

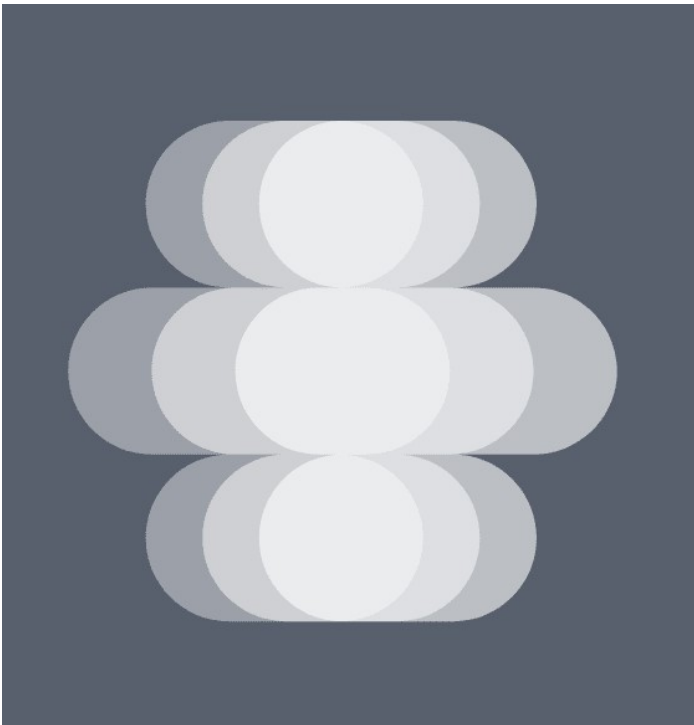


MAPA

MAPA

Multilingual Anonymization
for Public Administrations

MAPA AIMS



“To develop a toolkit for effective and reliable text anonymization in the medical and legal domains [**de-identification of personal data**].

MAPA addresses all EU languages.

The multilingual anonymization toolkit will be based on the Named-Entity Recognition (NER) technique, thus it will be able to focus not be restricted to European names and surnames but also covering non-European and mostly common in all EU countries.”

Final Objectives

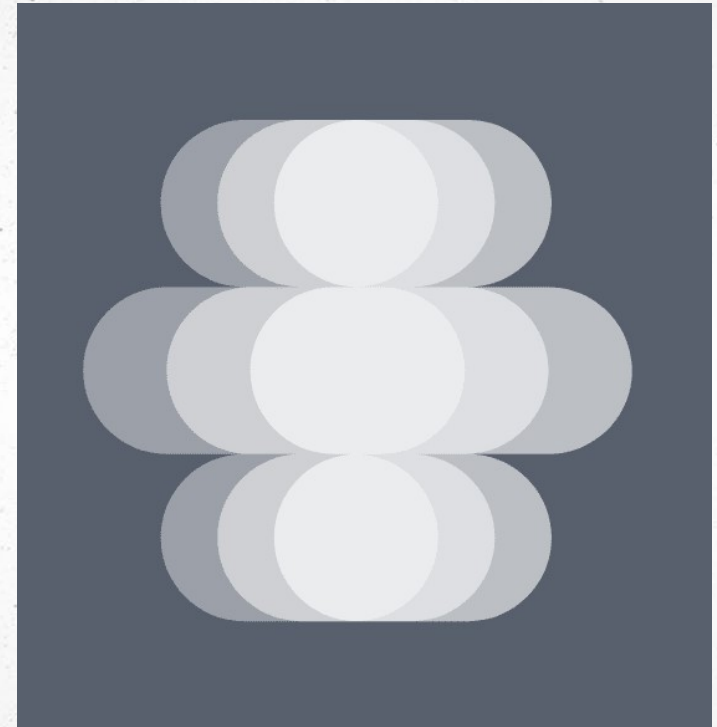
Multilingual Anonymization for Public Administrations



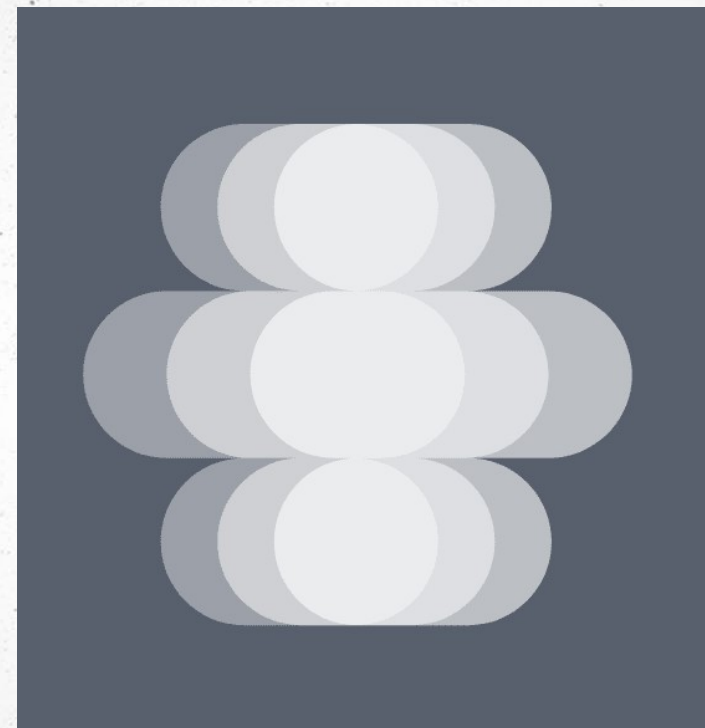
- 1 Deployment Cases**
At least three deployment cases at European PA's: Spain, Latvia, Malta
- 2 Potential Deployments**
France, Poland
- 3 Deliverable**
Docker ready anon [de-identification] kit to ELRC-Share
- 4 Dissemination**
At EU conference gatherings and engaging NAPs
- 5 Best Practices**
Report for the benefit of European PAs, ELRC community, etc.

Difference between CEF and other type of EU projects

- No research activities funded
- 50% in advance, 50% upon successful completion
- Strong focus on deliverables
- Implementation and deliverables **above all!**
- Final report: Milestone compliance



Q&A



ADVP NP VP PP NP NP VP NP NP VP PP VP
When you think of the words "risk" and "investment," what's the specific peril that comes to mind?

NP ADVP
Pencils down.

SBAR NP VP PP NP NP VP NP VP NP NP VP ADVP NP NP NP VP NP NP VP PP NP
If you're like most people, you said it's a holding that goes completely sour -- maybe a bond that defaults or a stock whose value disappears in a bankruptcy proceeding.

NP VP NP ADVP PP NP NP VP NP NP PP NP NP PP NP
"People tend to see risk primarily on that one dimension," says Timothy Kochis, national director of personal financial planning for accountants Deloitte, Haskins & Sells.

ADVP VP NP PP NP NP PP VP NP VP NP PP NP VP PP NP
But therein lies another aspect of investment risk: the hazard of shaping your portfolio to avoid one or more types of risk and being blind-sided by others.

NP VP ADVP NP PP NP NP VP PP NP NP VP PP NP PP NP VP ADVP PP NP
This is clearly not good news to all you people who sleep like babies every night, lulled by visions of your money sitting risk-free in six-month CDs.

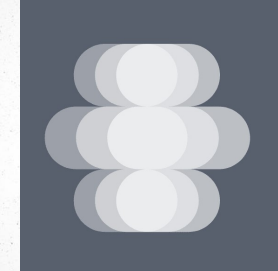
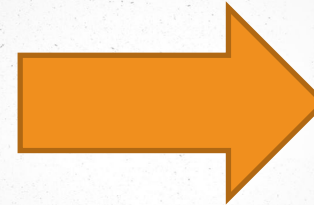
NP VP NP NP NP VP ADVP PP NP PP NP VP ADVP PP NP
Risk wears many disguises, and investments that are low in one type of obvious risk can be distressingly high in other, less obvious kinds.

NP
U.S. Treasury bonds,

PP NP VP NP ADVP NP VP PP VP NP PP NP
for example, are supersafe when it comes to returning money at maturity.

NP PP NP VP PP NP NP VP NP PP NP
But their value as investments can be decimated by inflation, which erodes the purchasing power of bonds' fixed-dollar interest payments.

NP VP ADVP NP PP NP
Risk is also a function of time.

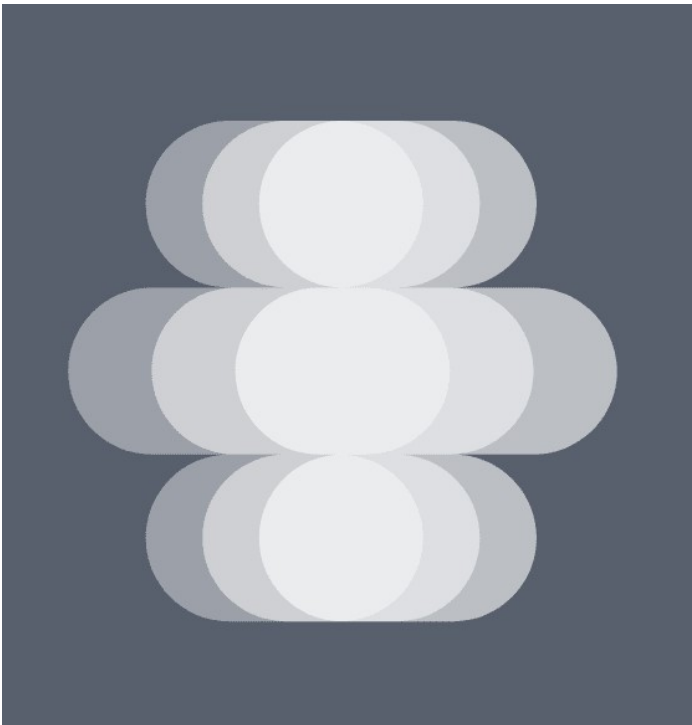


Data Acquisition

(discussion by all partners)

- Data annotation or annotated data?
- Repositories by data partners?
- Same annotation across languages
- Adding Brat (or similar) as an integrated tool for users' refinement?
- Definition of what is to be de-identified for use cases

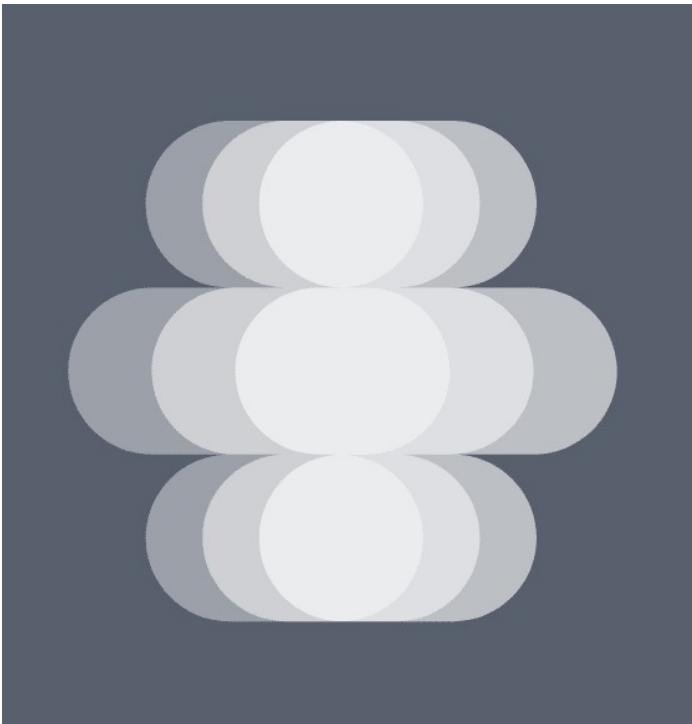
Methodology



There are several approaches and partner input is required

- Traditionally, dictionary based approach
- Of late, neural networks + Named Entity Recognition
- Need for pre-, postprocessing?
- No format handling, only plain text call

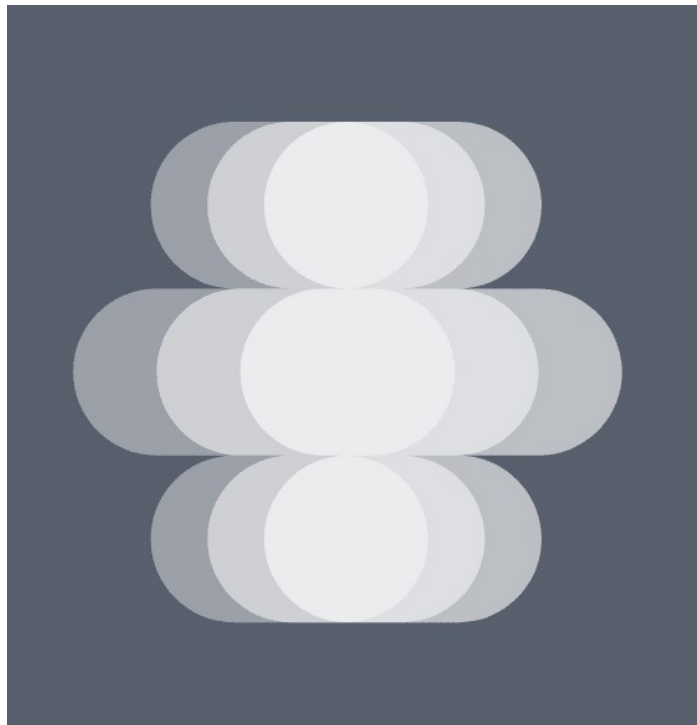
Task Forces



Committee-based approached has proven successful type of governance in other CEF projects.

- Engineering
- Testing / Quality
- Dissemination

Tonight

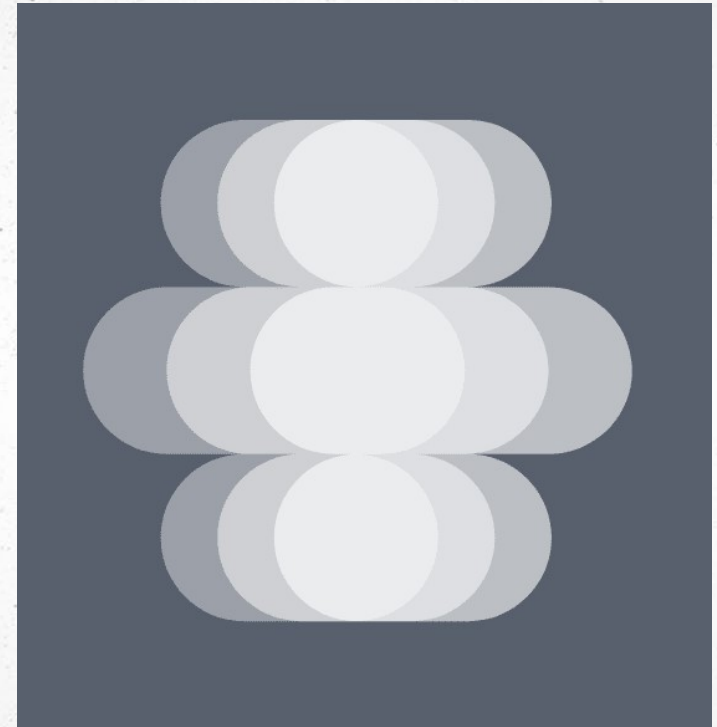


20:30

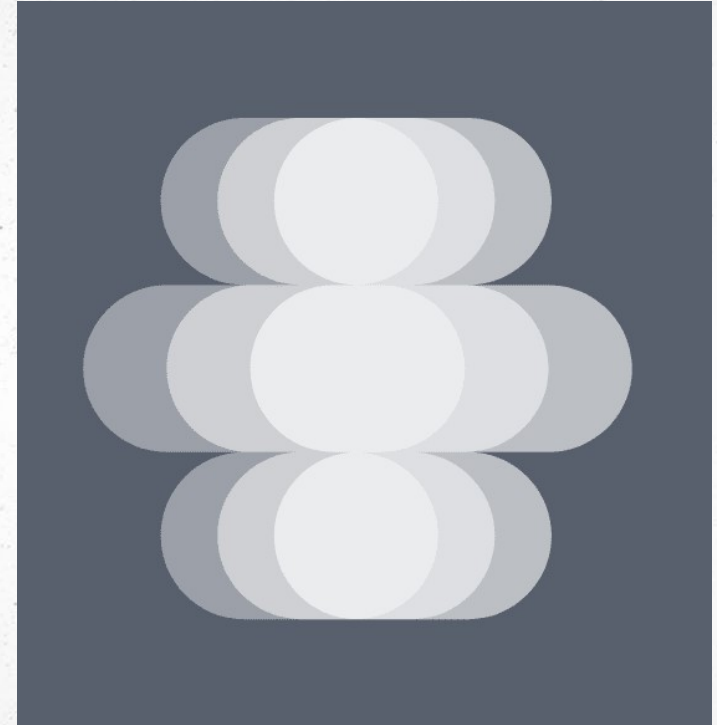
Dinner at Duna Restaurant
(walking distance)

<https://goo.gl/maps/Sd1WdVA8TXYjMD468>

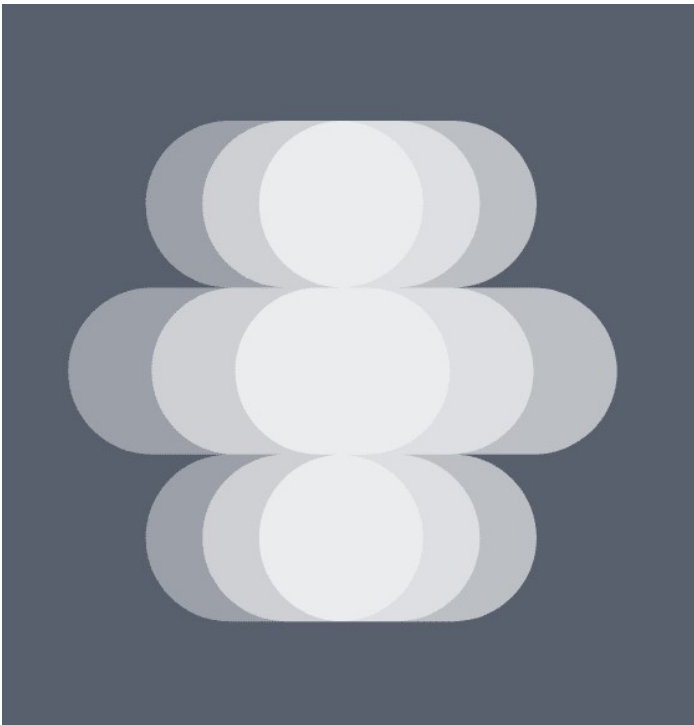
Day 2



Recap



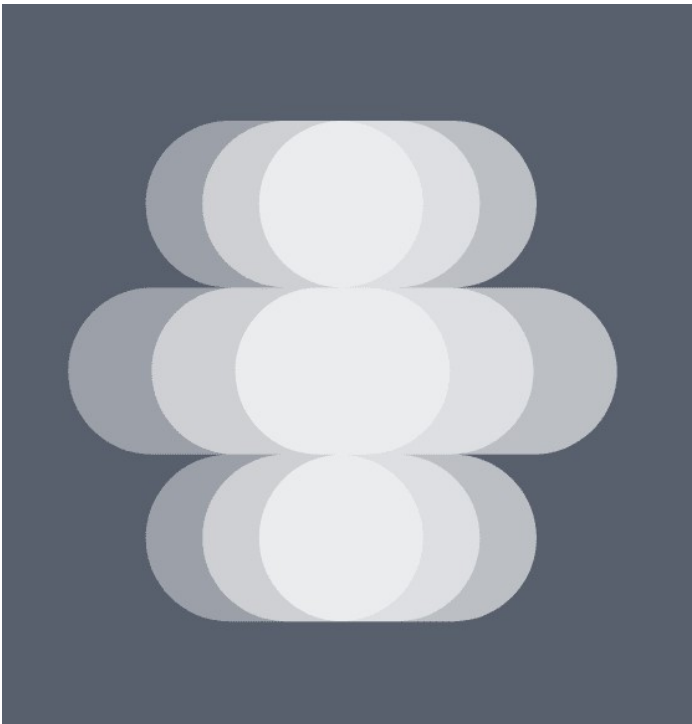
Deployment Cases



- Spain (SEAD)
- Latvia (Tilde)
- Malta (UM)
- France (CNRS and ELRA)

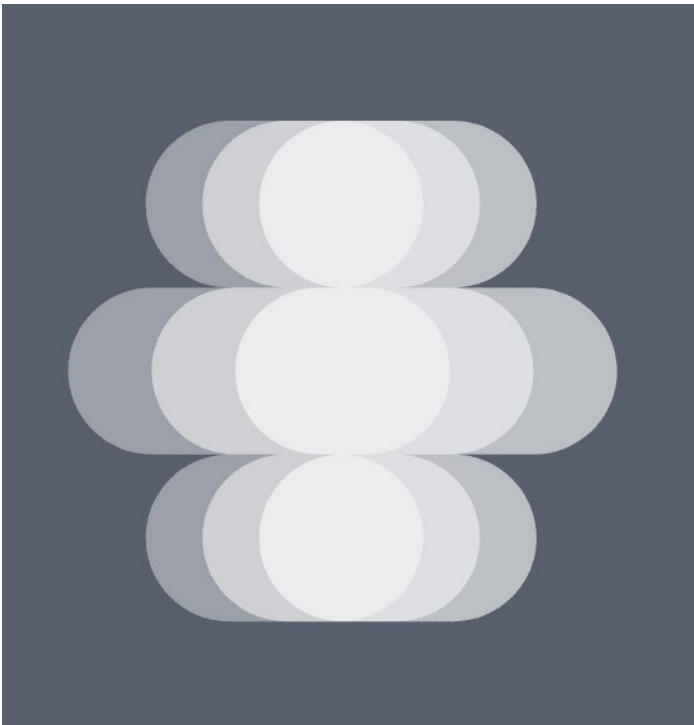
Remember to be social!

Feedback from INEA



- Signatures
- Budget OK (?)
- Approval expected end of January – this means potentially a 1 month or 2 month start delay. However, the consortium may decide to start 01.01.2020 and claim expenses back.

General Admin info

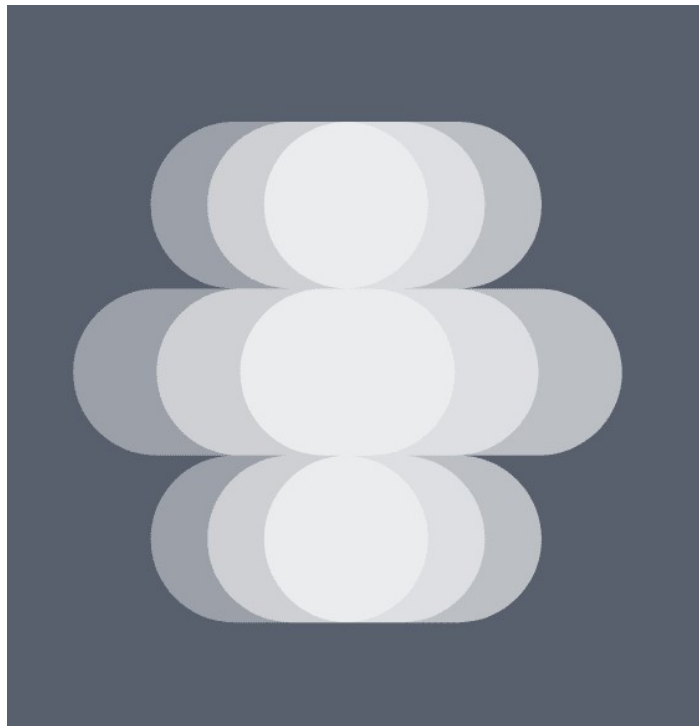


How to use onlyoffice for data storage and reporting

<https://office.pangeanic.com:7080>

Sources of finance

https://europa.eu/youreurope/business/finance-funding/getting-funding/access-finance/search/en/financial-intermediaries?shs_term_node_tid_depth=795



Thank you!