

AI in Medicine HPML

Bryan Cardenas

High Performance Machine Learning



What are we doing @SURF?

01.

Histopathology

Detecting cancer and
Precursor lesions

02.

Protein Design

Assist with protein design for
vaccines

03.

Hereditary

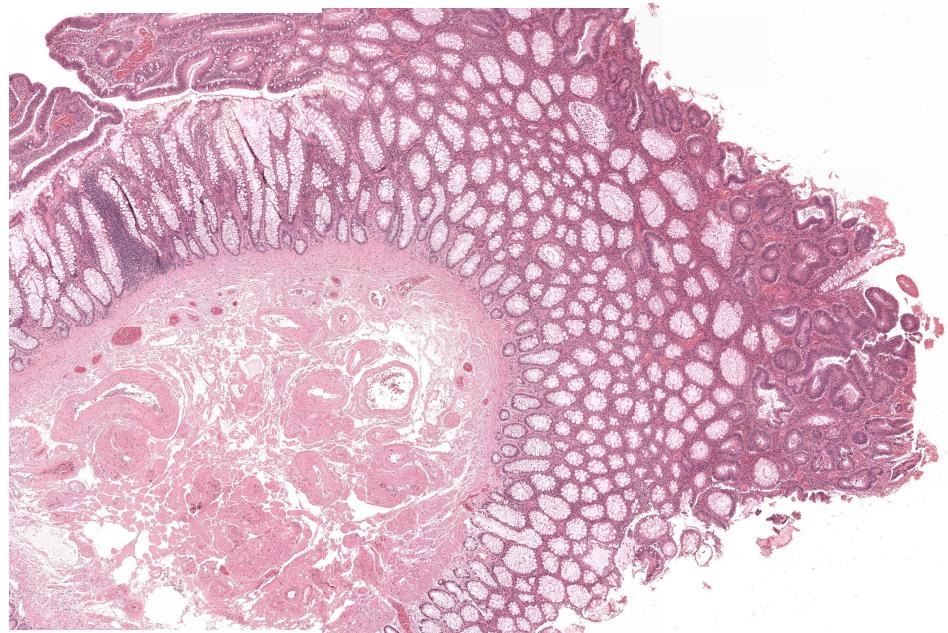
How to use Neural
Networks in multi modal
data environments across
multiple labs in the EU.

| Histopathology

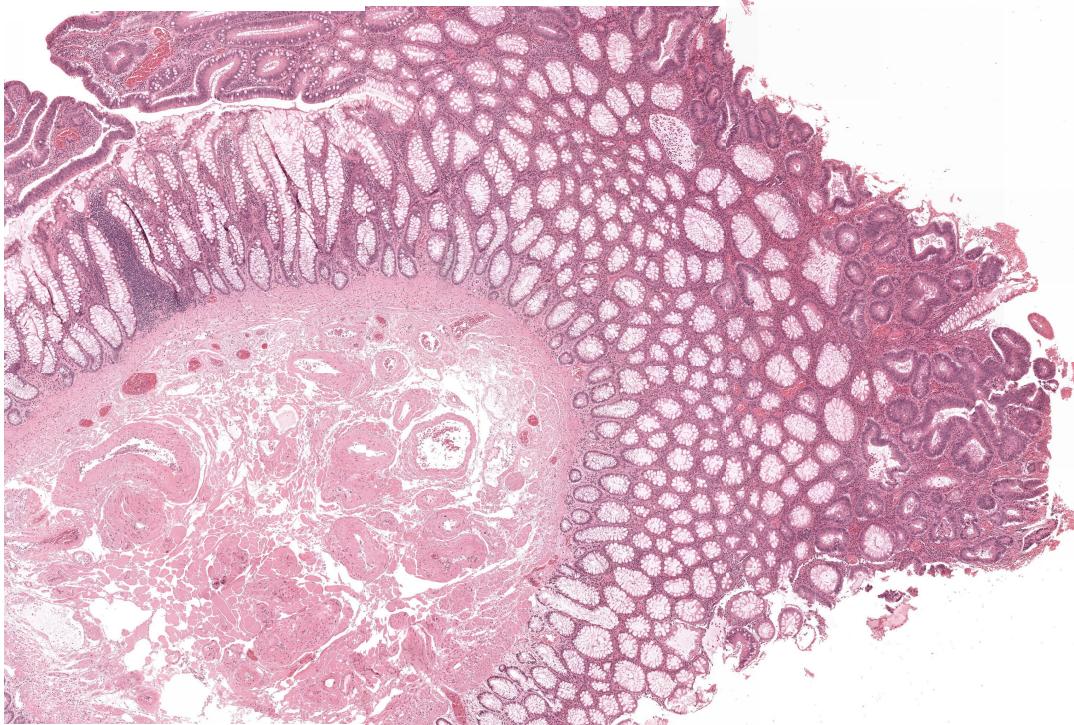


High Performance Machine Learning Group

Histopathology Data Collection

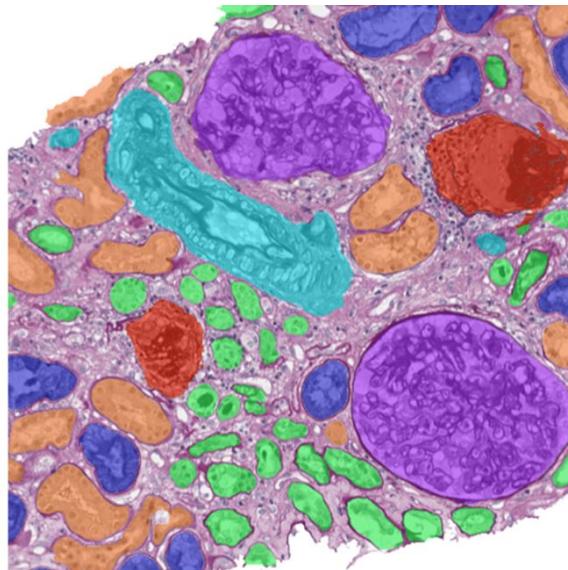
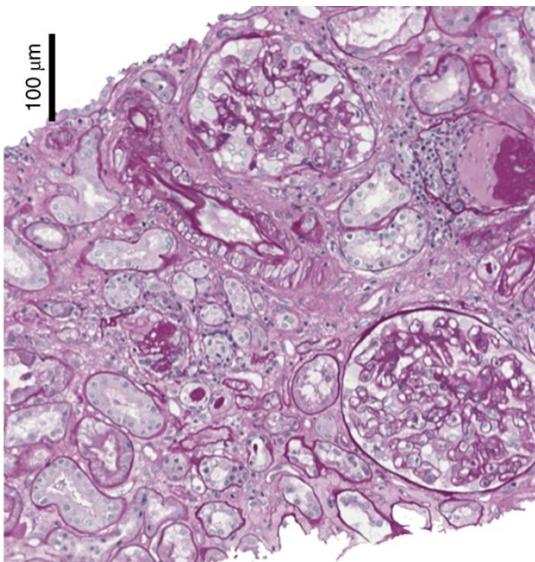


SURF Histopathology Data: Images and Diagnostic Reports



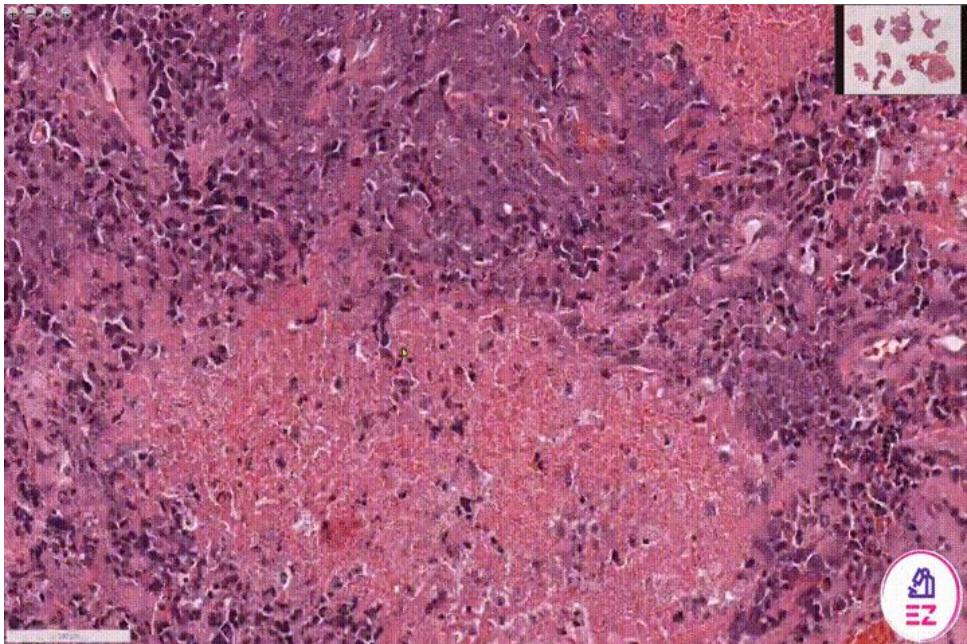
- ❖ “Colonbiops 21-25 cm en 15 cm: in beide subrieken adenocarcinoma metastase, morfologisch en immuunhistochemisch geschikt voor het al bekend primaire ovariumcarcinoom van de patiënt.”
- ❖ 2 - adenoma tubulare con displasia bassa, focalmente, alto grado (moderato, focalmente grave)

SURF Histopathology Data: Images and Diagnostic Reports



- ❖ “Colonbipt 21-25 cm en 15 cm: in beide subrieken adenocarcinoma metastase, morfologisch en immuunhistochemisch geschikt voor het al bekend primaire ovariumcarcinoom van de patiënt.”
- ❖ 2 - adenoma tubulare con displasia bassa, focalmente, alto grado (moderato, focalmente grave)

SURF Histopathology Data: Images and Diagnostic Reports



- ❖ “Colonbipt 21-25 cm en 15 cm: in beide subrieken adenocarcinoma metastase, morfologisch en immuunhistochemisch geschikt voor het al bekend primaire ovariumcarcinoom van de patiënt.”
- ❖ 2 - adenoma tubulare con displasia bassa, focalmente, alto grado (moderato, focalmente grave)

https://huggingface.co/docs/transformers/model_doc/marian



ExaMode

01.

Data Management

02.

HPC Consultancy

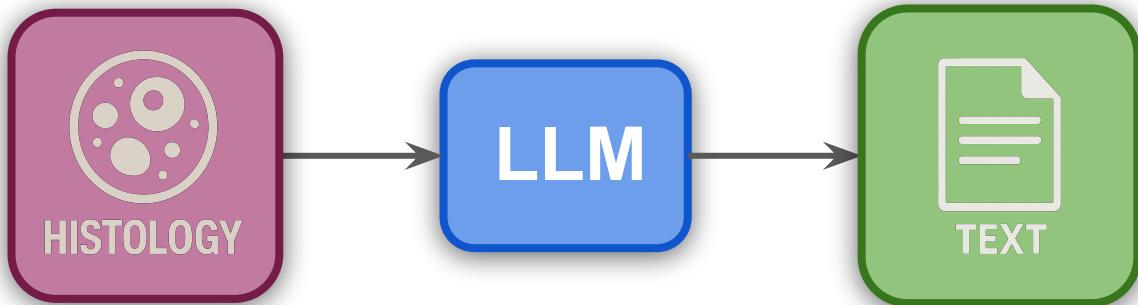
03.

Classification,
Segmentation,
Multimodal
Representation
Learning



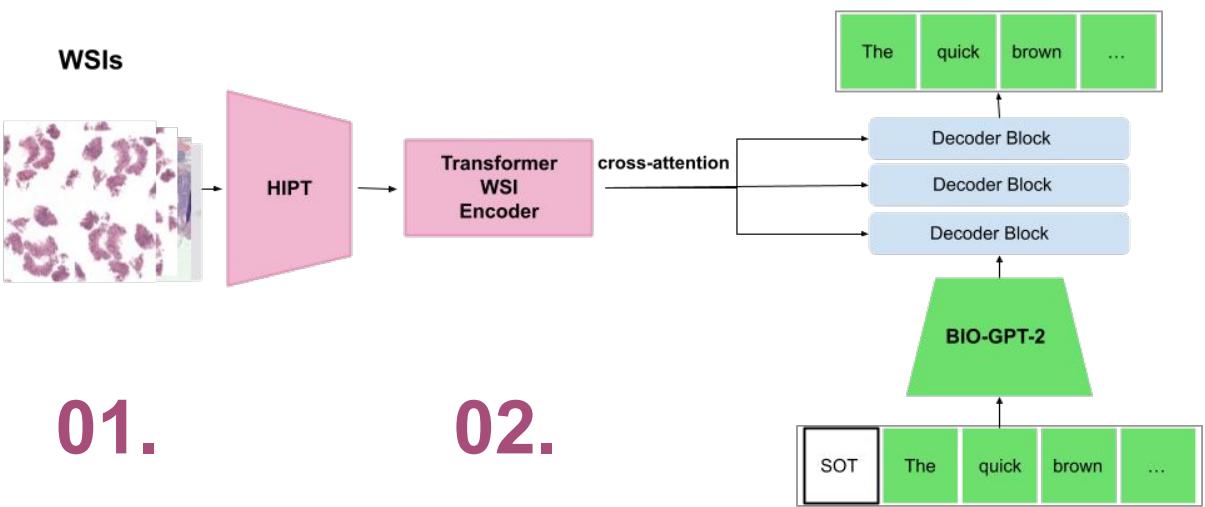


Histopathology Conditioned Language Modelling





Histopathology Conditioned Language Modelling



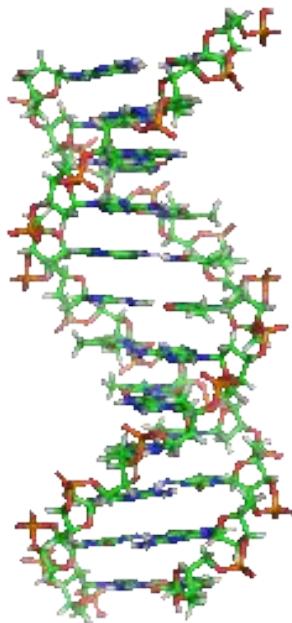
01. Data Management 02. Fitting generated text

| Protein Folding & Design

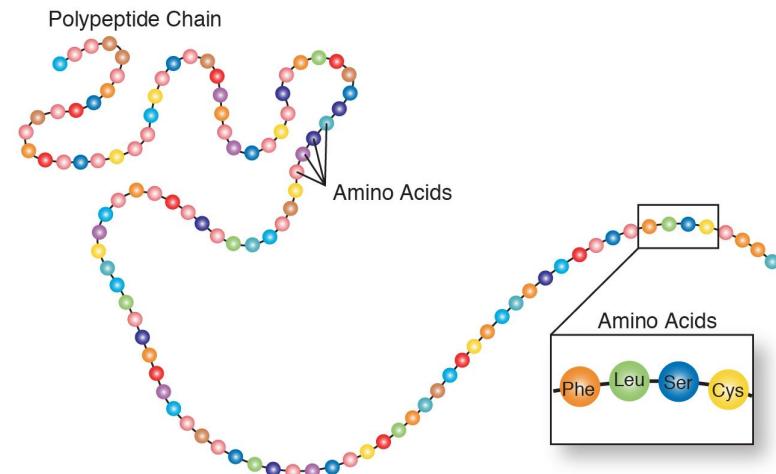


High Performance Machine Learning Group

SURF



ACTG



Amino Acids

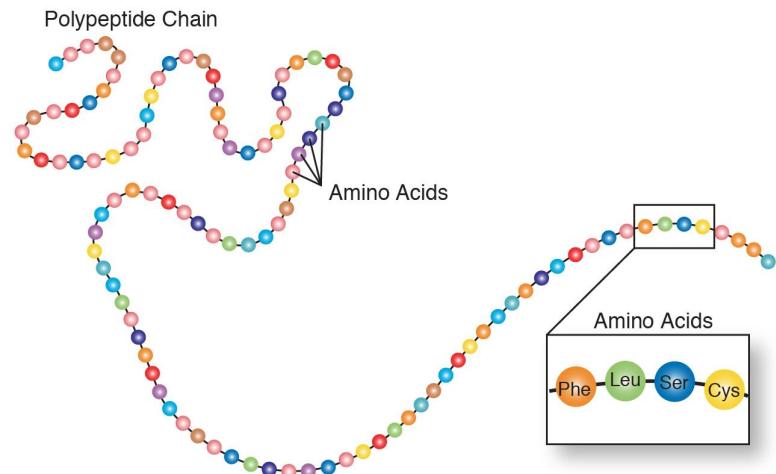
Ala: Alanine
Arg: Arginine
Asn: Asparagine
Asp: Aspartic acid
Cys: Cysteine

Gln: Glutamine
Glu: Glutamic acid
Gly: Glycine
His: Histidine
Ile: Isoleucine

Leu: Leucine
Lys: Lysine
Met: Methionine
Phe: Phenylalanine
Pro: Proline

Ser: Serine
Thr: Threonine
Trp: Tryptophane
Tyr: Tyrosine
Val: Valine

SURF



Amino Acids

Ala: Alanine
Arg: Arginine
Asn: Asparagine
Asp: Aspartic acid
Cys: Cysteine

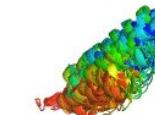
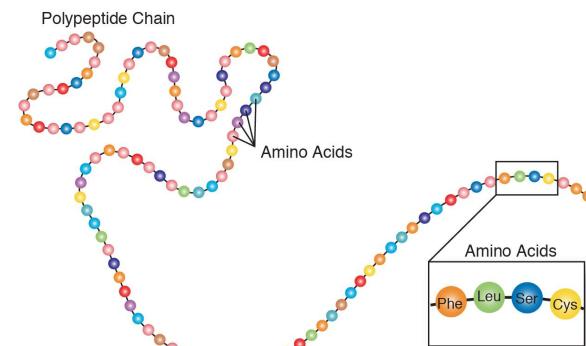
Gln: Glutamine
Glu: Glutamic acid
Gly: Glycine
His: Histidine
Ile: Isoleucine

Leu: Leucine
Lys: Lysine
Met: Methionine
Phe: Phenylalanine
Pro: Proline

Ser: Serine
Thr: Threonine
Trp: Tryptophane
Tyr: Tyrosine
Val: Valine



AlphaFold2



1



AlphaFold2

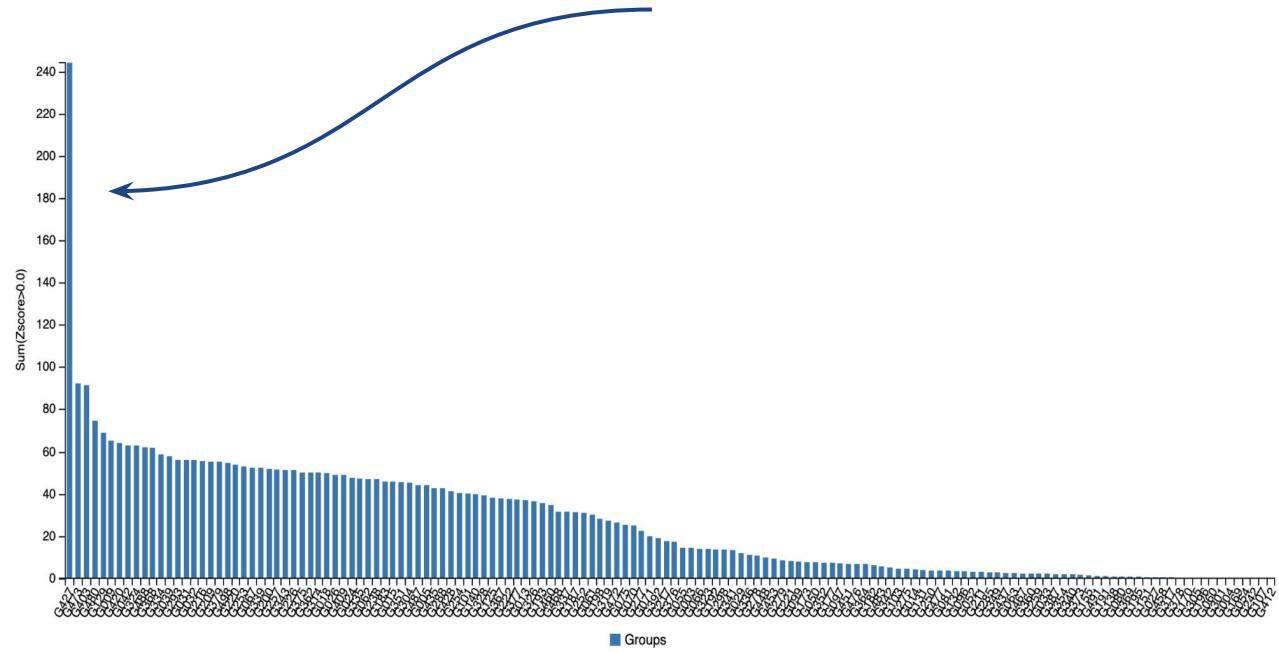


©nature

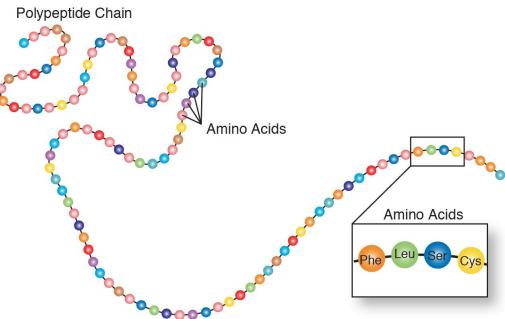
SURF



AlphaFold2



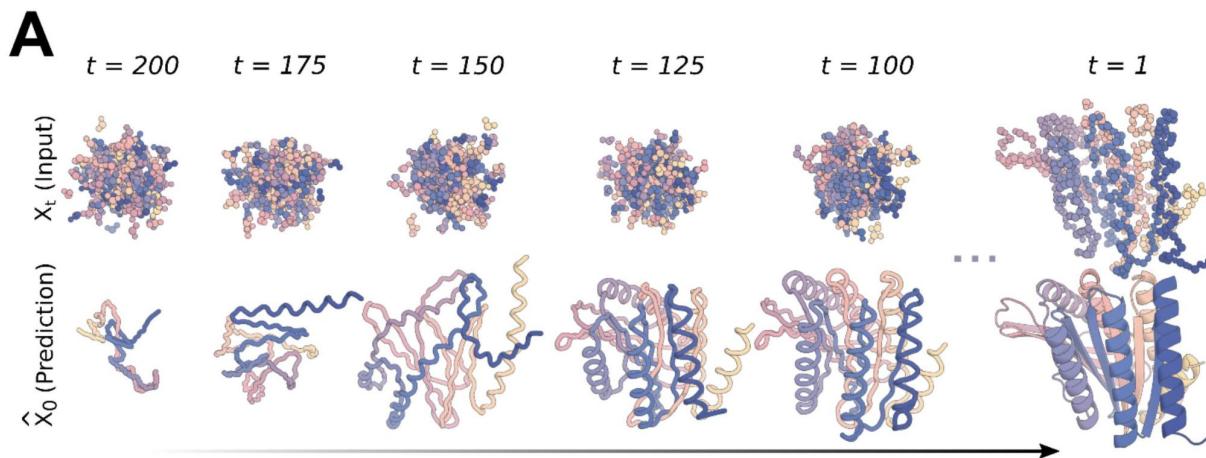
SURF



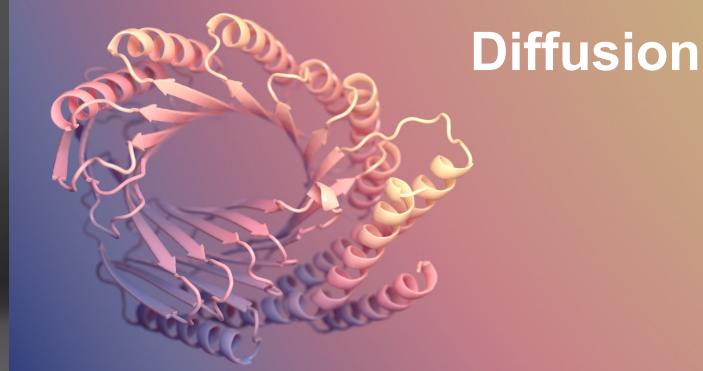
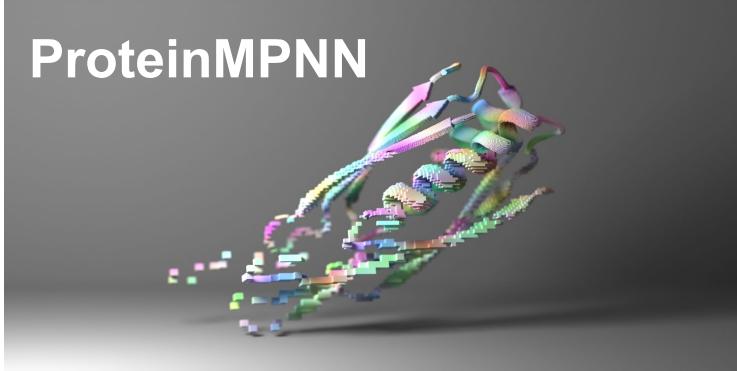
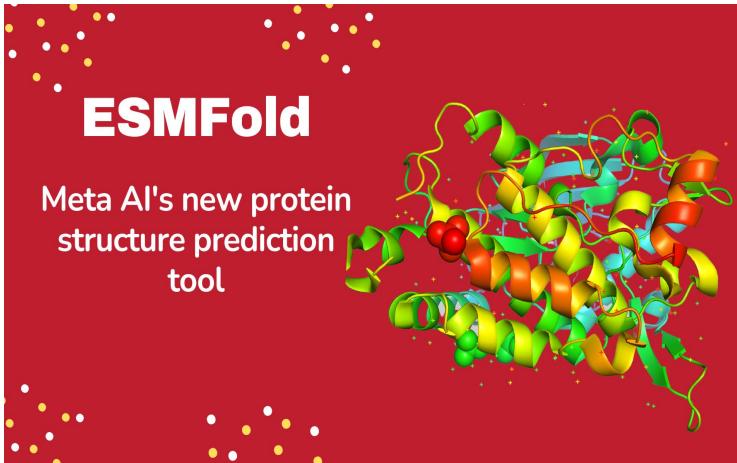
→ AlphaFold2 →

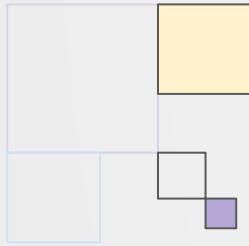


1

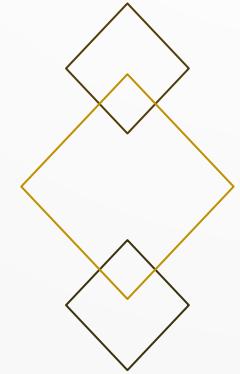


Suite of Protein Models





Experimental vaccinology@AMC

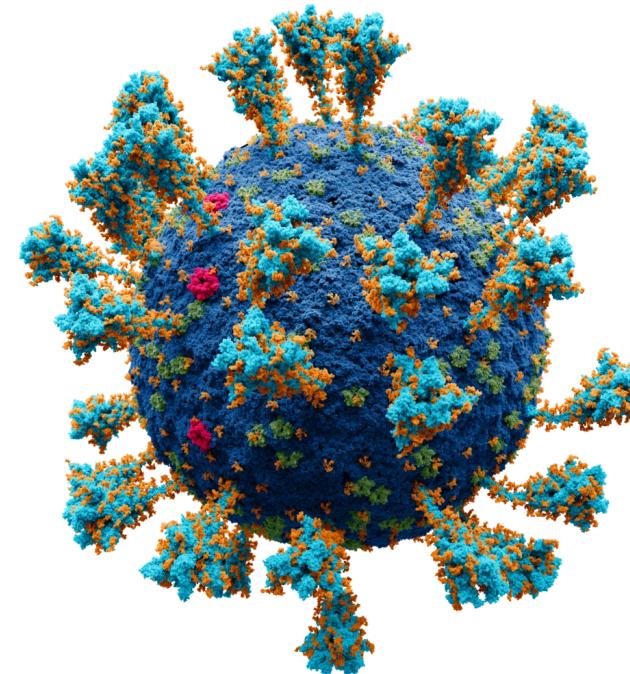


High Performance Machine Learning Group

SURF



SARS-COV-2

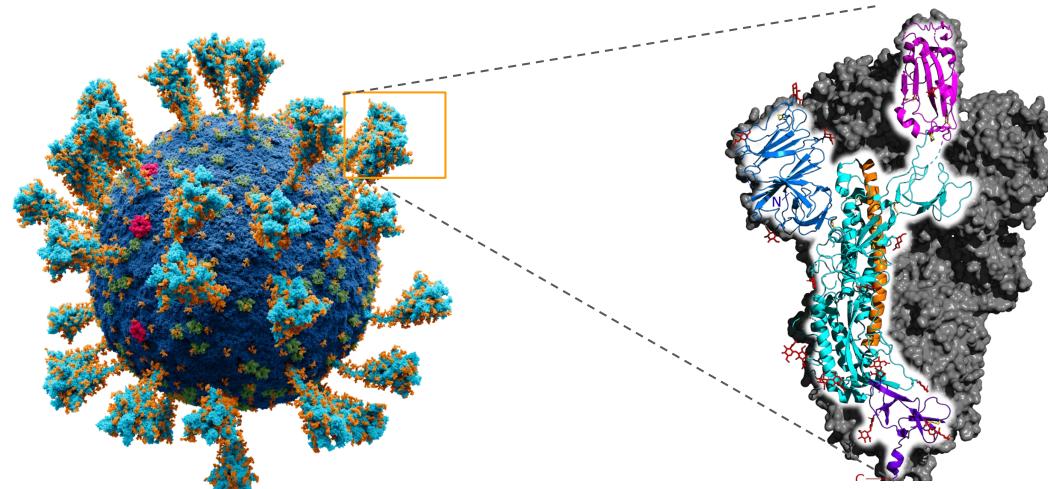


SURF



SARS-COV-2

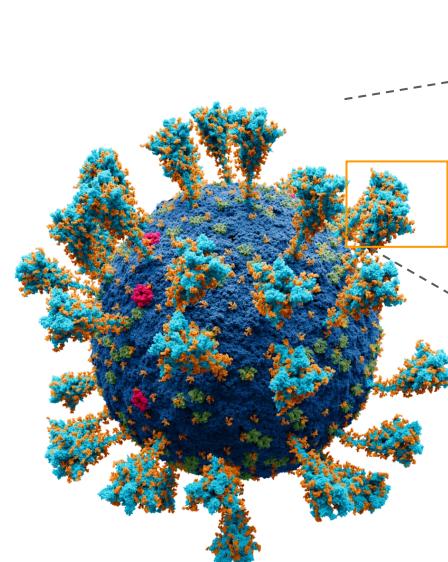
Spike Protein



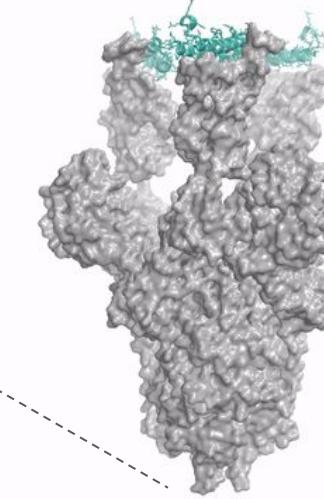
SURF



SARS-COV-2

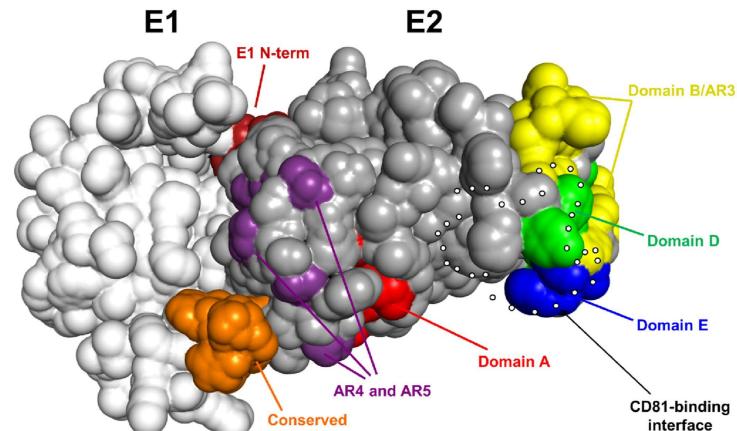
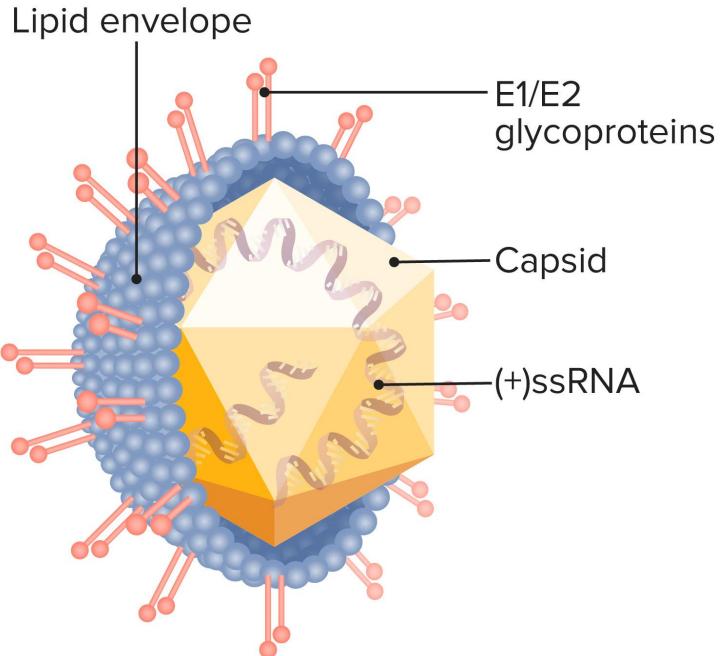


Spike Protein



SURF

Hepatitis C Virus



Recipe for Assistance

01.

Provide hardware &
storage

02.

Run software

03.

Efficiency

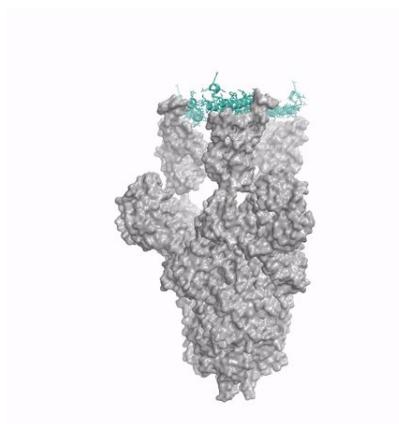
04.

Think of a new way
to tackle the
problem

Chaining Predictions

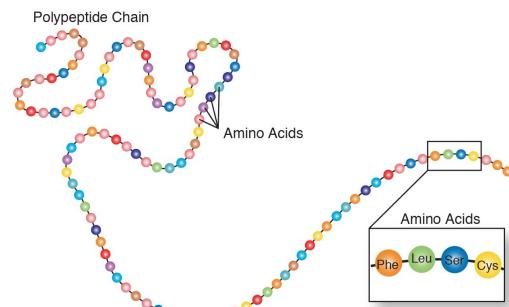
01. Generation

RFDiffusion



02. Predict Sequence

ProteinMPNN



03. Confirm Structure

ESMFold, ColabFold,
AlphaFold, OmegaFold,
RoseTTaFold



SURF

Hereditary

High Performance Machine Learning Group





Hereditary

01.

Federated Learning

02.

Provide compute
& HPC support

03.

Train Models on
multiple modalities

04.

Security & Compliance
monitoring System



Hereditary Data Types



GENOMICS



CLINICAL
DATA



TEXT



ONTO TEXT



EEG



MRI



CT



HISTOLOGY

Federated Learning

Improve diagnostic accuracy
but maintain privacy!

Make data comply with privacy
and ethics

Multiple centers benefit!



SURF



Gut-Brain Interplay

01. Neurological Diseases

02. Gut Microbiome Disorders

Multi-Modality
Multi-Linguality

Data Assembly
Privacy & Security

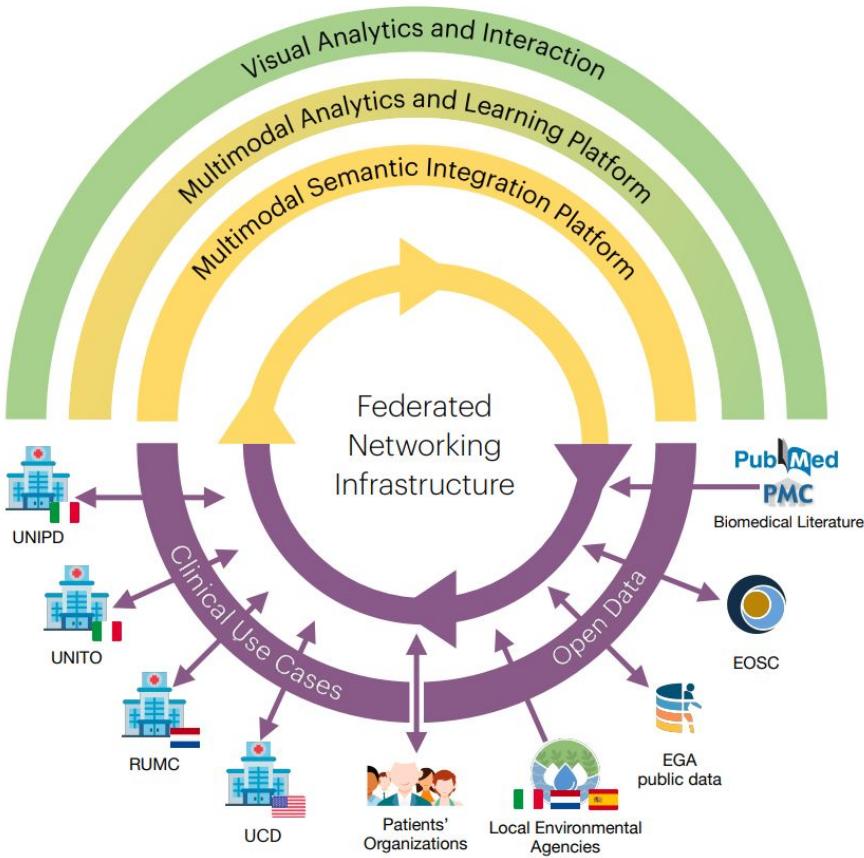


Funded by
the European Union



HEREDITARY

The Call & Objectives



Secure & lawful federated access to linked multimodal data across Europe

Healthcare data normalization

Multimodal Ontologies

Federated Polystore for querying
Federated Learning Models

Analytics platform for aggregation and discovery from multiple sources

Patient Recommendations

Thanks!

Mail me for Questions:

bryan.cardenasguevara@surf.nl

SURF

High Performance Machine Learning Group