



Chambre
des Députés
GRAND-DUCHÉ
DE LUXEMBOURG

Is artificial intelligence more efficient and safer than humans?

Addressing the challenges of using artificial intelligence to identify sensitive data in the archival collections of the Chamber of Deputies of Luxembourg

Sarah Marcq, research engineer, Université Paris 1 Panthéon-Sorbonne

Amandine Gorse, records manager, Chambre des Députés



**Chambre
des Députés**
GRAND-DUCHÉ
DE LUXEMBOURG

Introduction: Framing the project

AI and archival challenges
Ambitions
Use case: Inventory
Main constraints

Responding to the constraints

Operational and technical
response
Tactical answer

Lessons learnt

What can we achieve with
LLMs ?
Ethical considerations

Conclusions

A responsible use of AI
A procedure to minimise risks
Some resources

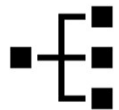


Chambre
des Députés
GRAND-DUCHÉ
DE LUXEMBOURG

Introduction: Framing the project

How can AI meet today's challenges in the archives' world ?

Availability of digital archives



- Archiving digital records, hierarchical structures
- The "**dark archives**" threat

Discoverability



- An inventory not only to make archives available but also **findable**
- AI has proven to be a great tool for the **exploration** and **mediation** of public data in cultural heritage
→ Digital records ? Sensitive data ?

Sensitive data



- The impact of **digitalization**
- **Definition** issues

Ambitions

Context

The law of August 17th, 2018, introduced specific access restrictions for each type of sensitive data.

A national-level description framework was established by the National Archives of Luxembourg (ANLux).

Sensitive data still needs to be identified within the Chamber's documents.

Goals

Automate archival inventory using artificial intelligence.

Objectives of the inventory :

- Produce archival **description** to support research and discovery
- Identify **sensitive data** to determine appropriate **access restriction periods**.

Benefits expected from the project :

- Ensure **legal compliance** in the communication of archival material
- Promote **transparency** within the Chamber of Deputies
- Provide a **description** of preserved collections
- **Reduce processing time and workload** for the archivists

Use case : generating the inventory

Introduction

Metadata

- Creation period: from
- Creation period: to

Title and Description

- Title
- Description

Sensitive Data

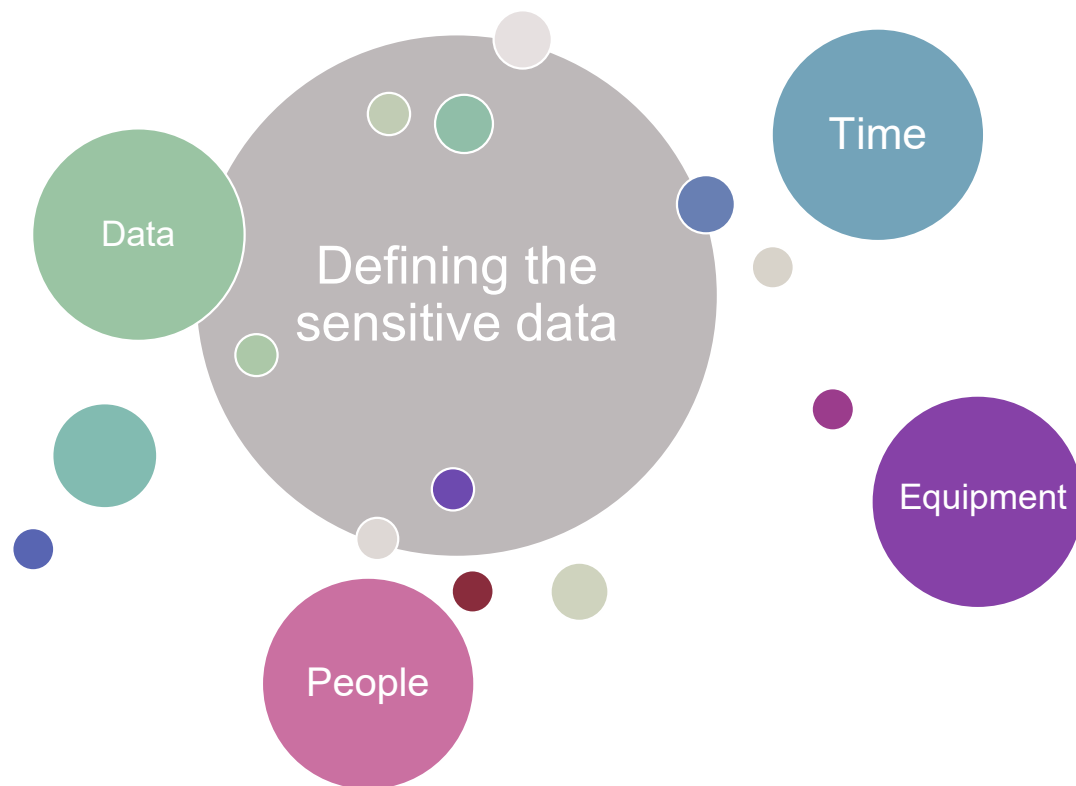
- Personal data
- Civil status records
- Notarial deeds
- Breach of foreign relations, national security or public order of the Grand Duchy
- Tax secrecy
- Cases brought before judicial, extrajudicial or disciplinary bodies
- Prevention and investigation of criminal offences
- Commercial and industrial data

Out of Scope

- Reference code
- Physical location
- Classification plan identifier
- Retention schedule series code
- Copyright information

Main constraints

Introduction



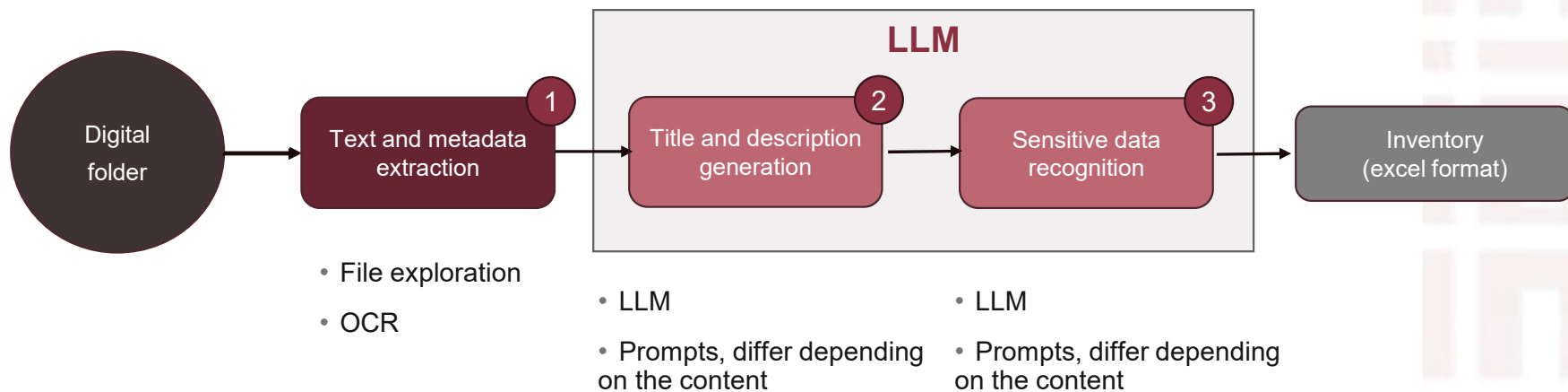


Chambre
des Députés
GRAND-DUCHÉ
DE LUXEMBOURG

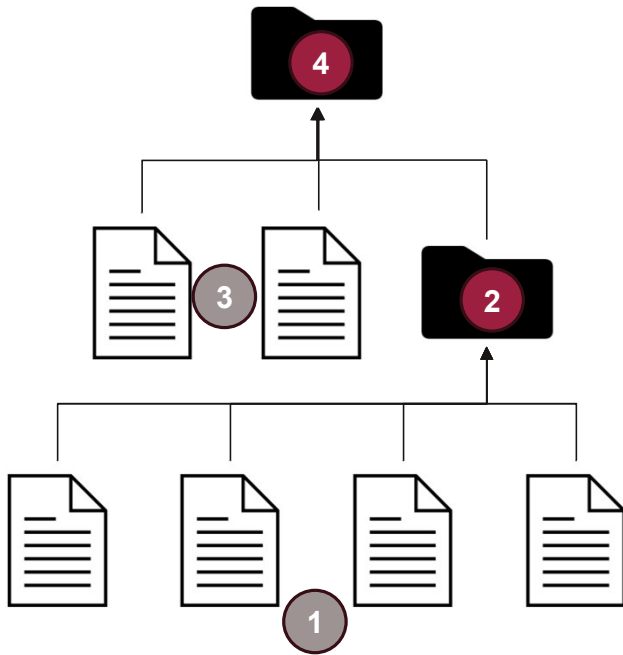
Responding to the constraints

Operational and technical response - App

- Python/Flask app
- Llama 3 7B, 8 and 4 bits *quantization*



Operational and technical response - Process



- 1 Generation of a title and description for **files** at the lowest level of the hierarchy
- 2 Generation of a **folder** title and description based on the titles, descriptions, and metadata of the files it contains
- 3 Generation of a title and description for **files** at the higher level
- 4 Generation of a **folder** title and description at the higher level, based on the titles, descriptions, and metadata of the files and folders it contains

Operational and technical response - Prompt

Input :

```
{  
  "path": "C:\\User\\Chemin\\du\\dossier_exemple\\document.txt",  
  "type": "file",  
  "fileName": "document.txt",  
  "creationDate": "Wed May 20 10:26:26 2024",  
  "modificationDate": "Wed May 29 10:35:06 2024",  
  "mimeType": "text/plain; charset=UTF-16LE",  
  "texte": "Ceci est le texte d'un document txt.",  
  "generatedTitle": "to be defined"  
}
```

File containing text +
metadata
(step 1)

Context :

You are an activist at the Chambre des députés of Luxembourg. You have been given a **set of metadata for a document** formatted as json as input. You are working on an archival inventory project.

Question :

Please generate one **title (max 20 words)** in french for this document for the inventory. Examine well the whole content for this. Please only write the **title** in french, nothing else, so I can directly reuse it in my inventory. Do not explain your answer or tell me about alternatives. I only need the title. Do not invent information, only use the ones you have. It's ok if your answer is minimalistic because you do not have enough info.

Title in french :

Tactical answer

- Create a high-impact strategy based on a culture of archiving and digital transition
- Providing a clear definition of sensitive data
- Points of attention :

Skateholders

- Training, *change management*
- External expertise

Equipment

- IT Infrastructure
- Nationwide collaboration

Data

- Quality
 - Quantity
- } AI data center



Chambre
des Députés
GRAND-DUCHÉ
DE LUXEMBOURG

Lessons learnt

What can we achieve with LLMs ?

Titles and descriptions

71% Correct
22% Incomplete/Imprecise
6% Incorrect

Sensitive data

79% Correct
6% Debatable
15% Incorrect

- Is AI system more efficient than human?
 - Time saving,
 - Processing of a very large set of data in a single operation,
 - Consistent and uniform application of rules,
 - Reduction of bias and subjectivity,
 - Traceability of decisions made by the AI system,
 - The human error rate is around 4% but it can raise to 25% in the case of complex files. AI system reduces human error from 30% to 60% in certain use cases.

It is difficult for humans to process mass-generated data by machines (automation of tasks goes hand in hand with automation of the documentary production).

Ethical considerations

Lessons

Safer than human when...



- ☐ Humans are a potential risk of leakage
- ☐ Humans may be threatened into revealing sensitive information
- ☐ Humans make mistakes

Safer than human if...



- ☐ Privacy and security are addressed
- ☐ Safety is guaranteed
- ☐ The quality of evaluation is monitored
- ☐ Best practice are part of project management



Chambre
des Députés
GRAND-DUCHÉ
DE LUXEMBOURG

Conclusions

Towards a responsible use of AI system

From theory...	... To practice
Controllability: monitor and steer AI system behaviour	<ul style="list-style-type: none"> Automatic and human evaluation methods Verification of results by regular quality controls and whenever diffusion of the archives is requested.
Privacy and security: appropriately obtaining, using and protecting data and models	<ul style="list-style-type: none"> Revision of the register of processing operations to include the collection and use of data by AI Systems. Curated dataset Data are not used to improve models
Safety: preventing harmful system output and misuse	<ul style="list-style-type: none"> System operated only by trained archivists and Data stored on a sovereign Luxembourgish cloud controlled by the State and use of a national supercomputer (Private endpoint maintained by State) Data and models are encrypted
Explainability: understanding and evaluating outputs generated by AI system	<ul style="list-style-type: none"> Track of usage metrics, KPI to demonstrate LLM is accurate.
Transparency: enabling stakeholders to make informed choices	<ul style="list-style-type: none"> All decisions are documented (use cases, AI system design, etc.)
Governance: introducing best practice into the project management	<ul style="list-style-type: none"> Risks analysis, impact analysis, KPI monitoring, quality management, data sourcing
Ethics: apply ethical safeguards, rise awareness on the ecological cost	<ul style="list-style-type: none"> Prompts are assessed for cognitive bias and attack modalities. Ecological cost is calculated, and compensation measures are proposed to management.

A procedure to minimise risks

1) Transparency

- Enabling the user who requested the archives to make informed choices.
- The user chooses the procedure (with or without AI processing).

2) Sharing the risks

- User is made responsible under a contract with Parliament for not disseminating sensitive data that they have identified.
- User must notify the Parliament if they find sensitive data in the documents.

3) No publication allowed

- Certain sensitive archives funds are never open to requests (HR funds because of their very long life cycle).
- Publication of the archives is never allowed if the AI has been used to grant an exemption.

Procedure
based on trust

Make the user
responsible

Double check

Some resources

Jason Franks, “Can ChatGPT and other large language models streamline records management?”, in *IRMS Bulletin*, Sept. 2023.

Adrian Cunningham with David Fricker, “AI and Archival Description”, *ICA tutorials*, video recorded on 17 June 2024, <https://www.ica.org/fr/resource/ai-and-archival-practice-on-line-tutorials/>

Viridiana García Marignón, Ernesto Ibarra Sanchez, with Claudia Escoto, “AI, Ethics and Human Rights”, *ICA tutorials*, video recorded on may 2024, <https://www.ica.org/fr/resource/ai-and-archival-practice-on-line-tutorials/>

J. Curzon, T. A. Kosa, R. Akalu and K. El-Khatib, “Privacy and Artificial Intelligence” in *IEEE Transactions on Artificial Intelligence*, vol. 2, no. 2, pp. 96-108, April 2021.

J. R. Baron and N. Payne, "Dark Archives and Edemocracy : Strategies for Overcoming Access Barriers to the Public Record Archives of the Future", *2017 Conference for E-Democracy and Open Government (CeDEM)*, Krems, Austria, 2017, doi : 10.1109/CeDEM.2017.27.



Chambre
des Députés
GRAND-DUCHÉ
DE LUXEMBOURG

Thank you very much for your attention



PANTHÉON SORBONNE
UNIVERSITÉ PARIS 1

Sarah Marcq

Experte en ingénierie
numérique
DIREVAL : Direction de la
recherche et de la
valorisation

Téléphone: +33144078000
Centre Ulm | 1 rue d'Ulm -
75005 PARIS - FRANCE

www.panthéonsorbonne.fr



Amandine GORSE

Service secrétariat général et archives



Chambre
des Députés
GRAND-DUCHÉ
DE LUXEMBOURG

23, rue du Marché-aux-Herbes | L-1728 Luxembourg
T (+352) 466 966-650
agorse@chd.lu | www.chd.lu