Integration and analysis of heterogeneous big data for precision medicine and suggested treatments for different types of patients.

iASiS: Big Data to Support Precision Medicine and Public Health Policy

Anastasia Krithara
Inst. of Informatics & Telecommunications
NCSR “Demokritos”, Greece
iASiS Basic Facts

• Title: Integration and analysis of heterogeneous big data for precision medicine and suggested treatments for different type of patients

• Topic: H2020-SC1-PM-18-2016 - Big Data supporting Public Health policies

• Contract No.: 727658

• Budget: € 4.3M
Vision and Objectives

iASiS Vision:

Turn clinical, pharmacogenomics, and other Big Data into actionable knowledge for personalized medicine and health policy-making

iASiS Objectives:

• Integrate automated unstructured and structured data analysis, image analysis, and sequence analysis into a Big Data framework

• Use the iASiS framework to support personalized diagnosis and treatment
The iASiS Framework

- iASiS focuses on two use cases:
  - Lung cancer
  - Alzheimer’s disease

- General-purpose drugs are often ineffective

<table>
<thead>
<tr>
<th>Drug Class</th>
<th>Percentage Ineffective</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anti-depressants</td>
<td>38%</td>
</tr>
<tr>
<td>Asthma drugs</td>
<td>40%</td>
</tr>
<tr>
<td>Diabetes drugs</td>
<td>43%</td>
</tr>
<tr>
<td>Arthritis drugs</td>
<td>50%</td>
</tr>
<tr>
<td>Alzheimer’s drugs</td>
<td>70%</td>
</tr>
<tr>
<td>Cancer drugs</td>
<td>75%</td>
</tr>
</tbody>
</table>
The iASiS Framework

- **iASiS analyses:**
  - EHRs (English & Spanish)
  - MRI & PET/CT images
  - Genomic data (e.g. liquid biopsy samples)
  - Related bibliography (e.g. PubMed)
  - Biomedical databases (e.g. DrugBank)
  - Biomedical ontologies (e.g. GO, UMLS)
The iASiS Framework

- Extracted knowledge is fused in the iASiS knowledge graph
  - Unified semantic schema
  - Linked data
  - Machine-processable knowledge

- iASiS end-users can:
  - Perform natural language questions
  - Receive answers along with justifications
  - Identify patterns in patient populations
  - Make informed decisions

- All steps of data management and analytics enforce privacy and access control
Lung Cancer Pilot Data

- EHRs in Spanish
- PET/CT Images
- Genomic Data/Liquid Biopsy Samples

+ Pharmacological knowledge extracted from publicly available datasets
- Biomedical ontologies and taxonomies
  - terminology standardization
  - semantically describing the EHRs
Lung Cancer Use case

Select treatment for long survival

Input: patient data

Patient categorization into survival groups

Output

Treatment Recommendation

online

offline

Analysis

Input: patient data

Output

Treatment Recommendation

Analysis
Alzheimer's Disease Pilot Data

- EHRs in English
- MRI Brain Images
- Genomic Data

- Pharmacological knowledge extracted from publicly available datasets
- Biomedical ontologies and taxonomies
  - terminology standardization
  - semantically describing the EHRs

Logos and symbols: Biobank UK, CRIS Network, St George's University of London, PubMed, PMC, UniProt, DrugBank, Gene Ontology, Disease Ontology, MeSH, Unified Medical Language System.
Alzheimer’s Use case

Which drug is most suitable for a particular patient?

Input: patient data

Positive response estimates for each available drug

Output

Drug Recommendation

online

offline

Analysis
**Current status: iASiS in numbers**

- **Electronic health records:** 7,146 reports, 171,878 clinical notes, 706 patients (LC)

- **Open data:** 266,170 articles (LC & AD), 168,831 concepts, 1,001,180 extracted relations

- **Genomic data:** 20,778 proteins x 98,608 RNAs interaction network

→ **Knowledge graph:** 231,693,984 triples
Beyond Data Analysis

Data Ownership
Explicit Consent

Data Access & Portability
Right to be Forgotten

Intended Use
Data Minimisation

Data Accuracy
Secure Storage
iASiS Partners
Thank you for your attention

http://project-iasis.eu

@Project_IASIS