redouté de **M^{me} Olleris**, l'absence des **Pilastre** ayant empêché la conversation traditionnelle sur l'enfantépatancologie Pilastro-Lefortique comparée.

M. Lefort a raconté ses voyages en découpant son poulet comme à l'ordinaire, mais les réussites traditionnelles n'ont pas paru.

Ce matin tentative infructueuse de verglas.

M. Daubrée, inspecteur général des mines, et **Madame Daubrée** resteront chez eux les samedis 22 janvier 4 et 19 février.

1. Il s'agit probablement de **Charles Comon**.↩

2. Il s'agit d'une allusion à la coiffure de cette dame qui avait sans doute les cheveux frisés ou qui portait de petits morceaux de papier (des papillotes) pour faire friser ses mèches de cheveux.↩

3. Une habitante de Pont-l'Évêque ? Le terme exact serait « pontépiscopoise ».↩

Noms cités ▾

Personne citée dans le document primaire.

ahpo:citeName

Eugénie Launois

Contenus
 Collections
 Médias

Françoise Malgaigne

Contenus
 Collections
 Médias

Marie-Anne Barthélémy

Contenus
 Collections
 Médias

Mme Brice

Contenus
 Collections
 Médias

Famille Olleris

Contenus
 Collections
 Médias

famille Lefort

Contenus
 Collections
 Médias

Marie Olleris

Contenus
 Collections
 Médias

Famille Pilastre

Contenus
 Collections
 Médias

Léon Le Fort

Contenus
 Collections
 Médias

Gabriel Auguste Daubrée

Contenus
 Collections
 Médias

Marie Laure Rosalie Daubrée

Contenus
 Collections
 Médias

Charles Comon

Contenus
 Collections
 Médias

+ Texte
 Ressource Omeka
 URI

Mots d'argot polytechnicien cités ▾

Argot polytechnique utilisé dans le document primaire.

ahpo:citeArgot

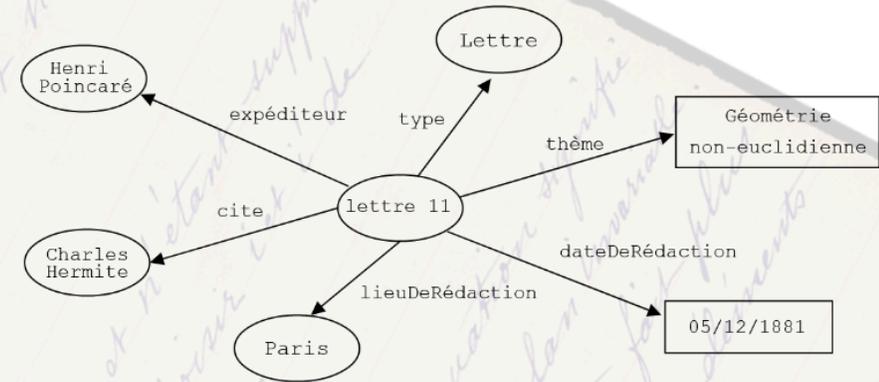
Taupin

Contenus
 Collections
 Médias

+ Texte
 Ressource Omeka
 URI

Le choix du web sémantique

- Métadonnées structurées, liées et réutilisables.
- Des outils et méthodes pour représenter, lier, requêter et raisonner avec des jeux de données.



```

SELECT ?l
WHERE {
  ?l a :Letter .
  ?l :sentBy :henriPoincaré . ?l :sentTo :felixKlein .
  ?l :hasTopic "Géométrie non euclidienne" .
  ?l :hasWritingDate ?y
  FILTER(YEAR(?y) >= 1890 && YEAR(?y) <= 1900)
}
  
```



Outils d'exploration Travaux de Nicolas Lasolle

Initial resource [Update](#)

<http://henripoincare.fr/api/items/5834> ⓘ

Henri Poincaré à Gösta Mittag-Leffler - 29 juin 1881

Search parameters

⊙ All conditions ○ One of

- Sent by Henri Poincaré 1107
- Sent to Gösta Mittag-Leffler 141
- Has subject "Équations aux dérivées partielles et espaces lacunaires" 8
- Written at Paris 1233
- Quotes Thiébaud 1
- Quotes Balthazar Mathis 20
- Has subject "Voyage d'étude de Poincaré en Scandinavie-1878" 2

[Query](#) [Add condition](#) [Generate condition](#)

Writing date

1854 1858 1862 1866 1870 1874 1878 1882 1886 1890 1894 1898 1902 1906 1910 1912

Results: 52 resources [Compact view](#)

Export results [CSV](#) Visualize distribution [Show](#)

March

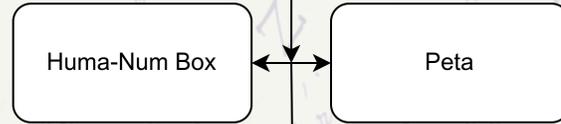
1889-03-01 Quotes Anders Lindstedt, Charles Eugène Delaunay, Gabriel Lamé, Hugo Gyldén, Karl Weierstrass Sent by Henri Poincaré Sent to Gösta Mittag-Leffler Has subject Polémique Poincaré avec Gyldén Details New Search	1889-03-05 Quotes Charles Eugène Delaunay, Hugo Gyldén Sent by Henri Poincaré Sent to Gösta Mittag-Leffler Has subject Polémique Poincaré avec Gyldén, Prix roi Oscar Details New Search	1889-03-24 Quotes Comte de Lewenhaupt Sent by Henri Poincaré Sent to Gösta Mittag-Leffler Has subject Prix roi Oscar Details New Search
--	--	---

Les données de la recherche aux AHP

- Des documents d'archives provenant de sources multiples, numérisés ou non,
- Des descriptions et structuration archivistiques des collections.
- Des enrichissements scientifiques (qui se transcrivent sous forme de métadonnées)
- Des bases de données
- Des modèles de reconnaissance et segmentation HTR

Mais aussi

- Des outils (d'exploration, de requêtage, de transcription (automatisée ou non), d'alimentation automatique de bases de données, etc.)
- Des visualisations



Campagnes de
numérisation



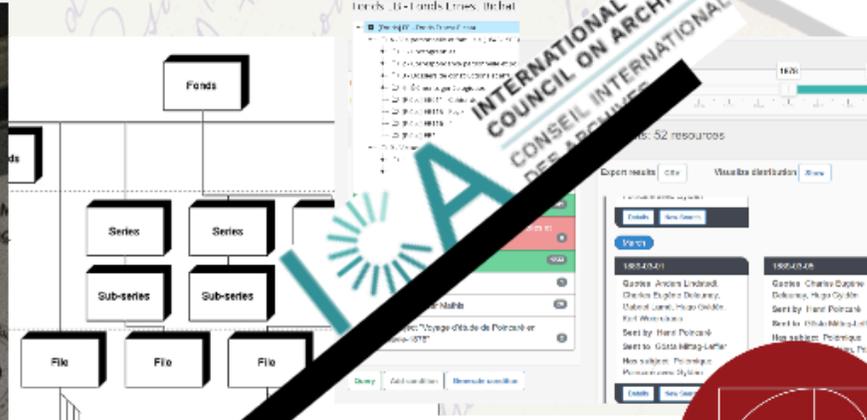
Valorisation / Exploitation



Réutilisabilité



Une séparation inopérante



PatriMaths
Patrimoines et patrimonialisation des mathématiques

HENRI POINCARÉ 1854-1912

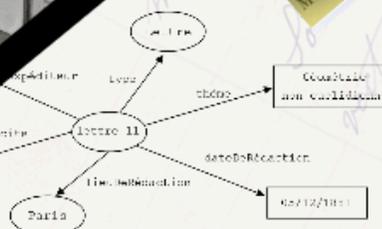
BataData

Accueil | Docu...

Filtres

Collection

- Bataville (538)
- Bataville Journal des collaborateurs des usines Bata en France (199)
- Batapresse (88)
- Bataville Avis et Echos (63)
- Formation (58)
- Bataville pour tous (45)
- Travail et relations professionnelles (37)
- Médiathèque (37)
- Vidéos (32)
- Sports et loisirs (21)



Les procès-verbaux du Bureau des longitudes

Un inventaire en 10 tomes (1793-1902)

1793-1902

1793-1902

1793-1902

EXPLOITATION NUMÉRIQUE ET SCIENTIFIQUE DES ARCHIVES DE JULES BOURBON



International Image Interoperability Framework

*intraire des valeurs
Z. ce qui démontre
de la dévotion
et*

*signifie
l'existence d'une infinité de
la 3^{me} sorte.*



Le projet DDS (Durability of Digital Storage)

*initiation de
Z. ce*

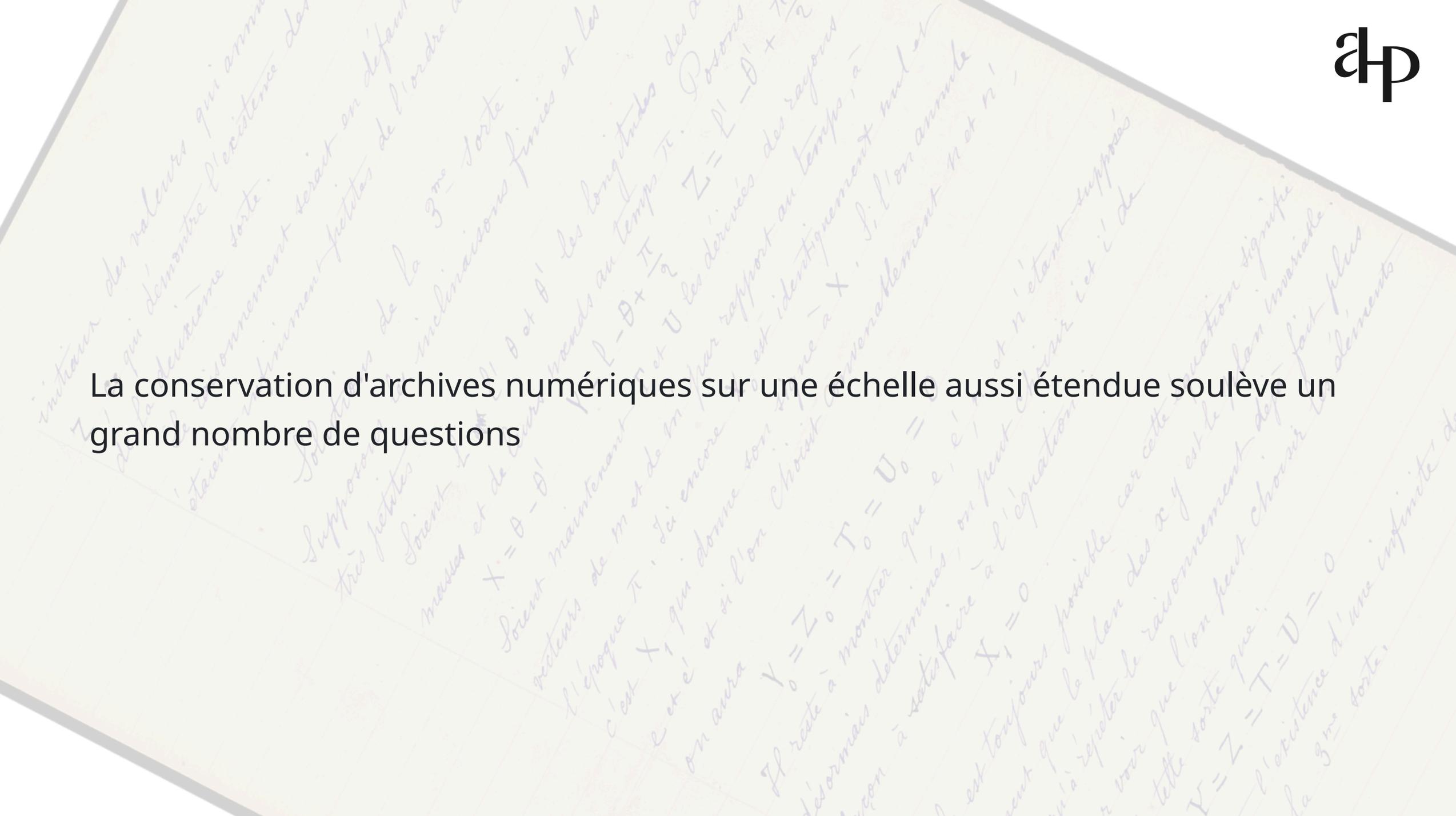
*plus
éléments
une infinité de
sortes.*

Constat

Pas la même échelle temporelle en informatique et dans les disciplines historiques.

- Stockage "long terme" en informatique = (au mieux) 30 ans.
- Préservation du patrimoine culturel = siècles, millénaires.
- La préservation des documents d'archives physiques est une problématique connue et éprouvée, bien que complexe à mettre en œuvre rigoureusement,
- il n'en est pas de même pour la conservation des artefacts numériques.

La conservation d'archives numériques sur une échelle aussi étendue soulève un grand nombre de questions



DDS

Un (petit) projet théorique et exploratoire.

- pour formaliser les questions
- pour entendre les acteurs de la conservation numérique sur ce sujet.

Projet rendu possible par la MSH Lorraine puis le pôle CLCS.

Question 1 : Les supports de stockage

Problem of discrete modifications of digital data, whether intended or not

the problem of digital data preservation is not simply reducible to the preservation of the corresponding binary code. One equally needs to preserve the corresponding software (especially for multimedia data) and sometimes specialised hardware.

Question 2 : La multiplication des artefacts numériques et la charge de la conservation

Intrinsically digital distributed archives : what to preserve when there are only copies? (NFT)

blockchain and NFT technologies aiming at reconstructing the concept of original ownership in a digital environment could lead to new forms of accountability

Question 3 : Aspects sociaux et cognitif de la conservation numérique à long terme

It should be stressed that on the top of the various technical issues related to the durability of digital data the problem of durability of digital archives has important social, cognitive and other human-related dimensions. Indeed, the archived data are always supposed to be in some sense meaningful to those who wants to preserve them. The endurance of digital data (along with the necessary soft- and hardware) does not, by itself, guarantee that the preserved data remain meaningful over time.



PROJECT DDS

DURABILITY OF DIGITAL STORAGE

Journée de lancement

"Préserver le patrimoine numérique à travers les siècles"

13. AVRIL 2023

9h-18h



<https://dds.hypotheses.org>

Contact : pierre.willaime@univ-lorraine.fr

Merci

pierre.willaime@univ-lorraine.fr