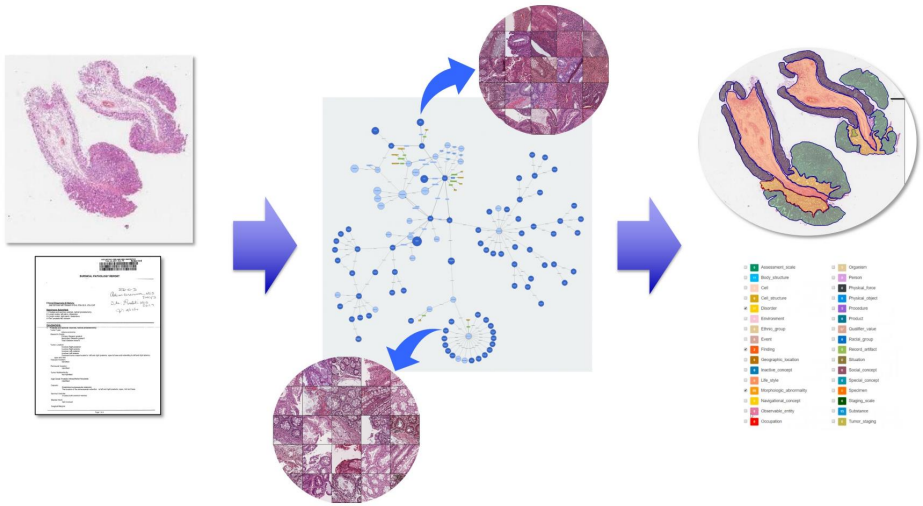
ExaMode

Extreme-scale Analytics via Multimodal Ontology Discovery & Enhancement



Project supported by European Union
Horizon 2020 grant agreement 825292

ExaMode

Project overview

Project supported by European Union
Horizon 2020 grant agreement 825292

ExaMode *(2019 – 2023)*

- Call
  - European Union Horizon 2020, ICT-12-2018-2020
- Topic
  - Big data technologies and extreme-scale analytics
- Consortium
  - 7 partners from 5 countries

High Performance Computing (HPC) resources: SURFSARA

ACADEMIC PARTNERS
(UNIPD, HESSO)

Scientific experience in extracting knowledge from:
- heterogeneous images
- text

MEDICAL PARTNERS:
(ASP, Radboudumc)

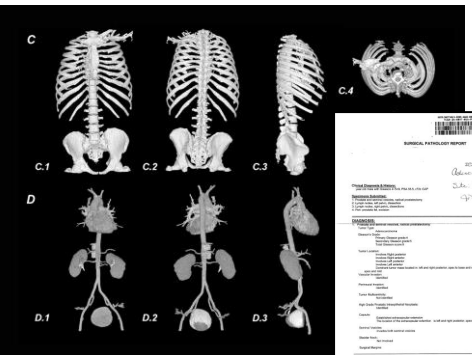
Worldwide unique providers of:
- medical data (imaging and text)
- digital pathology knowledge
- clinical evaluation experience

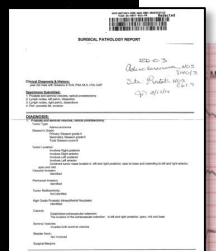
INDUSTRIAL PARTNERS:
(SIRMA AI, MICROSCOPEIT)

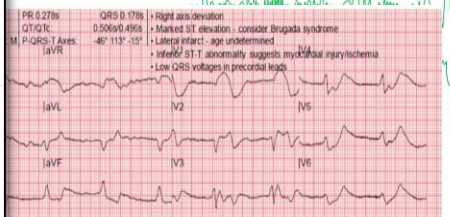
Solid industrial experience and market opportunities in:
- semantics-based services
- advanced AI-based image processing solutions

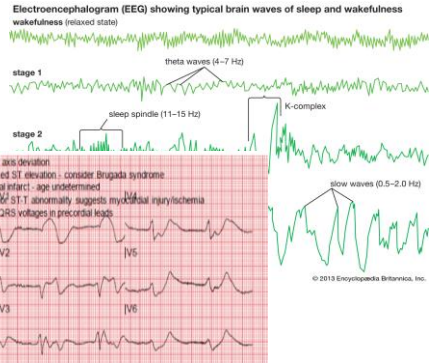
Background: biomedical data production

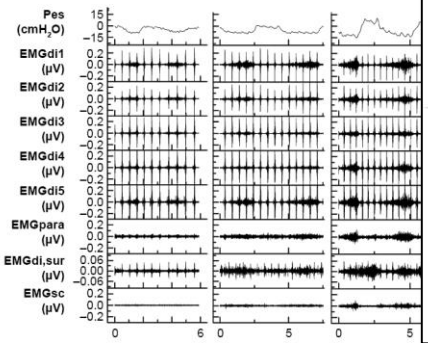
- Biomedical digital data
  - 2’000 exabytes in 2020
  - Include diagnostic reports
  - Have huge potential economical value

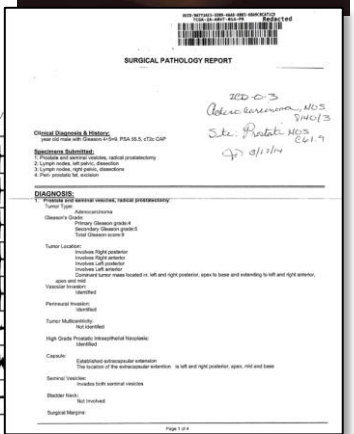


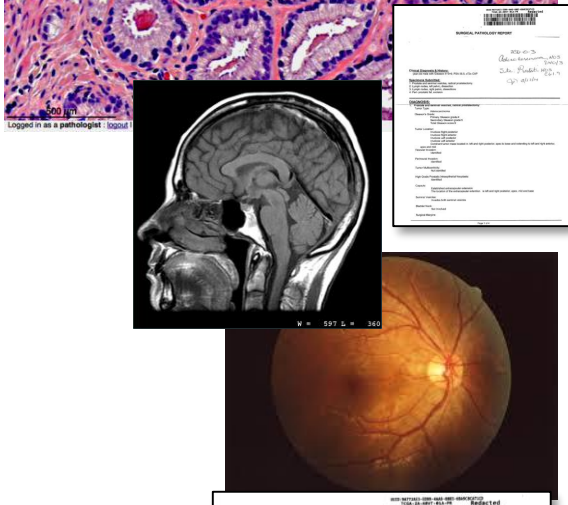










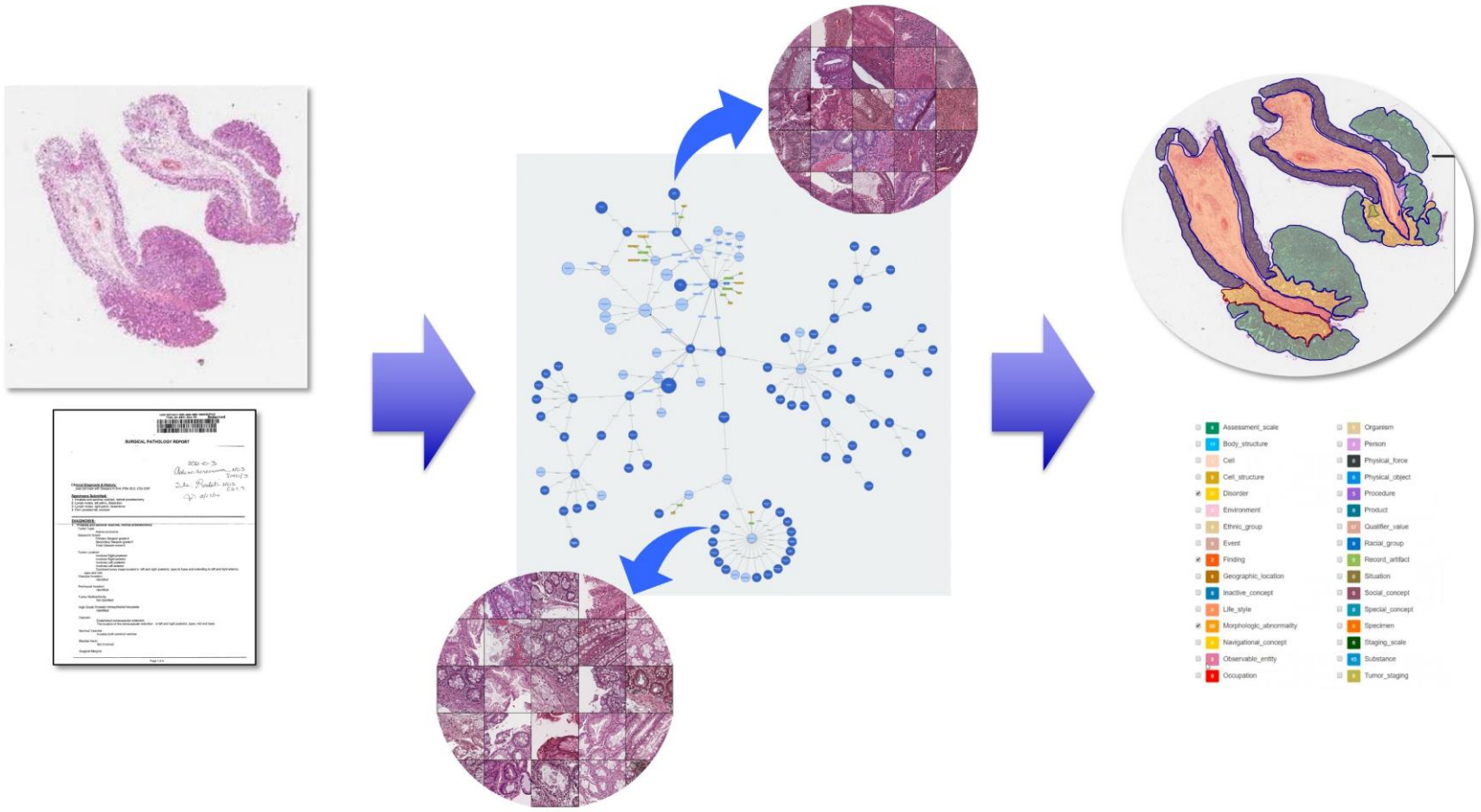


Motivation

- Several limits prevent extracting knowledge & value from extreme scale biomedical data, preventing
  - Clinical practice improvements
  - Data management improvements
  - The creation of an healthcare economy based on extreme scale data analytics

ExaMode's Objective

Easy and fast, weakly supervised knowledge discovery of extreme scale heterogeneous biomedical data



ExaMode

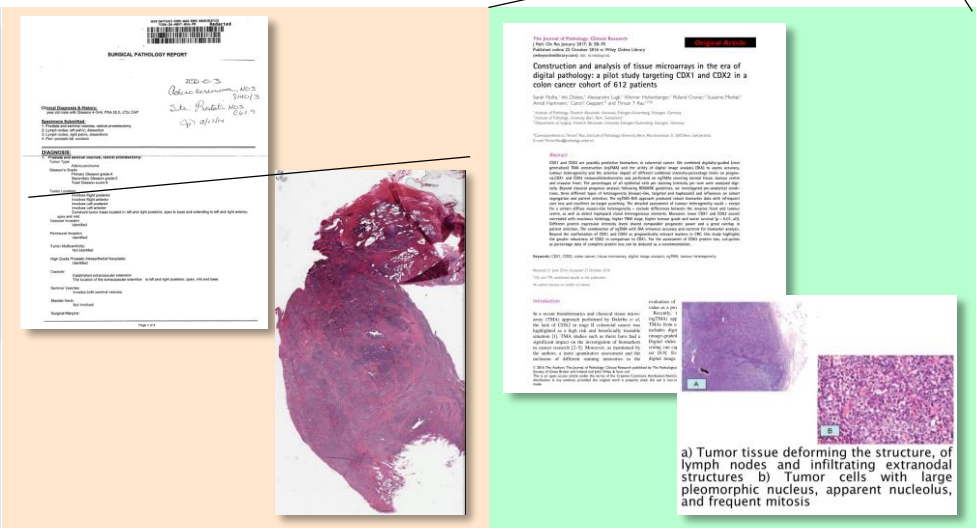
Key achievements

Project supported by European Union
Horizon 2020 grant agreement 825292

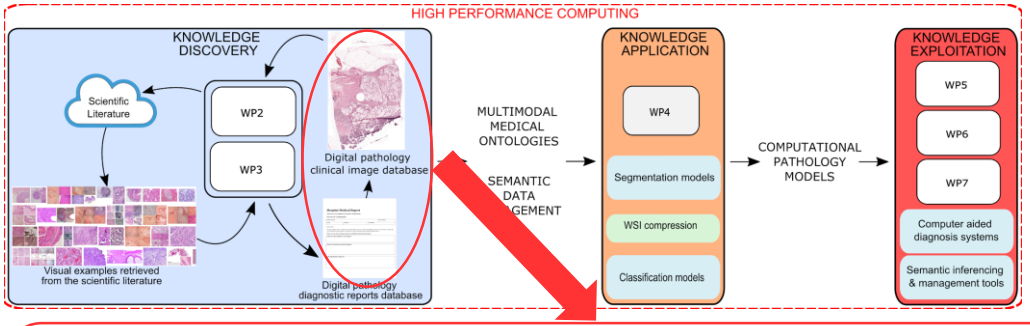
Data

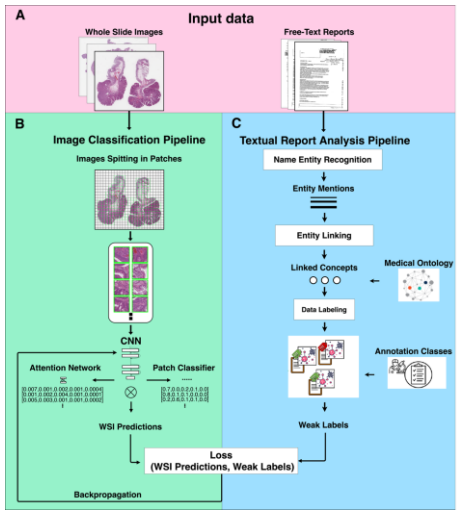
- Proprietary images & associated reports
- Publicly available images & text

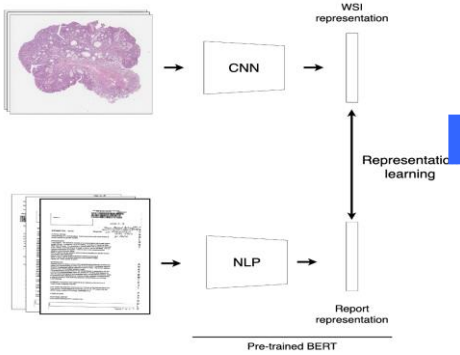
	TASK	First set of proprietary data				Final set of cured publicly available multimodal and multimedia data					
						Publicly available clinical data			Data from scientific literature		
		WSIs	TMA Images	Text	Source	Whole Slide	Text	Source	Images	Text	Source
COLON	Adenocarcinoma. Detection of cancer in polyps (in screening population).	2000		Diagnostic report, structured (table)	AOEC	50	Structured (table)	TCGAPubmc	2699	Image caption and article text	PMC Central
		9000		Synoptic report, structured (table)	Radboudumc						
		40	80	Structured (table)	Bern University						
UTERINE CERVIX	Squamous cell carcinoma	2000		Diagnostic reports, structured (table)	AOEC	45	Structured (table)	TCGA	962	Image caption and article text	PMC Central
		2500		Synoptic report	Radboudumc						
LUNG	Classification/detection of growth patterns related to cancer aggressiveness, prognosis	2000		Diagnostic report, structured (table)	AOEC	100	Structured (table)	TCGA	4151	Image caption and article text	PMC Central
CELIAC DISEASE	Celiac disease detection in duodenal biopsies	2000		Diagnostic report, structured (table)	AOEC				165	Image caption and article text	PMC Central
		1000		Synoptic report	Radboudumc						
PROSTATE	Gleason grading					50	Structured (table)	TCGA	1925	Image caption and article text	PMC Central
Additional data sources from publicly available datasets (Table 2)						12441		Various	2156		Various
TOTAL		20540	80			12686			12085		Various

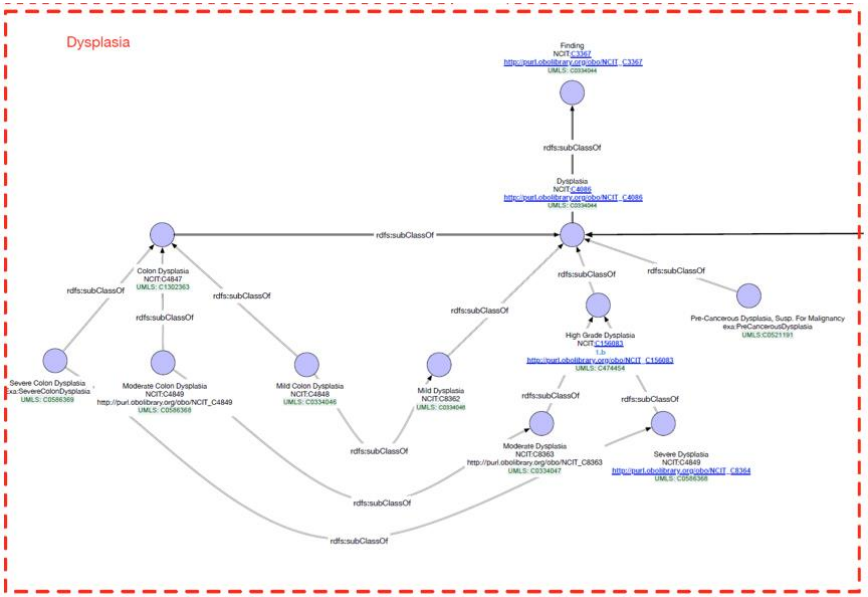


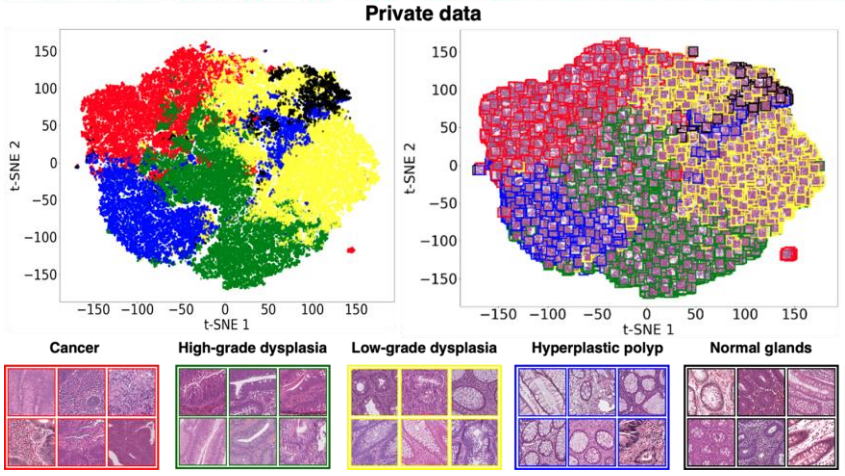
Knowledge discovery highlights: learning from clinical data





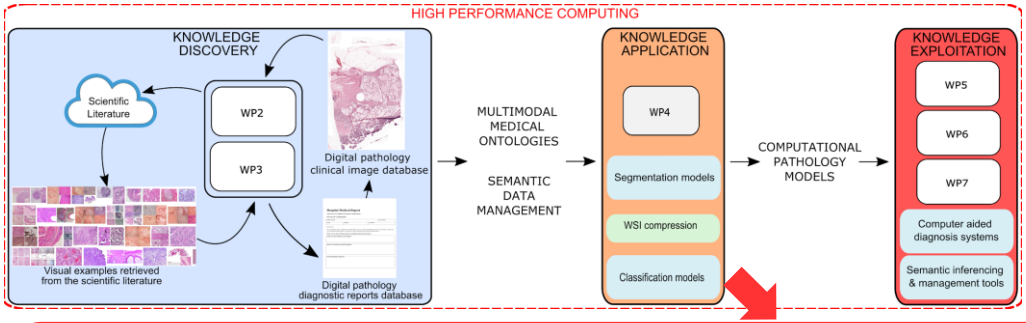






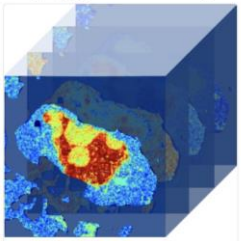
Marini N et al. "Unleashing the potential of digital pathology data by training computer-aided diagnosis models without human annotations." NPJ digital medicine 5.1 (2022): 102.
Marchesin S et al. "Empowering digital pathology applications through explainable knowledge extraction tools." Journal of pathology informatics 13 (2022): 100139.

Knowledge application highlights: computer assisted diagnostics



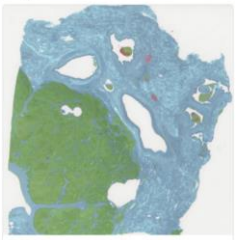
Pathology applications

https://grand-challenge.org/algorithms



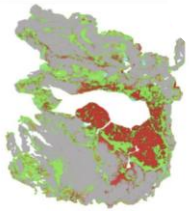
Neural Image Compression

Compresses whole slide images into much smaller volumes



HookNet-Lung

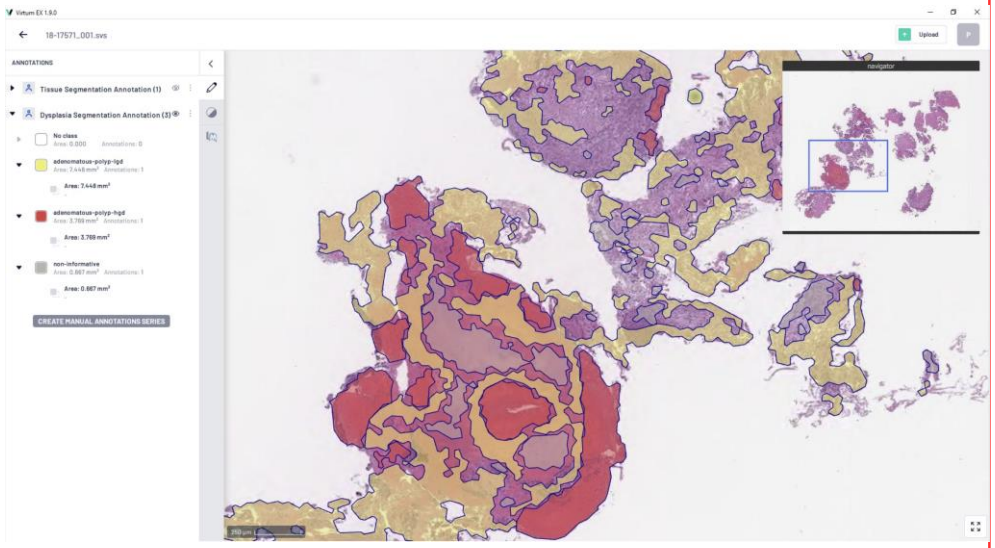
Segmentation algorithm for histopathology lung tissue.



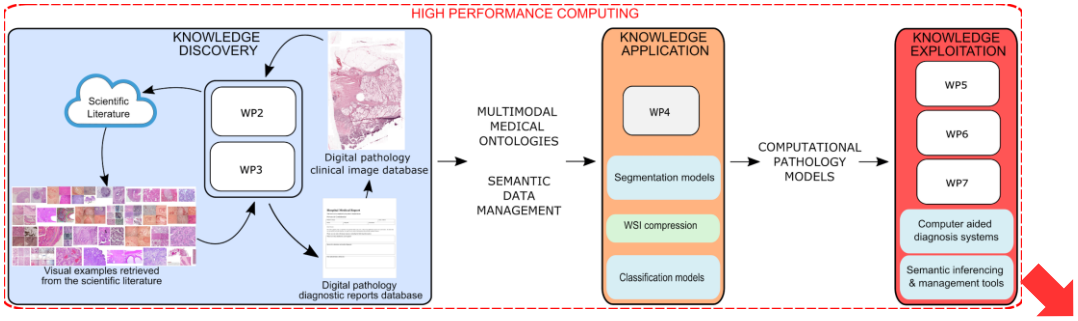
HookNet-Breast

Segmentation algorithm for histopathology breast tissue.

Product prototypes - Tooploox Virtum



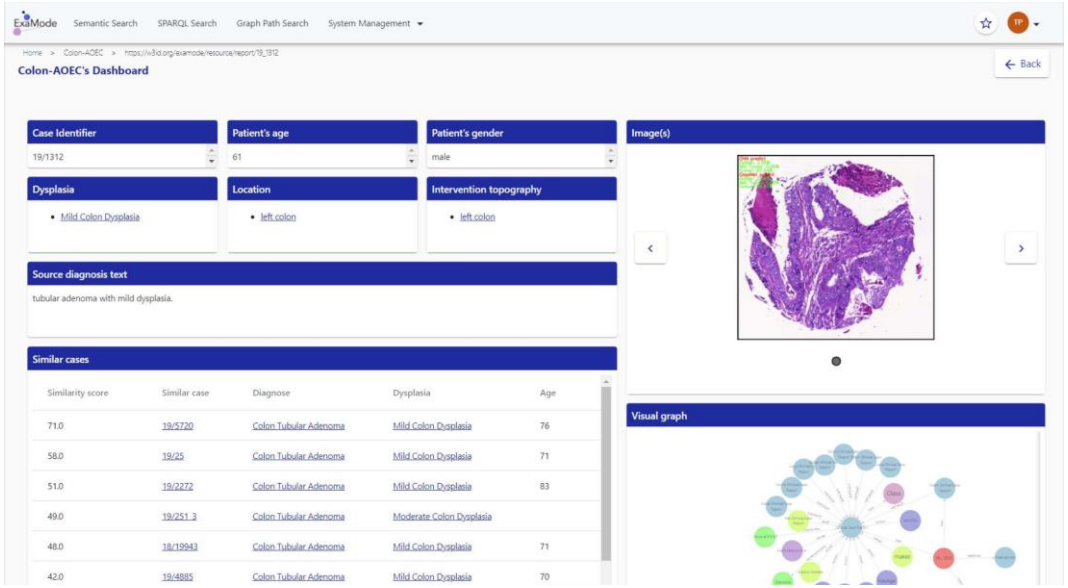
Knowledge application highlights: computer assisted diagnostics



Semantic knowledge discovery & exploitation prototype – SIRMA AI Histographer

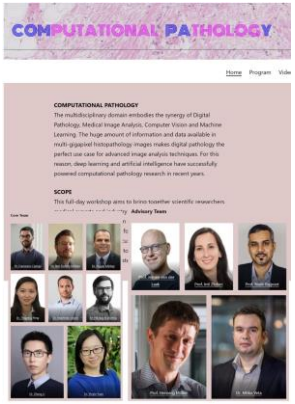
Supports pathologists in making more informed decisions based on:

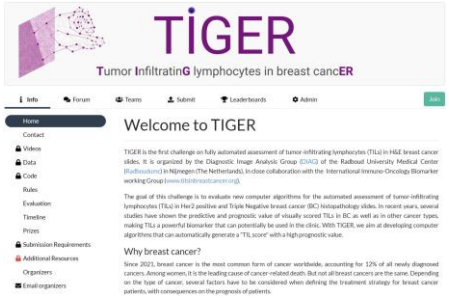
- similarity to other cases
- scientific literature

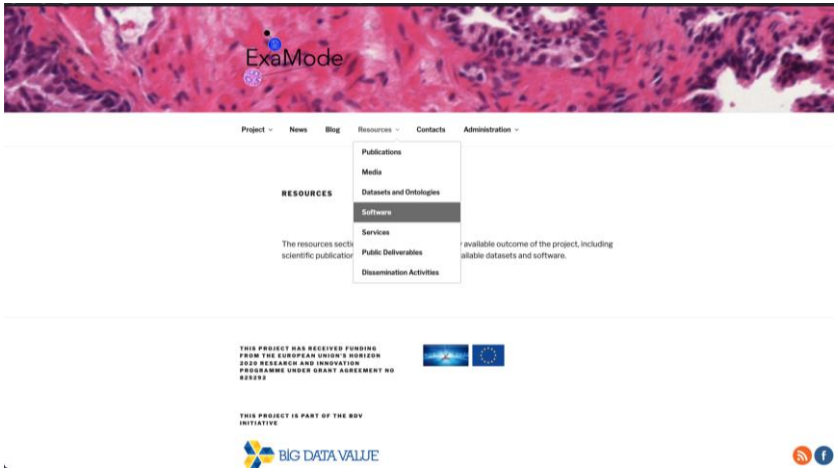


Dissemination, communication & exploitation

- Scientific dissemination
  - Over 70 journal or conference papers published
  - Several currently under review
  - Over 640 citations
  - 17 open access libraries
  - Several workshops, conferences, challenges







http://www.examode.eu/

ExaMode

Questions?

Project supported by European Union
Horizon 2020 grant agreement 825292