



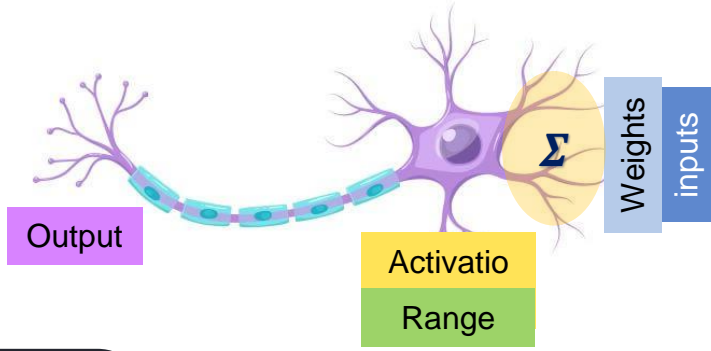
Co-financed by the Connecting Europe
Facility of the European Union



Transforming Healthcare

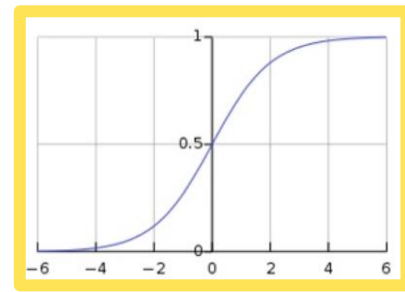
Data Science Process in Healthcare

Bonfitto Giuseppe
bonfitto.giuseppe@hsr.it



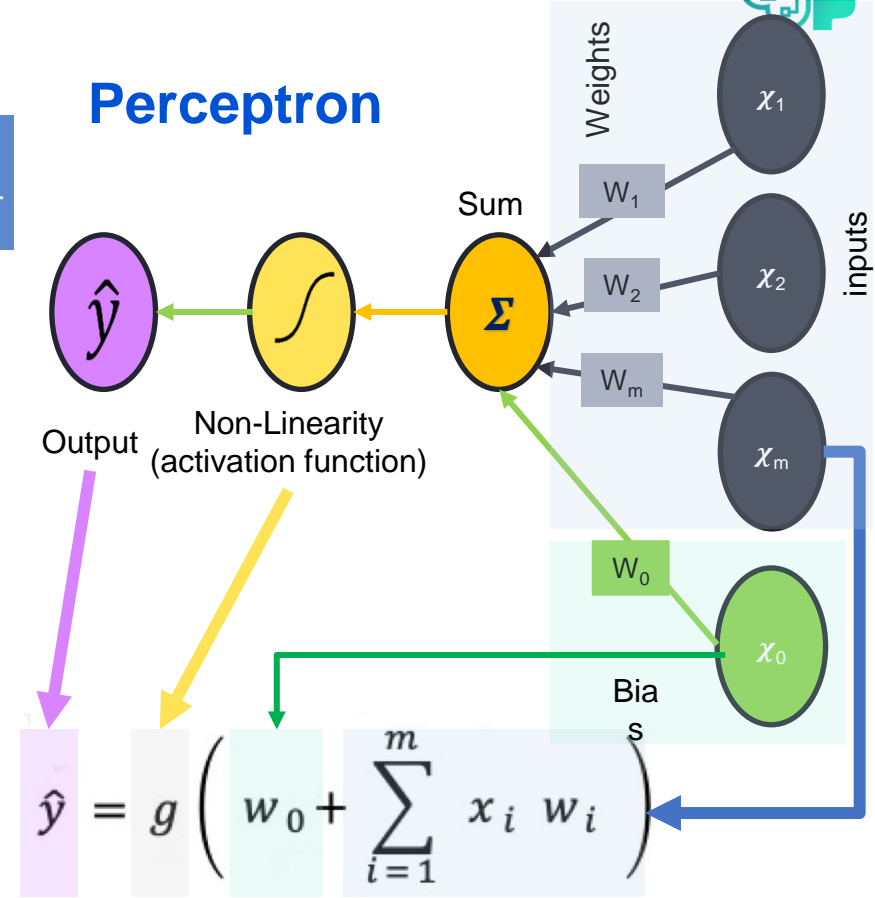
AI Modeling

Neural Network S



Bia

Perceptron

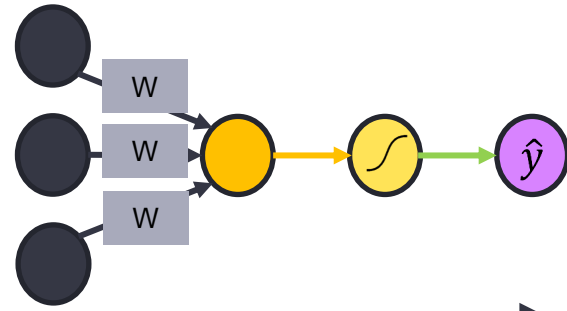


$$\hat{y} = g \left(w_0 + \sum_{i=1}^m x_i w_i \right)$$

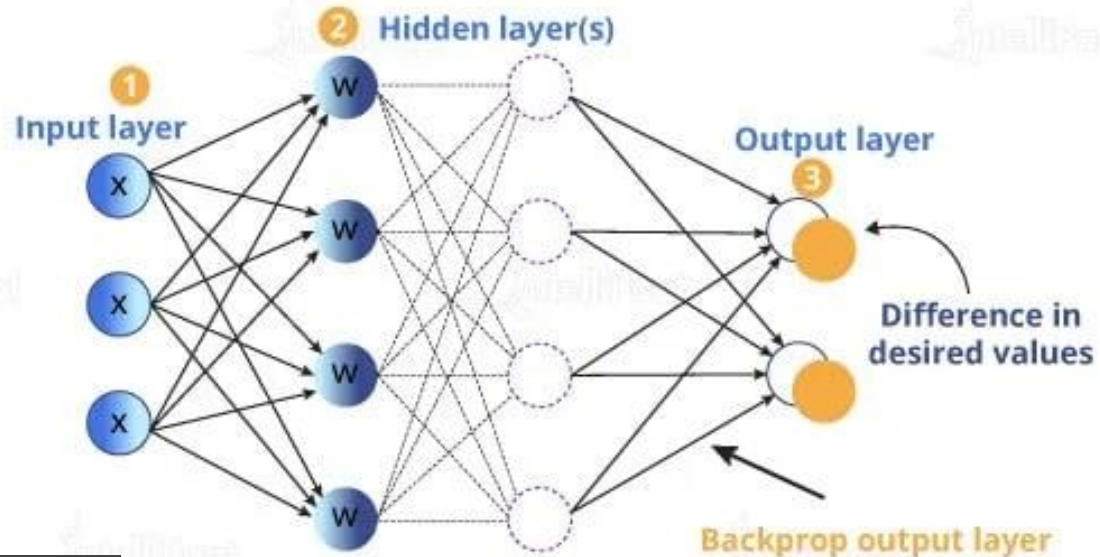
Linear combination of inputs

AI
Modeling

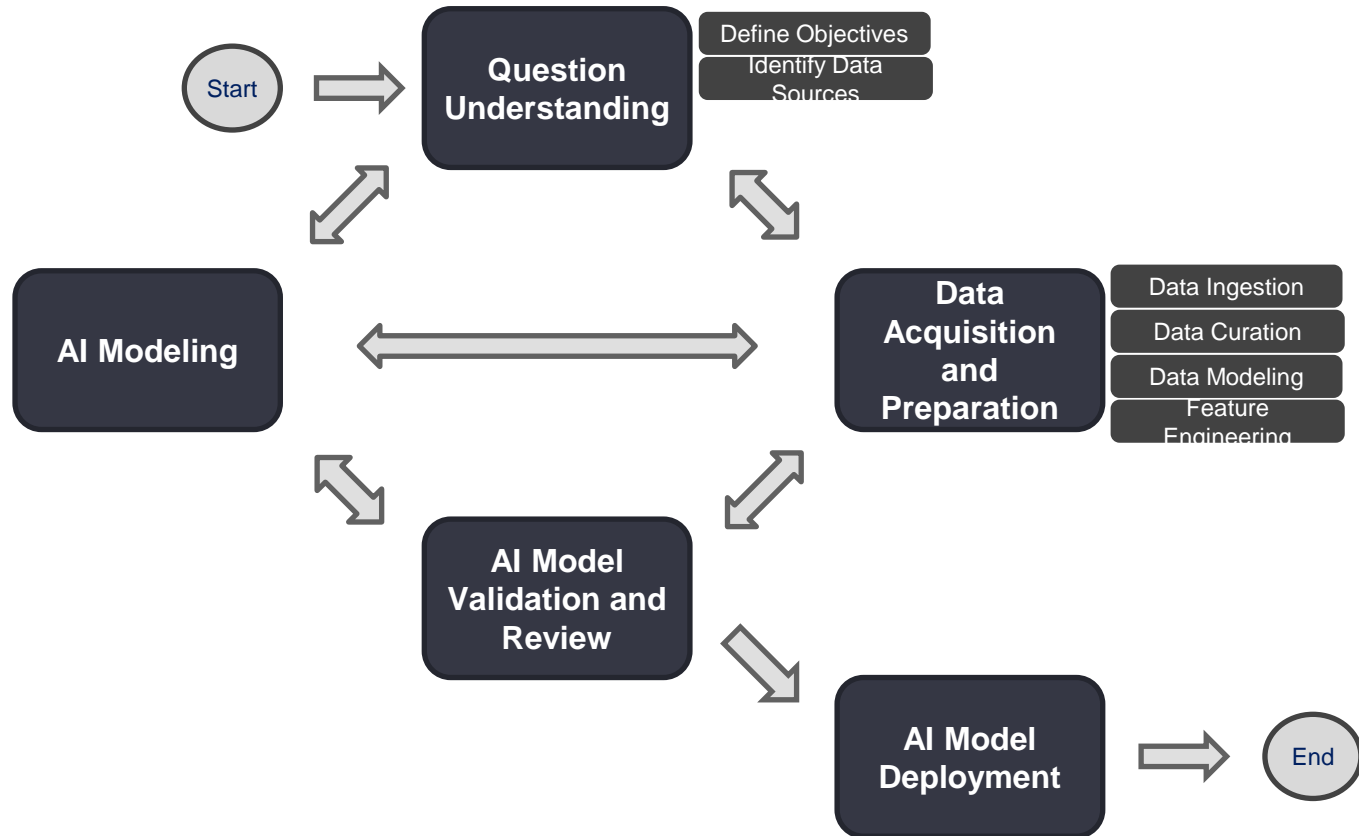
Neural
Networks

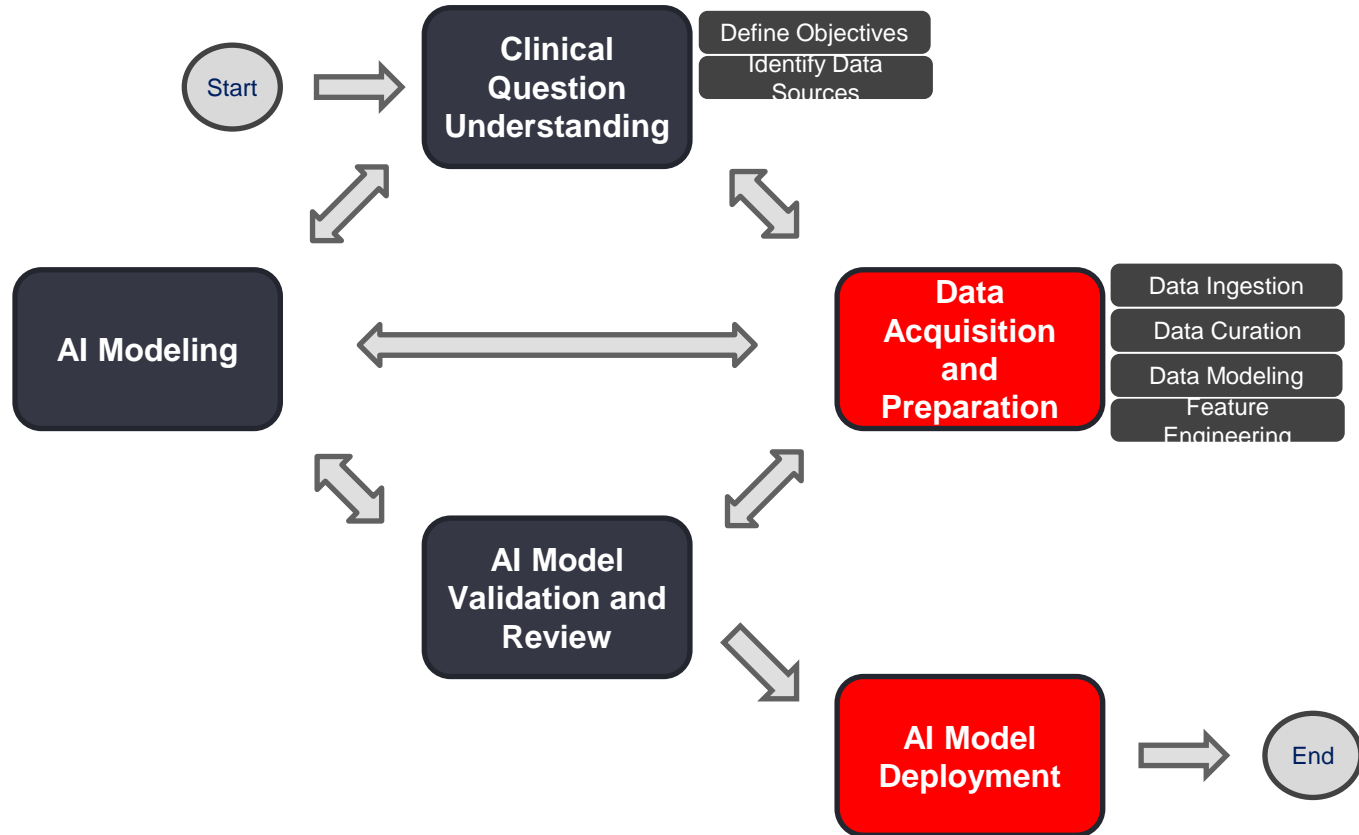


Forward propagation

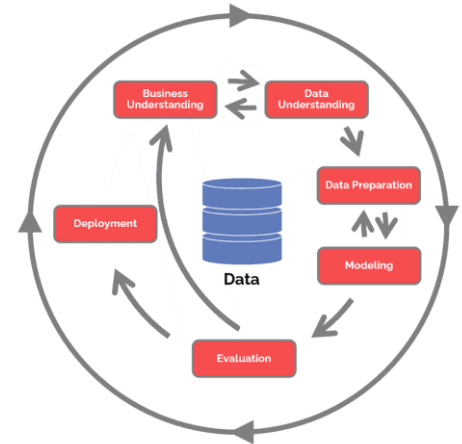
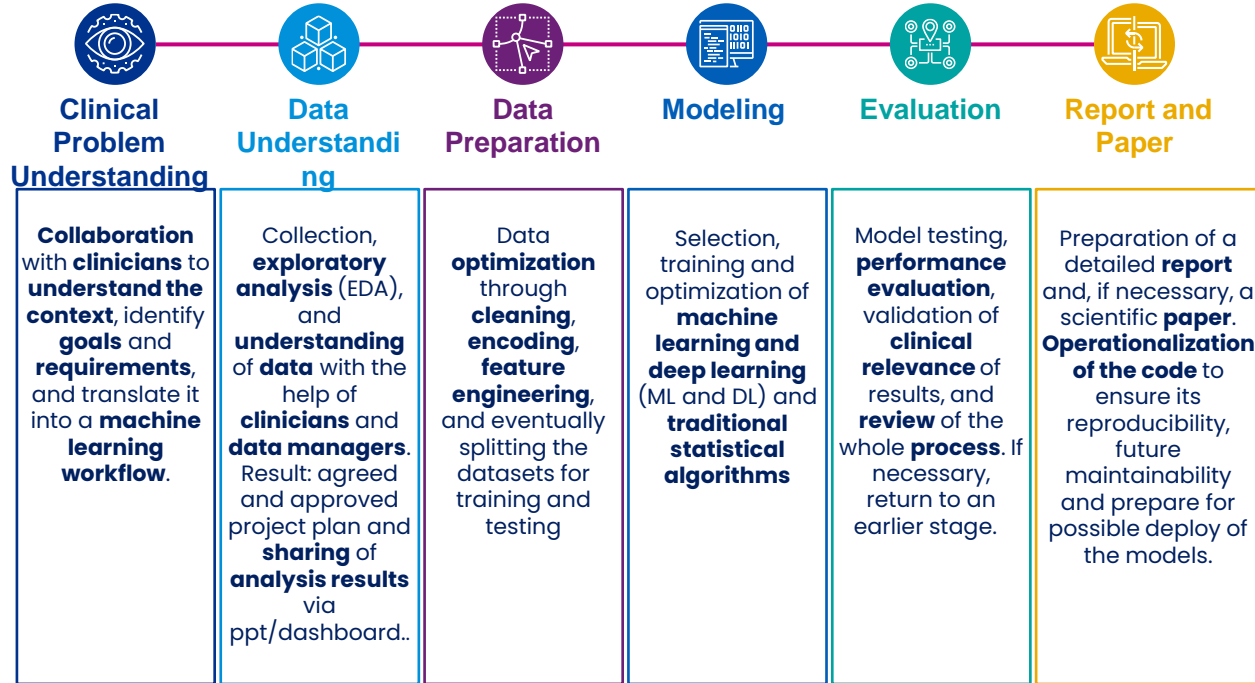


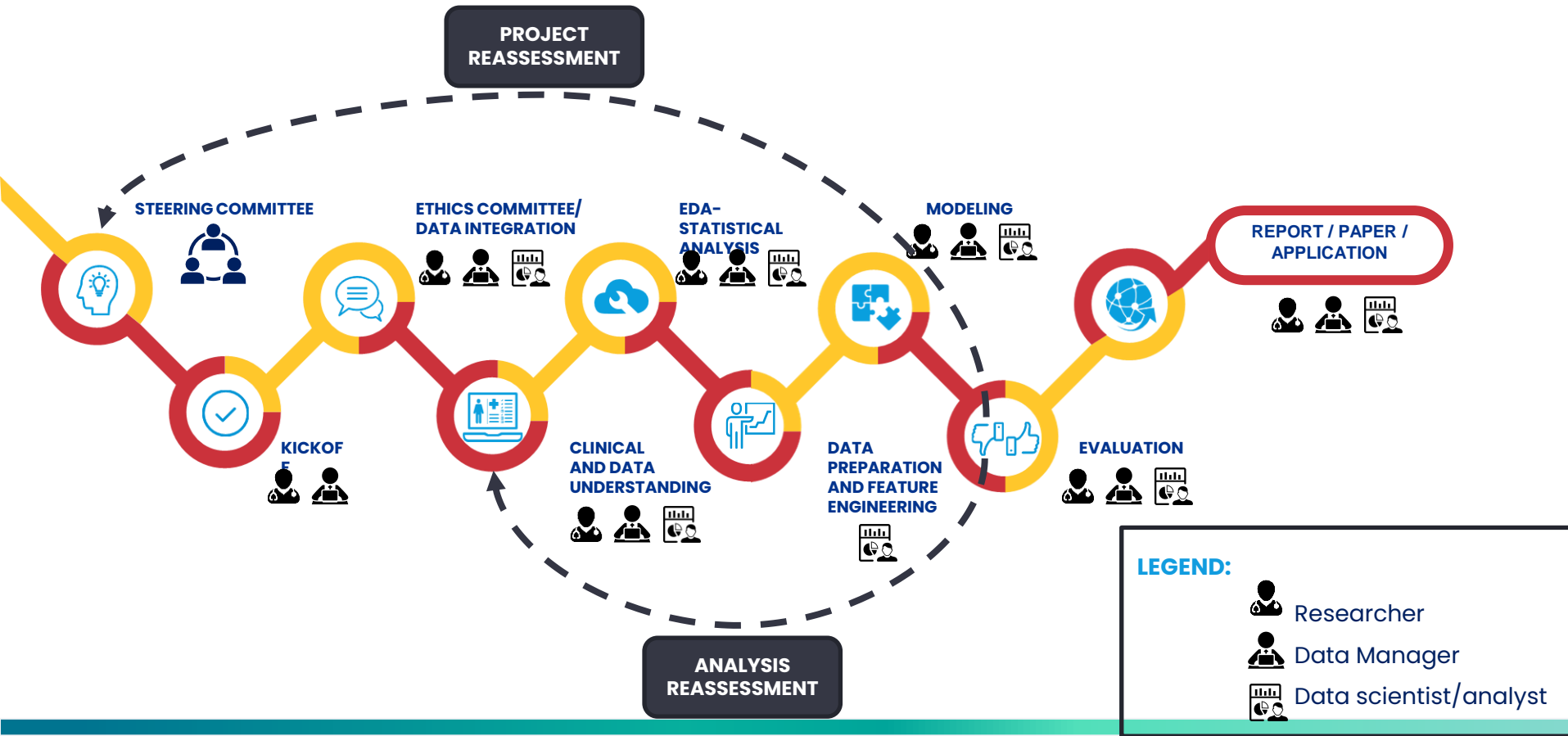
Backpropagation



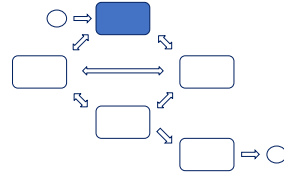


Cross Industry Standard Process for Data Mining (CRISP-DM)





**Clinical
Question
Understanding**



Confirmed cases: 59

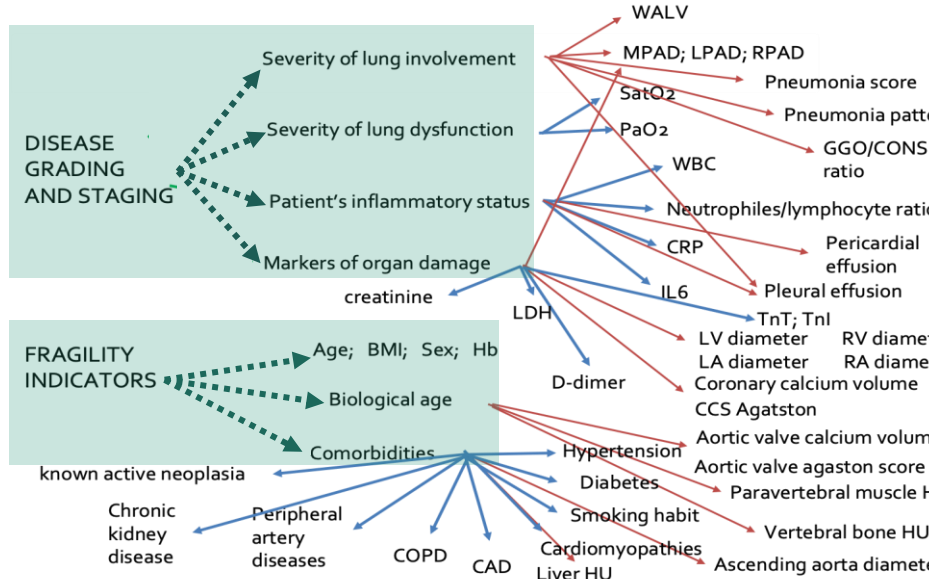
Deaths: 0

January 5, 2020

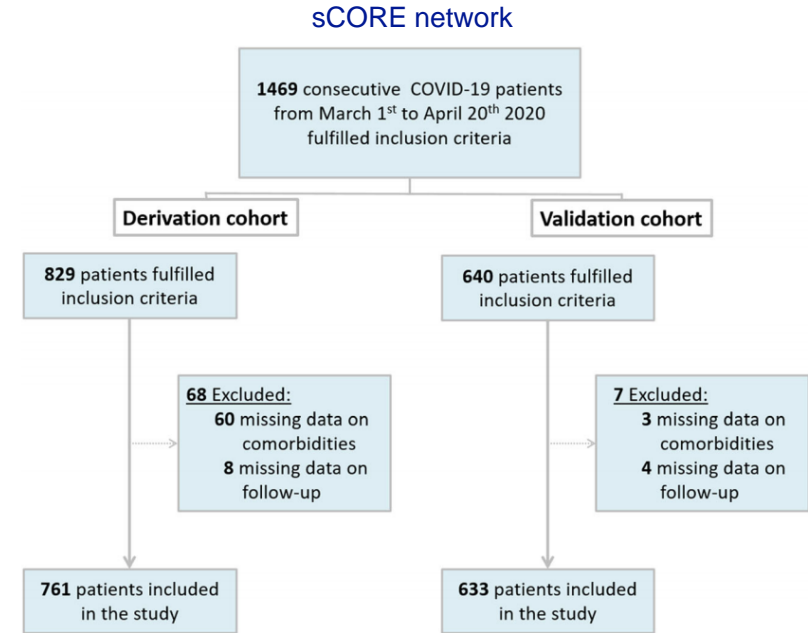
— Death Rate: 0.00%

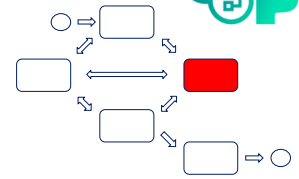
SARS COVID-19
worldwide pandemic





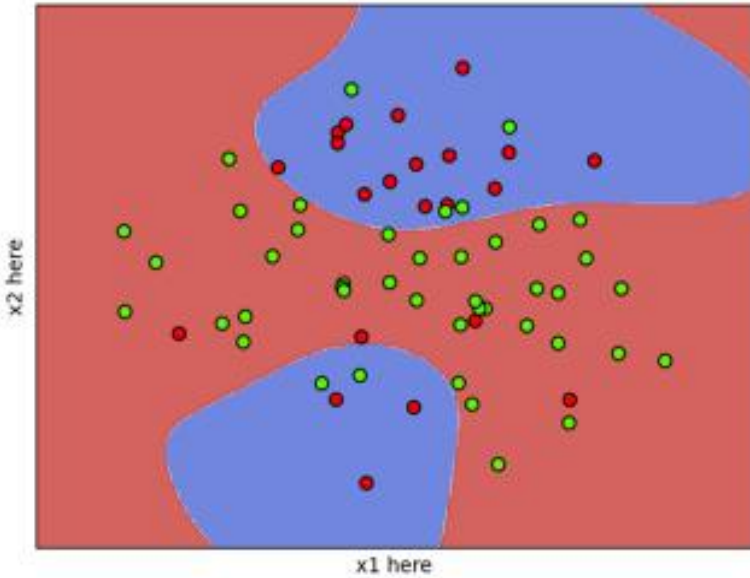
→ Thoracic CT parameters
 → Clinical and laboratory parameters



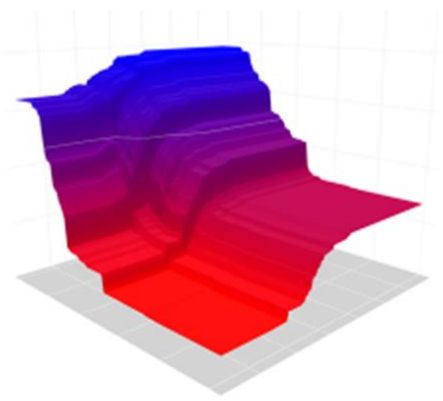


Data Acquisition and Preparation

Decision surface of rbf SVC



SVM model after dimension reduction



ADA Boost 3D after dimension reduction

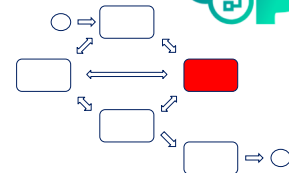
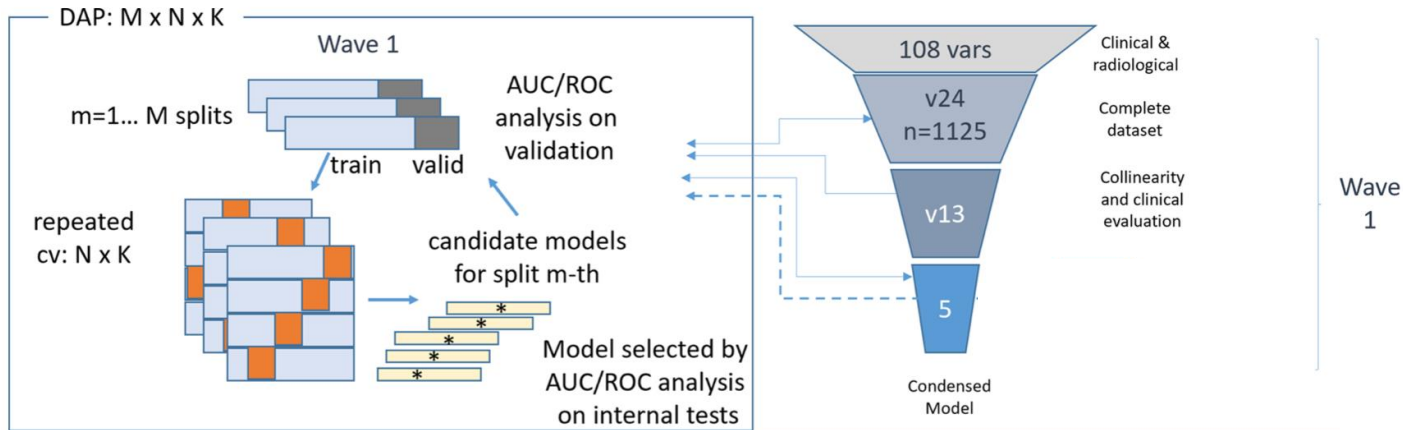
Other analysis:

Most informative content: analysis to find features with most discrimination power

Univariate: analysis to find correlation between single feature and class

Covariate: analysis to find correlation between group of features and class

Misclassification: analysis to find factors that encourage wrong class attribution

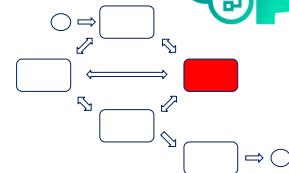
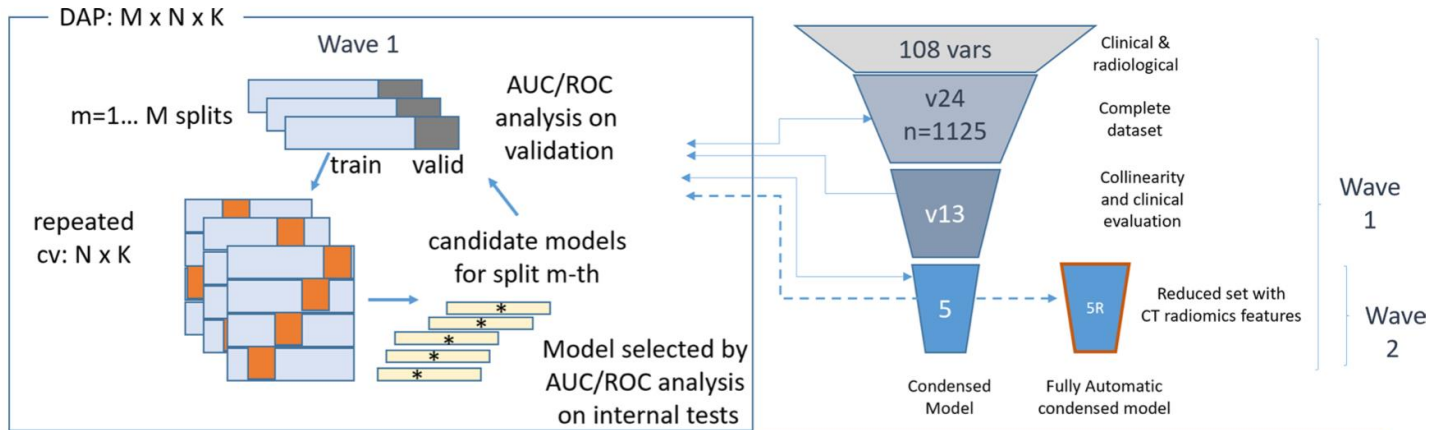


Data Acquisition and Preparation

**Bayesian Logistic Regression:
Classification using manual CT extraction**

| Target outcome | Number of selected features | Result |
|----------------|-----------------------------|---------------------|
| Survival | 24 | AUC = 0.839±0.009 |
| Survival | 13 | AUC=0.840±0.0093 |
| Survival | 5 | AUC = 0.834 ± 0.007 |





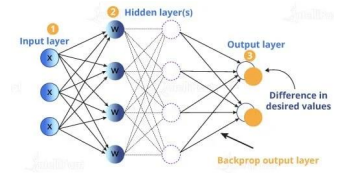
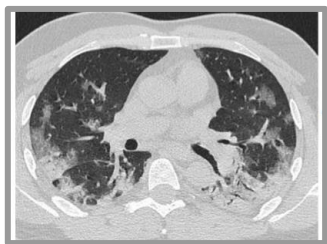
Data Acquisition and Preparation

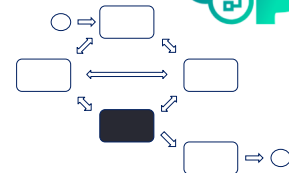
Bayesian Logistic Regression: Classification using **manual** CT extraction

| Target outcome | Number of selected features | Result |
|----------------|-----------------------------|---------------------|
| Survival | 24 | AUC = 0.839±0.009 |
| Survival | 13 | AUC=0.840±0.0093 |
| Survival | 5 | AUC = 0.834 ± 0.007 |

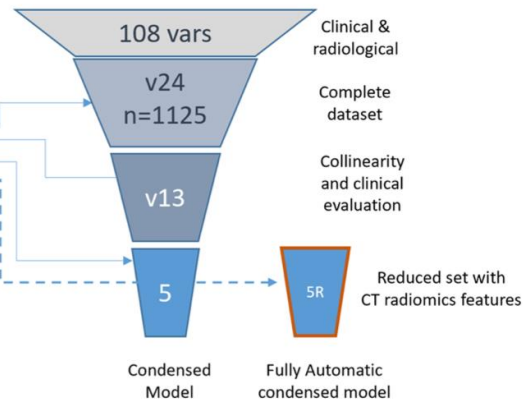
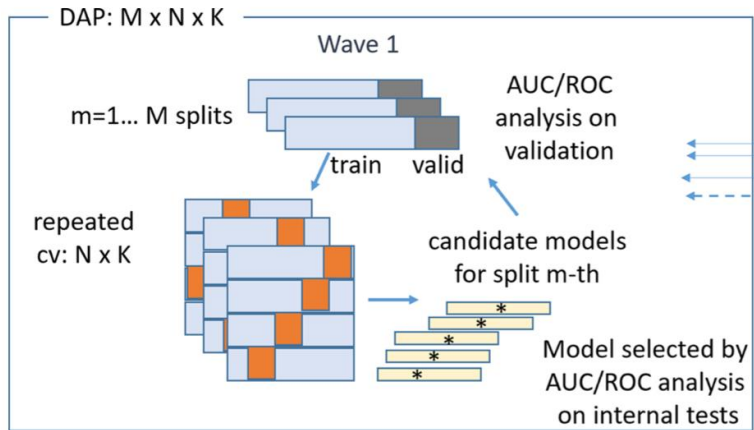
Bayesian Logistic Regression: Classification using **fully automatic** CT extraction

| Target outcome | Number of selected features | Result |
|----------------|-----------------------------|--|
| Survival | 5 | AUC=0.842 (DeLong 95% CI: 0.816-0.867) |





AI Model Validation and Review

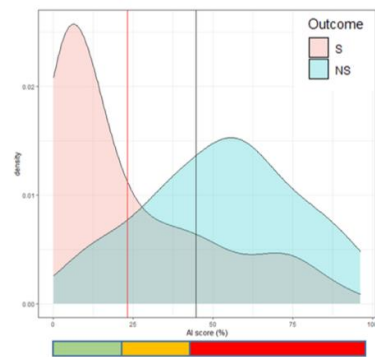
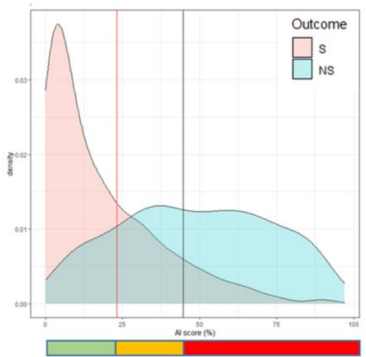
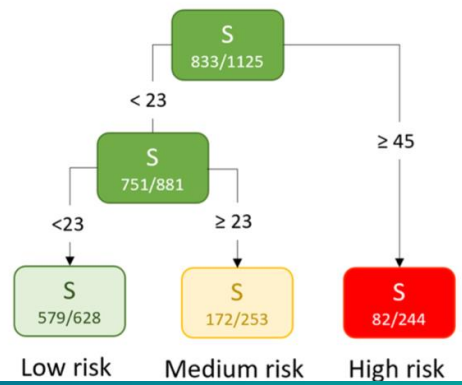


Model selection and validation by DAP on Wave 1: vars24 → vars13 → var5 → 5R

Tree-based binning in 3 risk classes

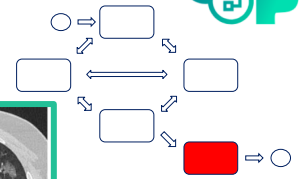
AI-SCoRE Wave1

AI-SCoRE Wave2



explainable Artificial Intelligence in healthcare Management

2020-EU-IA-00g8



AI Model Deployment

New Patient

Age

Sex

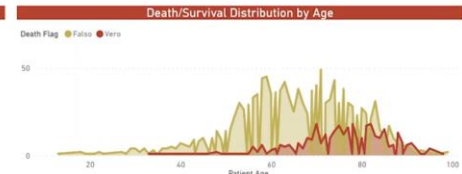
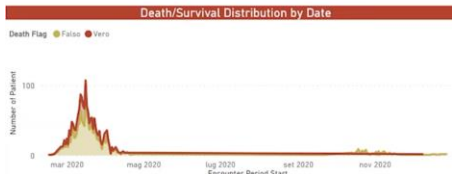
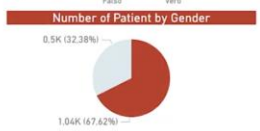
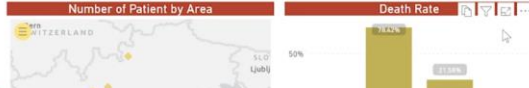
Sat O₂ %



Software interface for patient data entry and AI model deployment. Fields include Patient ID (ALPAT-00094), Last name (ALPAT), First name (ALPAT), Middle name, Gender (Male, Female, Other), and Date of birth (YYYY / MM / CC). It also lists Referring physician and Reading physician details. A file explorer shows a folder named 'AI-SCORE' containing a file 'AI-SCORE'.



| Number of Patient | Number of prediction available | % Death Male over Female |
|-------------------|--------------------------------|--------------------------|
| 1673 | 7 | 67.77% |



Software interface for patient details and observations. The patient is identified as BOL, male, 72 years old (Vuoto). The interface shows Patient Observations and Observation Comparison. The observation description is 'Body mass index (BMI) (Ratio)' with a value of 28.64. The observation comparison shows a scatter plot of 'Right ventricular cavity size (observable entity)' vs 'Body mass index (BMI) (Ratio)'. A table at the bottom shows the encounter period from 2020-03-17 to 2020-03-30 with 14 hospitalization days.



Microsoft Intelligent Data Platform, Azure Health Data Services, Azure Cognitive Services & Azure Machine Learning



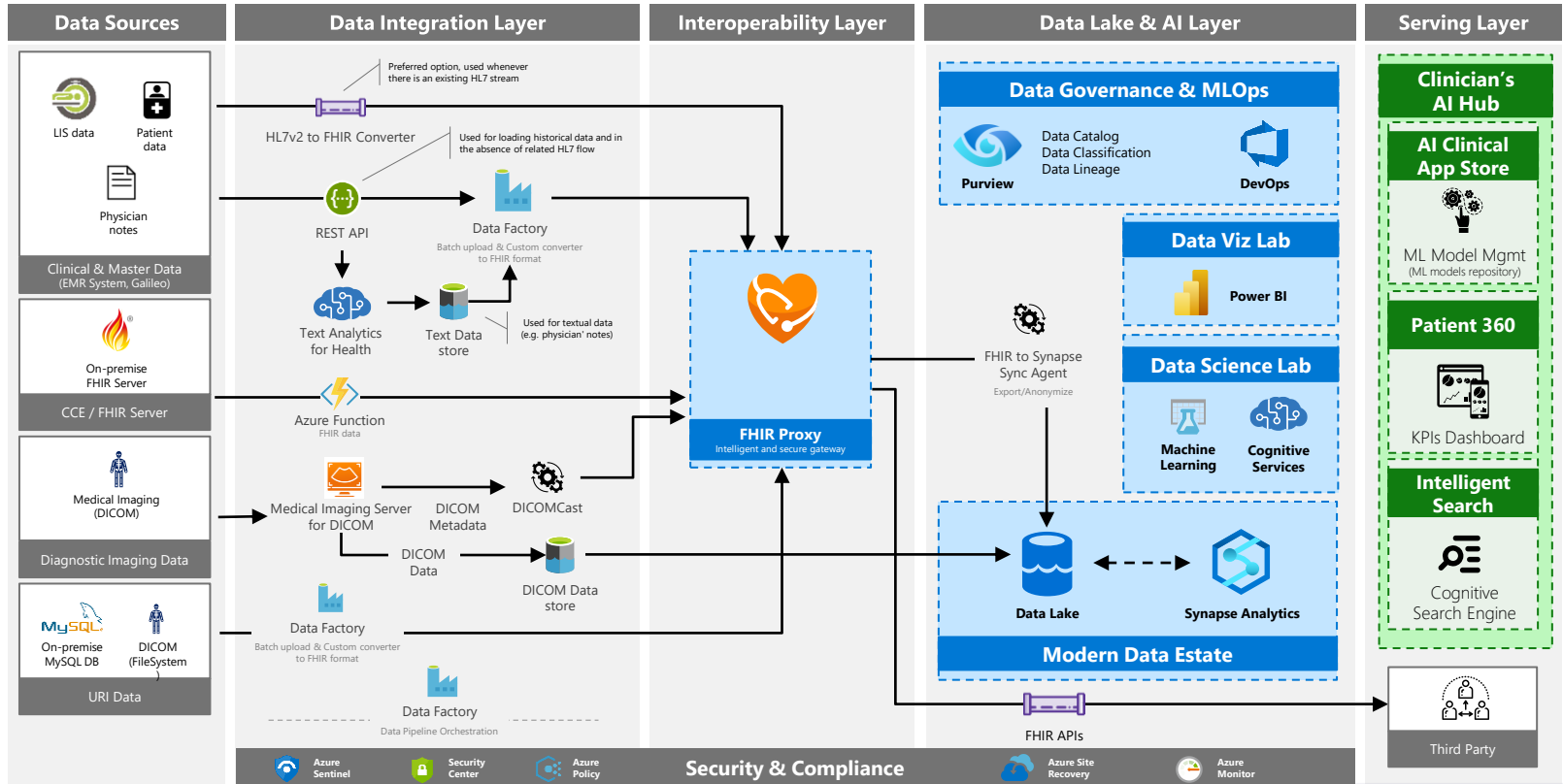
Data Engineer users



Data Scientist users



Researchers





Data Engineer users

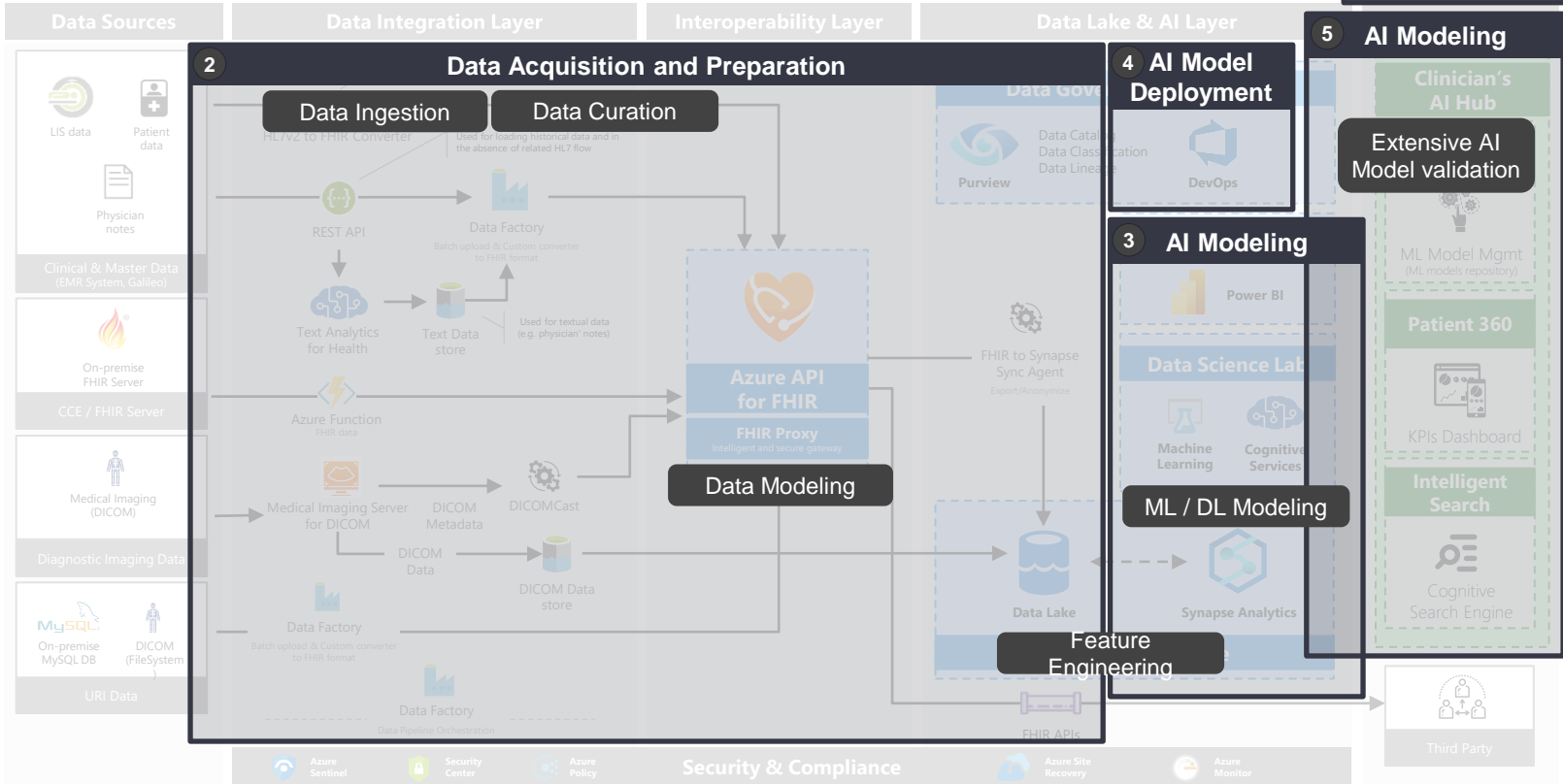


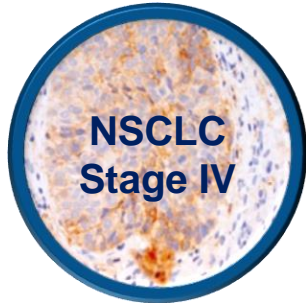
Data Scientist users



Researchers

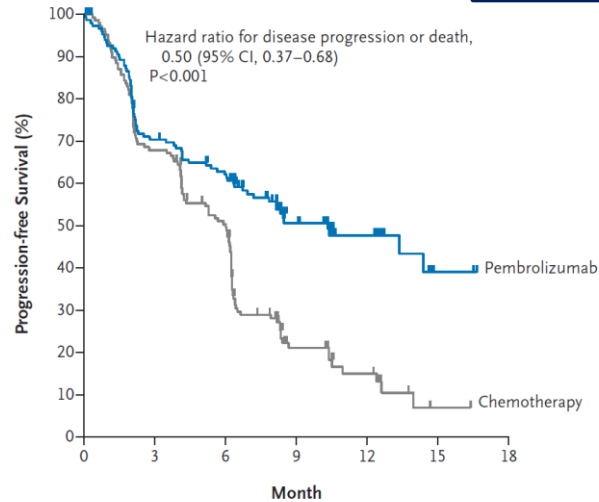
Clinical Question Understanding



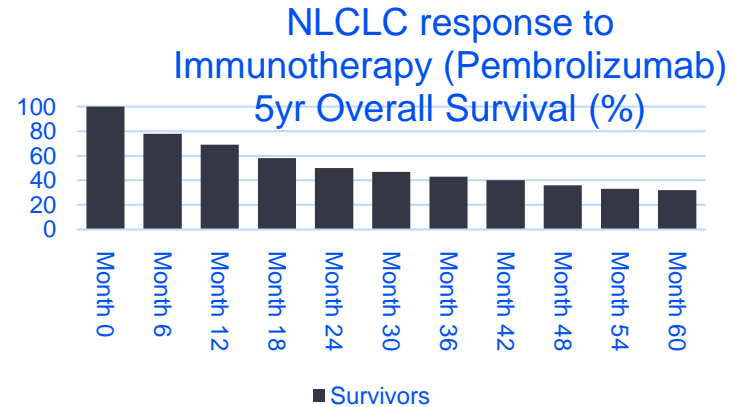
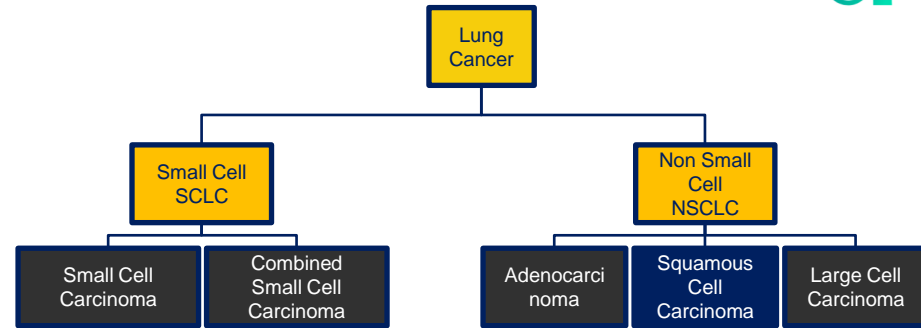


**PD-L1>50%
No EGFR mut**

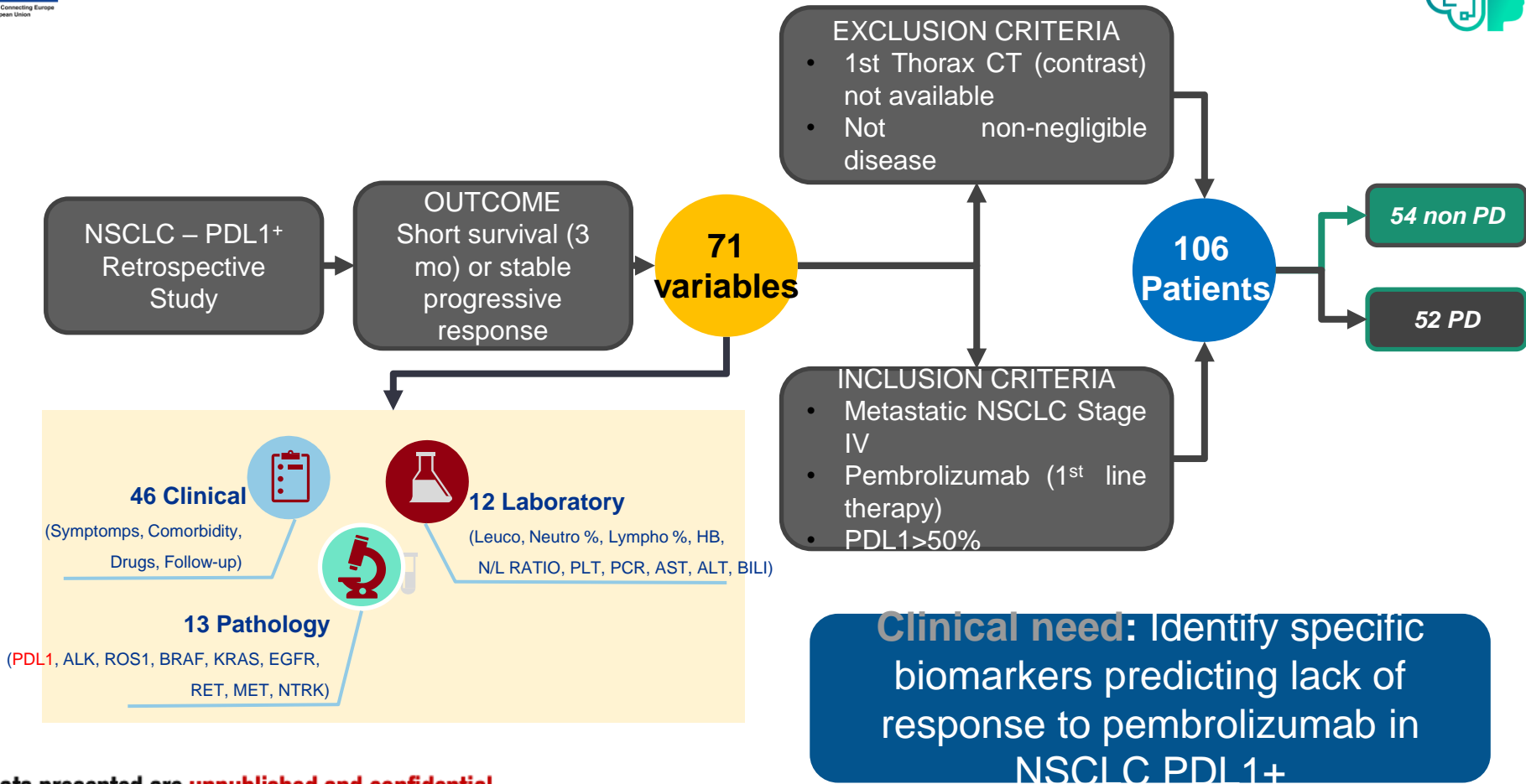
Pembrolizumab (first line)

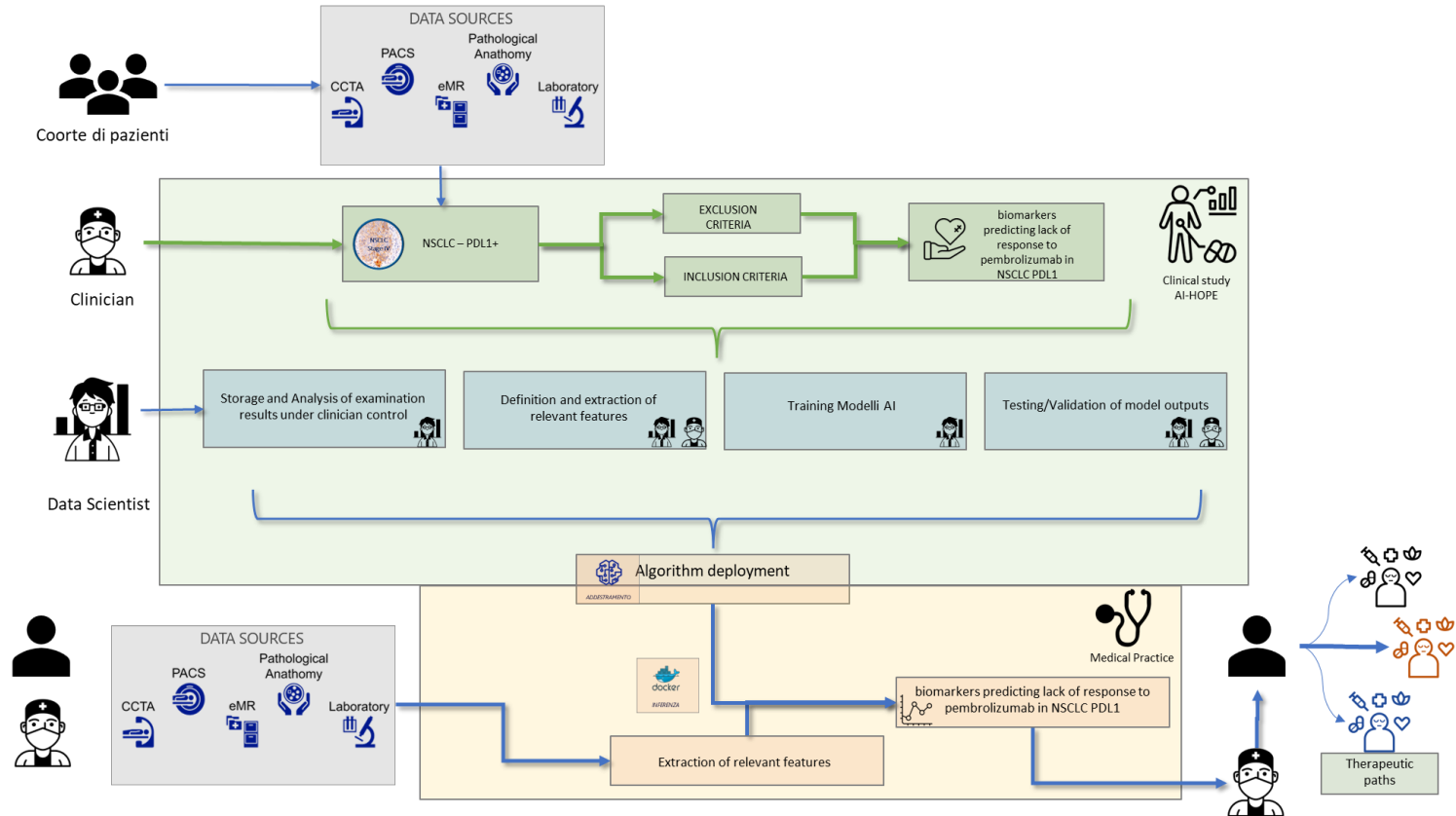


| No. at Risk | 0 | 3 | 6 | 9 | 12 | 15 | 18 |
|---------------|-----|-----|----|----|----|----|----|
| Pembrolizumab | 154 | 104 | 89 | 44 | 22 | 3 | 1 |
| Chemotherapy | 151 | 99 | 70 | 18 | 9 | 1 | 0 |



Data Source: Reck et al New England J Medicine 2016,375:1823



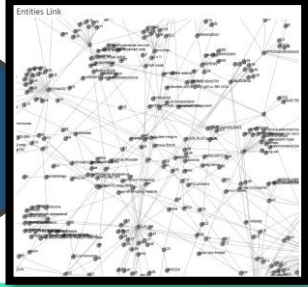
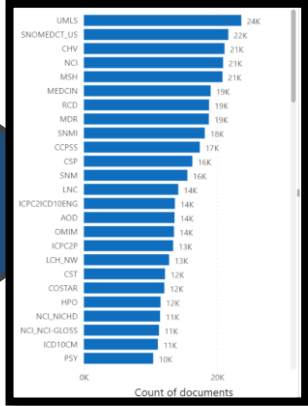
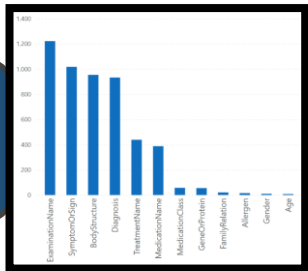
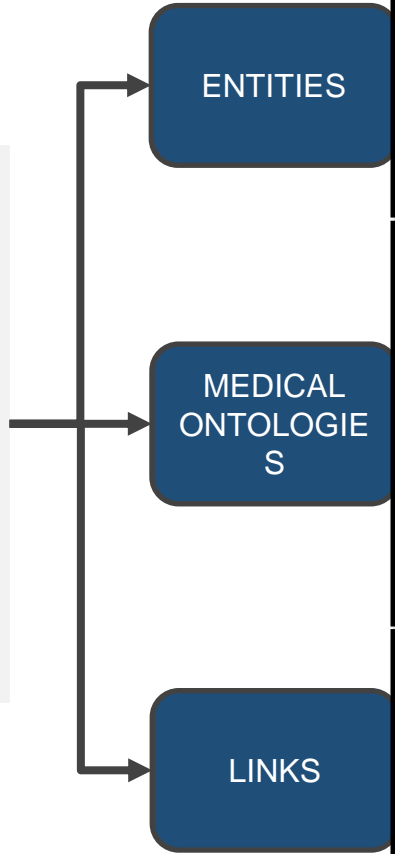


Data Sources

- LIS data
- Pathological data
- Patient data
- Physician notes
- Medical Imaging
- Diagnostic Imaging Data



- detect **words** and **phrases** mentioned in **unstructured text** as **entities** that can be associated with **semantic types** in the **healthcare** and biomedical domain
- connect entities to **medical ontologies** and domain-specific coding systems (example UMLS, SNOMED, LOINC)
- identifies **meaningful links between concepts** mentioned in text





Co-financed by the Connecting Europe Facility of the European Union

explainable Artificial Intelligence in healthcare Ma 2020-EU-IA-00g8

Clinician's AI Hub

Platform Management

Data & Model Management

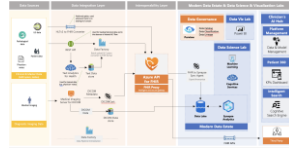
Patient 360

KPIs Dashboard

Intelligent Search

Cognitive Search Engine

Third Party



Clinicians AI Hub

Francesco Rossi
Medico Primario - Oncologia clinica

MSR - Clinical Research Office

My contribution T4 Sort

PRIME 26 Nov 2022

Integrazione Opedare - Testando uno studio diagnostico prognostico multicentrico farmaceutico all'arret.

2 Principal Investigators

MYOCAR 1 Nov 2022

Role of Endomyocardial Biopsy and Autologous based Treatment in Patients With Inflammatory Heart Disease in Antiphospho and ...

3 Principal Investigators

ACRONYM-1 28 Oct 2022

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Con quonque at massa in accumsan. Cras at congue eros, ut dictum lectus. Vestibulum venenula.

2 Principal Investigators

ACRONYM-2 11 Sep 2022

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Con quonque at massa in accumsan. Cras at congue eros, ut dictum lectus. Vestibulum venenula.

Mara De Luca

ACRONYM-3 8 Feb 2022

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Con quonque at massa in accumsan. Cras at congue eros, ut dictum lectus. Vestibulum venenula.

Mara De Luca

Next meetings

November - December 2022

| M | T | W | T | F | S | S |
|----|----|----|---|---|---|---|
| 28 | 29 | 30 | 1 | 2 | 3 | 4 |

Index - Tue, Nov 29

- 1:30pm MYOCAR - Alignment meeting Luigi Rossi
- 4pm PRIME - Data analysis
- 5:30pm MYOCAR - Alignment meeting ...

Your tasks

MYOCAR

- MYOCAR - Invitations update 06/12
- MYOCAR - Cohort update 06/12
- MYOCAR - Data analysis 06/12

Ongoing Studies

Clinicians AI Hub

Study Creation

1 Study ID card 2 Team composition 3 Confirmation

Discard Next

Main descriptive data

Acronym

Protocol full title

Study description

Study category

Retrospective Prospective

Classification will be available in further releases

Centers involved

Monocentric study (IRCCS Ospedale San Raffaele) Multicentric study

New Study

Clinicians AI Hub

Study Creation

1 Study ID card 2 Team composition 3 Confirmation

Discard Next

Team composition

All team members will receive an invitation by their own hospital San Raffaele email account. You'll be able to send again the invitation email later. ✕

| Team members | Study role | Permissions |
|---|---|---|
| <div style="display: flex; align-items: center;"> <div style="margin-left: 5px;"> <p>Francesco Rossi Cardiology</p> </div> </div> | <input type="text" value="Co-investigator"/> | <input type="text" value="Administrator"/> |
| <div style="display: flex; align-items: center;"> <div style="margin-left: 5px;"> <p>Alexandra Robertinelli Mastrogiacomio Oncology</p> </div> </div> | <input type="text" value="Principal investigator"/> | <input type="text" value="Editor"/> |
| <div style="display: flex; align-items: center;"> <div style="margin-left: 5px;"> <p>Alberto De Micheli Data Analyst</p> </div> </div> | <input type="text" value="Collaborator"/> | <input type="text" value="Viewer"/> |
| <div style="display: flex; align-items: center;"> <div style="margin-left: 5px;"> <p>Simone Barbieri Data Scientist - Lorem ipsum ceaser</p> </div> </div> | <input type="text" value="Select a study role"/> | <input type="text" value="Select a platform role"/> |
| <div style="display: flex; align-items: center;"> <div style="margin-left: 5px;"> <p>Monica Gandelli Cardiology</p> </div> </div> | <input type="text" value="Co-investigator"/> | <input type="text" value="Collaborator"/> |
| <div style="display: flex; align-items: center;"> <div style="margin-left: 5px;"> <p>Alberto Conte Psychology</p> </div> </div> | <input type="text" value="Collaborator"/> | <input type="text" value="Collaborator"/> |
| <div style="display: flex; align-items: center;"> <div style="margin-left: 5px;"> <p>Gianantonio Conviolando IT Core</p> </div> </div> | <input type="text" value="Principal Investigator"/> | <input type="text" value="Select a platform role"/> |
| <div style="display: flex; align-items: center;"> <div style="margin-left: 5px;"> <p>Martina Casquariciale Neurology</p> </div> </div> | <input type="text" value="Select a study role"/> | <input type="text" value="Select a platform role"/> |

Team



Co-financed by the Connecting Europe Facility of the European Union

explainable Artificial Intelligence in health 2020-EU-IA-0098

Clinician's AI Hub

Platform Management

Data & Model Management

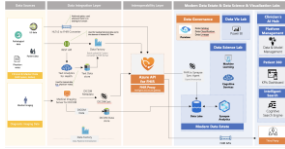
Patient 360

KPIs Dashboard

Intelligent Search

Cognitive Search Engine

Third Party



Clinicians AI Hub
PRIME PRIME0123

Overview
Data set
Data analysis
AI Models

Study details

Study ID
Integratioe Ospedale "Territorio uno studio osservazionale prospettico multicentrico lorem ipsum dolor sit amet nunc nuncquam rosae rosa tenemus.

Main dates
28-Oct-2022 (Study created)

Centers involved
UNICS Ospedale San Raffaele
ASST Grande Ospedale Metropolitano Spigarelle
+ 2 more

Team

Principal Investigator

- Prof. Arturo Domenicchielli (Oncology)
- Prof. Ramondina Bertolazzi Capra (Oncology)

Collaborators

- Francesco Rossi (Medico Primario - Oncologia clinica)
- Mona Kane (Software Engineer)
- Simone Barbieri (Data Analyst)
- Stefano Ranieri (Screening - Dolor Amet ut)
- Riccardo Fagoli (Pharm)

Study cohort

Study Population See inclusion and exclusion criteria

432 patients 28-92 age range

88 Female
51.4% (222 patients)

Male
48.6% (210 patients)

767 x 0

28-Nov-2022 (Last updated)

File browser

25 files - Latest updated See all files

- File_name Lorem Ipsum 09 (File size)
- File_name Lorem Ipsum 09 (File size)

Your tasks

- MYOCAR - Document preparation Doing (due Dec '22)
- MYOCAR - 517 assignment Doing (due Dec '22)
- MYOCAR - Cohort selection To do (due Dec '22)

Clinicians AI Hub

PRIME PRIME0123
Overview
Data set
Data analysis
AI Models

[Explore Data Catalog](#)

Cohort composition

Patients 123 Select a patient row to see a detailed report: [View Patient Details](#)

| Patient Code | Files | Exams | Dicom | Files2 | Files3 |
|--------------|-------------|-------------|--------------|--------|--------|
| PRIME-001 | 135 | 12 | | | |
| PRIME-002 | 14 | | | | |
| PRIME-003 | 279 | 20 | | | |
| PRIME-004 | 66 | 10 | | | |
| PRIME-005 | 83 | 36 | | | |
| PRIME-006 | 94 | 8 | | | |
| PRIME-007 | 124 | 34 | | | |
| PRIME-008 | 1 | 1 | | | |
| PRIME-009 | 13 | 5 | | | |
| PRIME-010 | 2 | 2 | 273 | | |
| PRIME-011 | 56 | 22 | | | |
| PRIME-012 | 76 | 24 | | | |
| Total | 7551 | 2024 | 20767 | | |

Medical ontologies

Data Source UMLS **Category** Diagnosis

Patients by id and text

| | | |
|--|---|---|
| C0012634 Disease | <div style="background-color: #0070C0; height: 15px; width: 100%;"></div> | <div style="background-color: #0070C0; height: 15px; width: 100%;"></div> |
| C0006826 Malignant Neoplasms | <div style="background-color: #0070C0; height: 15px; width: 95%;"></div> | <div style="background-color: #0070C0; height: 15px; width: 95%;"></div> |
| C0238339 Hereditary pancreatitis | <div style="background-color: #0070C0; height: 15px; width: 95%;"></div> | <div style="background-color: #0070C0; height: 15px; width: 95%;"></div> |
| C0497156 Lymphadenopathy | <div style="background-color: #0070C0; height: 15px; width: 95%;"></div> | <div style="background-color: #0070C0; height: 15px; width: 95%;"></div> |
| C0024115 Lung diseases | <div style="background-color: #0070C0; height: 15px; width: 95%;"></div> | <div style="background-color: #0070C0; height: 15px; width: 95%;"></div> |
| C0152013 Adenocarcinoma of lung (di... | <div style="background-color: #0070C0; height: 15px; width: 95%;"></div> | <div style="background-color: #0070C0; height: 15px; width: 95%;"></div> |
| C0011849 Diabetes Mellitus | <div style="background-color: #0070C0; height: 15px; width: 95%;"></div> | <div style="background-color: #0070C0; height: 15px; width: 95%;"></div> |
| C0001418 Adenocarcinoma | <div style="background-color: #0070C0; height: 15px; width: 95%;"></div> | <div style="background-color: #0070C0; height: 15px; width: 95%;"></div> |
| C0024121 Lung Neoplasms | <div style="background-color: #0070C0; height: 15px; width: 95%;"></div> | <div style="background-color: #0070C0; height: 15px; width: 95%;"></div> |
| C0004093 Asthenia | <div style="background-color: #0070C0; height: 15px; width: 95%;"></div> | <div style="background-color: #0070C0; height: 15px; width: 95%;"></div> |
| C0020538 Hypertensive disease | <div style="background-color: #0070C0; height: 15px; width: 95%;"></div> | <div style="background-color: #0070C0; height: 15px; width: 95%;"></div> |

| Diagnosis | Patients |
|--------------|------------|
| disease | 114 |
| HP | 97 |
| cancer | 94 |
| lung disease | 70 |
| Total | 122 |

| Group | Patients |
|-------------------------------------|------------|
| Acquired Abnormality | 12 |
| Anatomical Abnormality | 8 |
| Bacterium | 23 |
| Biomedical Occupation or Discipline | 1 |
| Total | 122 |

Study Overview