Project - OPEN ACCESS AND DIGITAL ETHICS

CFU 6 – Monica Palmirani

GOAL: To acquire the competences:

1. to formulate an hypothesis of re-use of datasets free from cognitive biases, fair, legal valid, consistent, accurate;
2. to manage a mash-up of different datasets, with different licenses, following the Guidelines for Open Data provided by the EU, with particular regard to the curation of data;
3. to analyse legal and ethics issues, economical and sustainable aspects, technical and metadata foundations;
4. to publish the results using one page web site that includes: original datasets, output datasets, documentation, visualization, metadata using DCAT_AP, license;
5. to visualize and communicate the results to the citizens in human readable manner;
6. to figure out the lifecycle of the dataset and its maintainment over the time.

TASK:

1. **Select datasets for provenience, typology, format, metadata, license. The selection is made for coping with the main goals of the scenario that you intend to achieve. The scenario could be in the cultural heritage, social, economic, politics and public interest domain.**
2. **Analyse the nature of the dataset, their legal basis, the ownership, the quality and accuracy, the compliance with the privacy rules in EU. In case they are not privacy-free, anonymize them using an allowed technology in order to avoid the de-anonymization phenomena.**
3. **Analyse the datasets from ethics point of view in order to avoid discrimination, cognitive bias, prejudice.**
4. **Analyse the datasets from technical point of view and produce Linked Open Datasets as output including DCAT_AP metadata and license RDF.**
5. **Design and develop visualization in order to explit the results and to support the communication.**
6. **Document all the process, inputs, outcomes using one-page web site and github.**

Template of documentation:

1. Introduction
2. Scenario
3. Original dataset and mushup datasets
4. Quality analysis of the datasets
5. Legal analysis (privacy, license, purpose, etc.)
6. Ethics anlysis
7. Technical analysis (formats, metadati, URI, provenance)
8. Sostenibility of the update the datasets over the time
9. Visualization
10. RDF assertion of the metadata

The code should be available as well under an open license for software.