

# CENL “AI in Libraries” Network Group

**Jean-Philippe Moreux**  
CENL AGM, 19 Oct 2020

## **Libraries and AI:**

- Libraries have expertise on data and data is central in modern AI.
- A lot of experimentation and some achievements.
- Even large institutions can't experiment in all areas.
- Barriers exist (access to technical skills, transition from legacy systems to AI driven processes).

**Needs for collaboration and sharing!**

# Objective

As a reminder, the objective of this group are:

- To leverage the use of AI through **better understanding and use**.
- To promote **sharing, collaboration and standardisation** on AI between institutions.
- To identify areas of AI with **particular significance for libraries**.

# The members

## How are we going to achieve this?

8 institutions, 9 members (so far)



**Jean-Philippe Moreux, Gallica  
Scientific Advisor**  
Bibliothèque nationale de France /  
National Library of France



**Svein Arne Brygfeldt, Head of Ai-  
lab**  
Nasjonalbiblioteket / National Library  
of Norway



**Max Kaiser, Head of Research  
and Development**  
Österreichische Nationalbibliothek /  
National Library of Austria



**Neil Fitzgerald, Head of Digital  
Research**  
British Library



**Jane Makke, Head of Library  
Services Development  
Department**  
Eesti Rahvusraamatukogu / National  
Library of Estonia



**Love Börjeson, Director of KBLab**  
Kungliga Biblioteket / National  
Library of Sweden



**Petr Kukak, Lead of Digitisation  
and Digital Collections**  
Národní knihovna České republiky /  
National Library of the Czech  
Republic



**Ralph Marschall, Digitalisation  
Project Manager**  
Bibliothèque nationale du  
Luxembourg/National Library of  
Luxembourg

+1



**Yves Maurer, Lead on  
Webarchiving program**  
Bibliothèque nationale du  
Luxembourg/National Library of  
Luxembourg

+1

# The actions

- Publication of an '**AI Cookbook for Libraries**', based on real use cases, designed in a way that allows a deep understanding of the core technologies and the test and reproducibility of the technical implementation (2020-2021).
- Launch of a **AI survey** (January 2021), to be sent to the CENL Network, about AI projects, initiatives, and roadmaps that are already active; it is hoped that this survey will also contribute to enlarging the membership of the Network Group with new member libraries.
- Publication of a '**Strategy and Roadmaps**' report, to help practitioners and managers to think about and build their own AI strategy (2021).
- 6 weeks frequency e-meetings (sharing, planning). Strong coordination with the **AI4LAM** international initiative

- A **Github Organisation** is on line (May 2020). It will disseminate the AI Cookbook, the Strategy and Roadmaps report, the CENL Survey results.
- **Workshops, panels and hands-on** will be organized to leverage the Cookbook and the Strategy resources. Some international events had already been listed as opportunities (Fantastic Futures 2021, December 2021; NewsEye european project final conference, March 2021)
- **CENL Letter and CENL website** for institutional communication.
- Slack channel or similar if membership increases



Search or jump to...

Pull requests Issues Marketplace Explore

CENL-Network-Group-AI / Recipes

Code Issues Pull requests Actions Projects Wiki Security Insights Settings

## AI Cookbook

CENL-AI-WG edited this page 2 hours ago · 43 revisions

### List of AI Recipes

The recipes are classified by technical domain, and for each recipe, classical use cases in libraries are provided.

## T

### Natural Language processing (NLP)

#### Recipe #1: Named entity extraction & linking

- enrichment of digital collections (creating new metadata such as person names, organizations, locations) for information retrieval, scientific objectives, etc.
- establishing links between documents and authority data

#### Recipe #2: Topic modeling

- understanding large collections of unstructured text documents (text mining use case)
- enrichment of digital collections with topics (information retrieval)

#### Recipe #3: Text classification, computational semantics, language models

- enrichment of digitized collections with genres (novel, poetry, science, ...) or other classification schema (Dewey...)
- cataloguing of born-digital materials

#### Recipe #4: OCR post-correction

- correction of ocerized collections
- on the fly correction of OCR



### Optical Character Recognition (OCR)



Search or jump to...

Pull requests Issues Marketplace Explore

CENL-Network-Group-AI / Recipes

Watch 1 Star 1

Code Issues Pull requests Actions Projects Wiki Security Insights Settings

## Images Classification Recipe

CENL-AI-WG edited this page 2 hours ago · 64 revisions

Edit New Page

### Images Classification on going

Pages 1

*Convolutional Neural Networks (CNN) and Transfer Learning for Heritage Images Classification*

Keywords: image classification

Approaches: convolutional neural networks, transfert learning

Tools: IBM Watson Studio, Google Cloud AutoML, Google Colab, Tensorflow, Python, IIIF

NLP

HTR

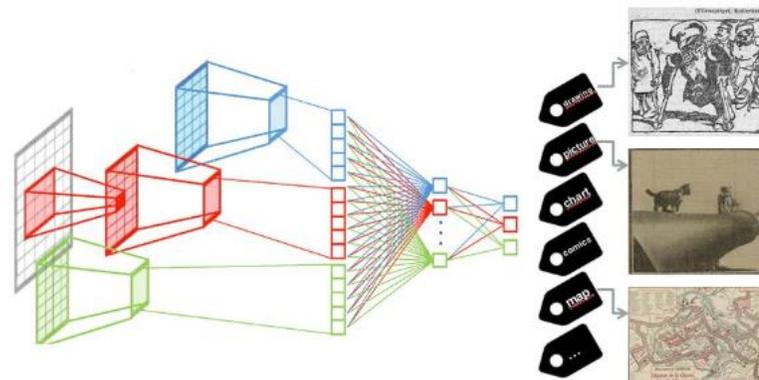
Document Analysis

Computer Vision

- Image Classification

Recipe Template

This recipe exposes an image classification scenario, aiming to deduce the technique or genre of heritage images (picture, drawing, map...) using a Convolutional Neural Networks pretrained model (i.e. a supervised approach with transfer learning). We leverage the classification ability of these models, trained to detect objects, to use them on our documents classification scenario.



Clone this wiki locally

<https://github.com/CEILL-1>

This recipe may have various library use cases, particularly for cataloguing and information retrieval systems.

*This recipe is based on BnF materials*

### Goals

This recipes includes a basic introduction to neural networks and deep learning (formal neuron model, neural networks, convolutional neural networks, transfer learning) and a hands-on session:

# Thanks!

jean-philippe.moreux@bnf.fr

svein.arne.brygfjeld@nb.no

<https://github.com/CENL-Network-Group-AI>

The screenshot shows the GitHub repository page for 'CENL - AI in Libraries Network Group'. The repository description states: 'AI in Libraries is a network group of The Conference of European National Librarians'. The page includes navigation tabs for Repositories (3), Packages, People (10), Teams, Projects, and Settings. A search bar is present with the text 'Find a repository...'. The main content area features three repository cards: 'Strategy' (AI strategy development), 'Survey', and 'Recipes' (AI Cookbook for Libraries). On the right, there are sections for 'Most used topics' (ai, cenl, digital-libraries) and 'People' (10 members). An 'Invite someone' button is located at the bottom of the 'People' section.