Judiciary Investigations on the AI Era

Daniel Camara – daniel.camara@gendarmerie.interieur.gouv.fr

Daniel Camara – daniel.camara@gendarmerie.interieur.gouv.fr

Center for Forensic Artificial Intelligence – CFIA/PJGN

Objective of the presentation

- Introduce some present and possible uses of AI for investigative tasks

Center for Forensic Artificial Intelligence – CFIA/PJGN

Objective of the presentation

FIRST POINT to pay attention:

- The problems of law enforcement agencies are not technological but human.

- We are not capable of treating the existing volume of data available!

Objective of the presentation

FIR... ...T: ...pay attention:

- ...

- ...

volume of data avail...

> "Before, It was hard to find clues; today is hard to discriminate what is important and what is not! Too much data is available!!!"
>
> -- GN investigator

AI declaration

CHARTE IA
AU SEIN DE LA
GENDARMERIE NATIONALE
Pour une IA de confiance au service de la sécurité

LOYAUTE
La gendarmerie s'engage à faire évoluer les systèmes d'IA afin de conserver un niveau de performance élevé et une utilisation intègre.

CONNAISSANCE
La gendarmerie s'engage à développer la connaissance en IA pour prévenir l'effet boîte noire.

RESPECT
La gendarmerie s'engage à une exploitation de l'IA respectueuse des données à caractère personnel et garante des libertés individuelles.

RESPONSABILITE
La gendarmerie s'engage à utiliser l'IA dans un cadre d'aide à la décision et non en lieu et place d'une décision humaine.

CONFIANCE
La gendarmerie s'engage dans l'utilisation d'une IA de confiance au profit du citoyen comme de son personnel.

TRANSPARENCE
La gendarmerie s'engage à rendre disponible les méthodes comme les évaluations des systèmes employés.
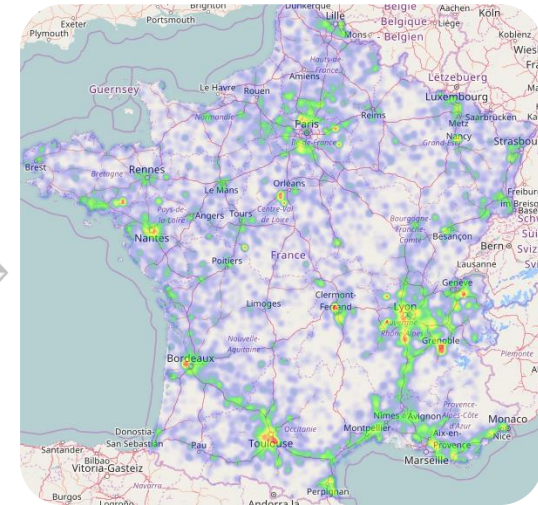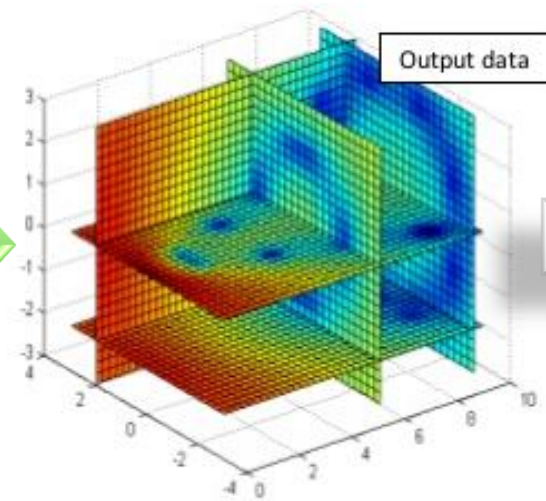
Center for Forensic Artificial Intelligence – CFIA/PJGN

Generic model

Some Important concepts

- ## Error
  - Inherently linked to the methods
  - A model of the reality, by definition, is not THE reality! (probabilities)

- ## Bias
  - "Bias is inherently bad and should be eliminated!"
  - No, Bias is inherent in all data. Current AI works on FINDING bias in data

- ## Explainability
  - Be able to explain HOW the method works and reaches the results. Extremely important in a legal framework
  - How does our brain work? And the other person in front of you?
  - Aren't we being MORE demanding from AI than we are from humans?
  - Some advanced AI models, today act as "black boxes". BUT are not less interesting!
  - Efforts towards explainable AI are crucial but should we wait for that?

Prompt: Please create an image that expresses the main concepts of machine learning. Use the style that you think could fit the best for a nice visual effect. - DALL-E 3 by OpenAI.

Center for Forensic Artificial Intelligence – CFIA/PJGN

Some Important concepts

- Error
  - Inherently linked to the methods
  - A model of the reality, by definition, is not THE reality! (probabilities)

- Bias

> However, we know, humans are never wrong, they do not commit errors!!!!

  - How does our brain work? And the other person in front of you?
  - Aren't we being MORE demanding from AI than we are from humans?
  - Some advanced AI models, today act as "black boxes". BUT are not less interesting!
  - Efforts towards explainable AI are crucial but should we wait for that?

Prompt: Please create an image that expresses the main concepts of machine learning. Use the style that you think could fit the best for a nice visual effect. - DALL-E 3 by OpenAI.

Center for Forensic Artificial Intelligence – CFIA/PJGN

AI and the intelligence cycle

Example of activities that could benefit from AI data treatment
-- The intelligence cycle

The good information at the perfect timming

- Adaptative alerts

- Suggestion of difusion channels

Provide high level correlation of multidimentional data

- Big Data visualisaton

- Correlation between data

Definition of a research plan
- Lessons learned from past events

- Suggestion of needs for new cases, based on past ones

Distinguish noise from useful information
- What to collect on which context

- Which is the source reliability

Automatic data treatments
- Cleaning

- Fusion

- Feature engineering

Planning & direction

Collection

Processing

Analysis and production

Dissemination

CFIA

Gendarmerie nationale

Opportunities and dangers

- Increase the capacity and decrease the time of action
- Amazing opportunities for improving and perfecting services
- The technology is agnostic
  - It does not distinguish usages

The same methods that can be used for creating spotless documents -> undetectable phishing mail

Center for Forensic Artificial Intelligence – CFIA/PJGN

Uses in production - Gendarmerie

- ## ValidIA
  - Verification of the viability of genetic samples
  - Accuracy of 99.16%
  - Decrease the time of treatment for 86 samples from 3:30 hours to 30 min

- ## TAJ – National Database of Criminal Records
  - Fingerprinting identification
  - Facial comparison

- ## Speech to text
  - Transcription of auditions
  - Decrease the time of treatment from 8 hours to 1:30 hour

Center for Forensic Artificial Intelligence – CFIA/PJGN
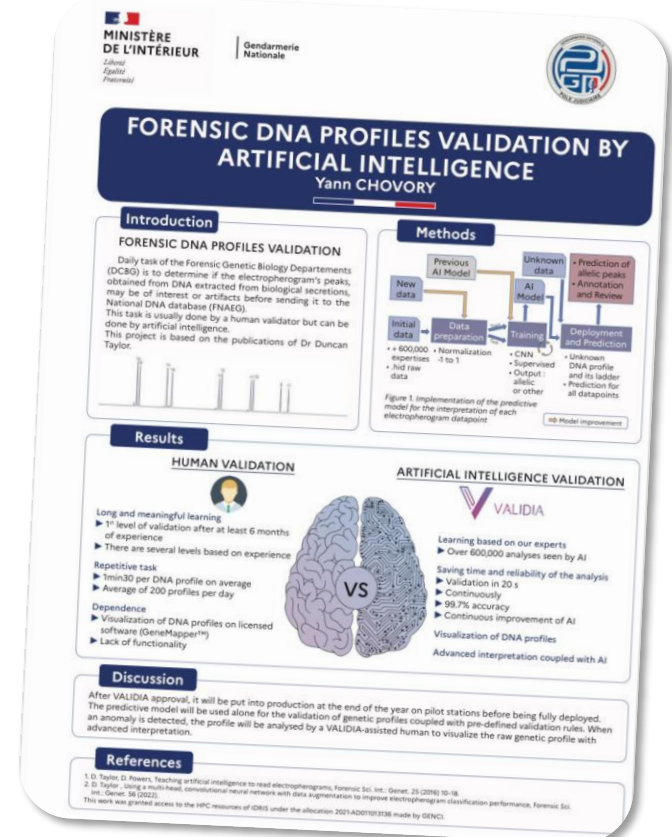
Present available for usage

- Some of the tools we have developed
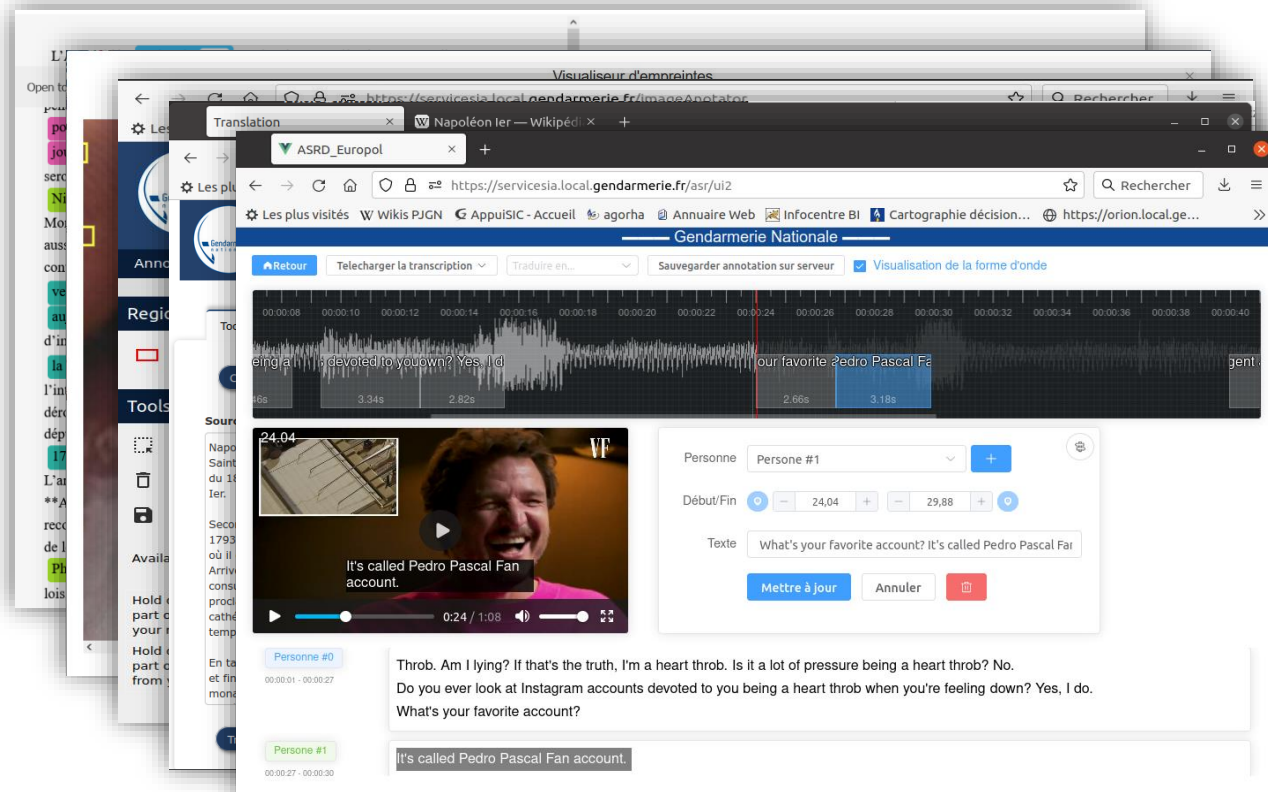
Present available for usage

- Some of the tools we have developed
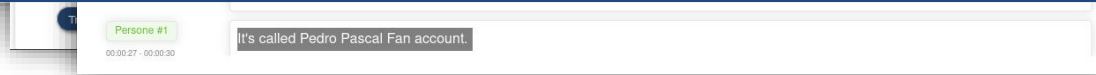
Great fear and source of distress for LEAs personnel

- Are we allowed to use it in the judicial context?
- Will this usage fragilise/invalidate my investigation?
- To what extent can I go?
- Sure I can use, but I will never declare I did!

Are not the same kind of fears we had with the advent of computers?

Persone #1

00:00:27 - 00:00:30   It's called Pedro Pascal Fan account.

Criminality forecast

- Usual Comments
  - I am quite suspicious of it!
  - Did not work in the USA!
  - The past is not enough to predict the future
- Preventive Analysis of criminality
  - Time series analysis methods
  - Probability, not crystal ball
  - Performance of > 80%
  - In experimentation
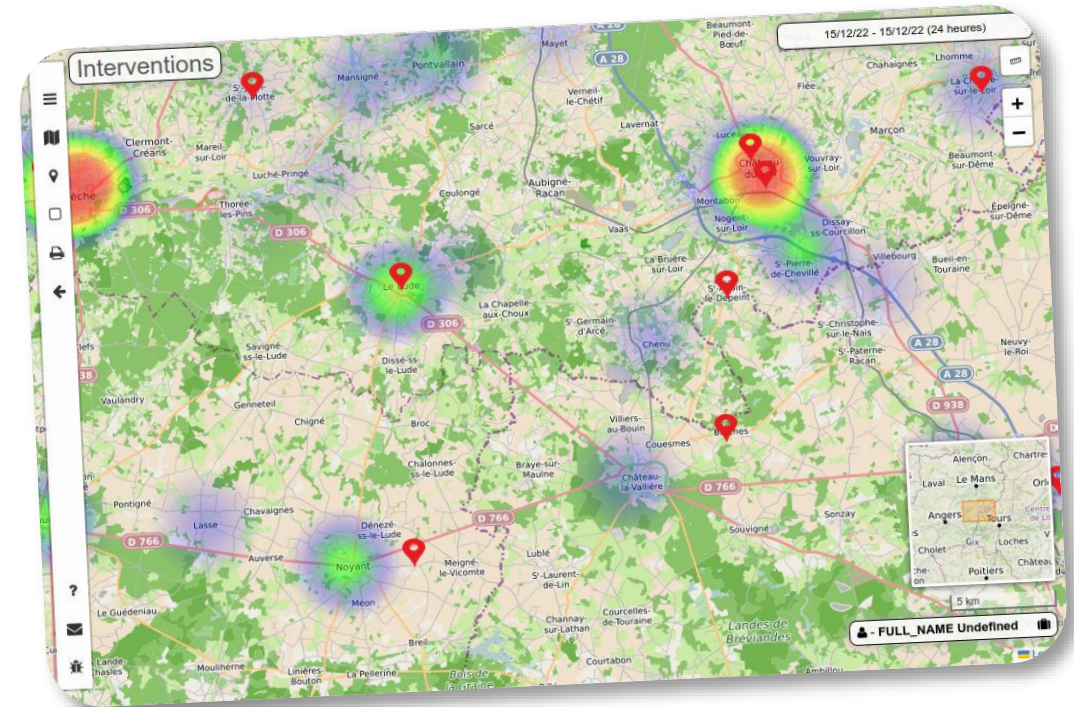
- # Generative Models

  - ## Whole new paradigm

  - ## Comparable to the advent of the Internet

  - ## Can be used as programming tools

    - ### Still underestimated

Prompt: Create an image contrasting traditional media with the internet age, highlighting the change in power dynamics and content quantity. Depict a small-scale publisher's office with limited media, books, and paper, symbolising controlled information dissemination. Contrast this with the internet era, showing a vast array of diverse individuals using digital devices, laptops, phones, and tablets, surrounded by an abundance of varied content and social media icons, representing content democratisation and a massive increase in information. The color scheme should distinctively separate these eras, emphasizing the internet's transformative effect on content quantity and distribution - DALL-E 2 by OpenAI.

- # IDeXtend

- # Retrieval Augmented Generation (RAG)

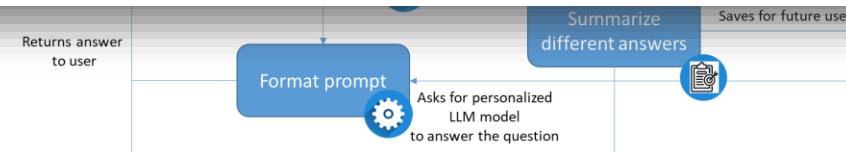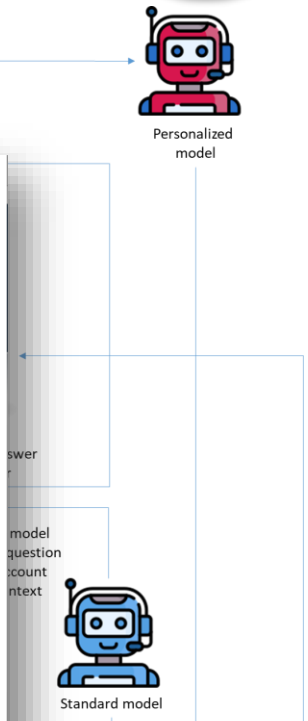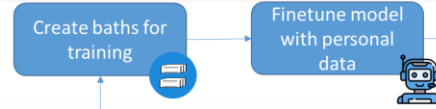 - ## LLMs are quite bad with specific and precise answers

- # Objective

  - ## Create an auxiliar for experts and investigators to help them explore and understand their data

  - ## Old term, Expert System

CFIA

- IDeX

- Retri
  Gene

  - LLM
    prec

- Obje

  - Cre
    inve
    und

  - Old term, Expert System

IDeXtend - Investigative Data Exploration and Extraction Tool

Get abstract file: export-1.pdf

Get abstract file: export1.pdf

Title: IDeXtend: A User-Focused Text and Speech Analysis Tool for Law Enforcement: An Application ProposalThis text is an invitation from Europol for organizations to apply for a grant, in collaboration with the Federal Criminal Police Office of Germany, to develop IDeXtend - an investigative data analysis tool. The project objectives involve creating a user-centric solution for document analysis using advanced natural language understanding and speech-to-text functionalities.IDeXtends features include:1. Feasibility study, prototyping, and deployment.2. Collaboration with co-applicants (e.g., French Gendarmerie) for testing and making the tool available on Europol Tools Platform.3. Advanced text and speech analysis using Large Language Models (LLMs).4. Real-time analytic features.5. User-friendly design, multilingual capabilities, and open-source software use (e.g., Python, Vue.js).Applications require details of the lead applicant and co-applicant(s), adhering to Europols eligibility requirements and data protection rules. The project duration is 7 months and involves a total contribution of €63,442 from Europol and €119,000 from the beneficiaries.

export1.pdf   2   Document abstract

Send Question

Thank you for your attention!

Judiciary Investigations on the AI Era

Daniel Camara – daniel.camara@gendarmerie.interieur.gouv.fr

Center for Forensic Artificial Intelligence – CFIA/PJGN