SISONKE BIOTIK

GROWING AI FOR HEALTHCARE IN AFRICA

TELLING OUR STORY

HOUCEMEDDINE TURKI
DATA ENGINEERING AND SEMANTICS RESEARCH UNIT, UNIVERSITY OF SFAX, TUNISIA

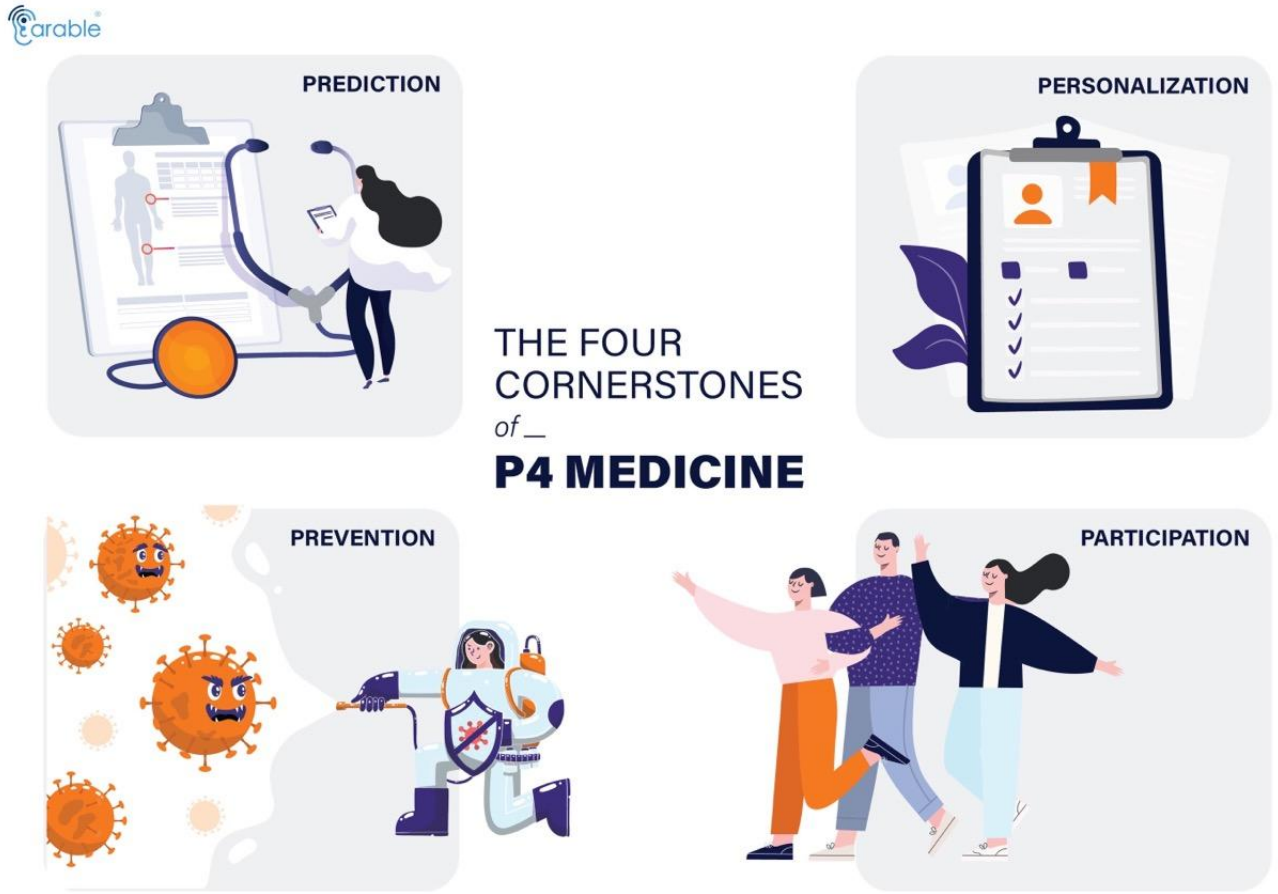
BONAVENTURE DOSSOU
QUEBEC ARTIFICIAL INTELLIGENCE INSTITUTE, CANADA

CHRIS EMEZUE
TECHNISCHE UNIVERSITÄT MÜNCHEN, GERMANY

ABRAHAM OWODUMNI
UNIVERSITY OF ILORIN, NIGERIA

GROWING AI FOR HEALTHCARE IN AFRICA

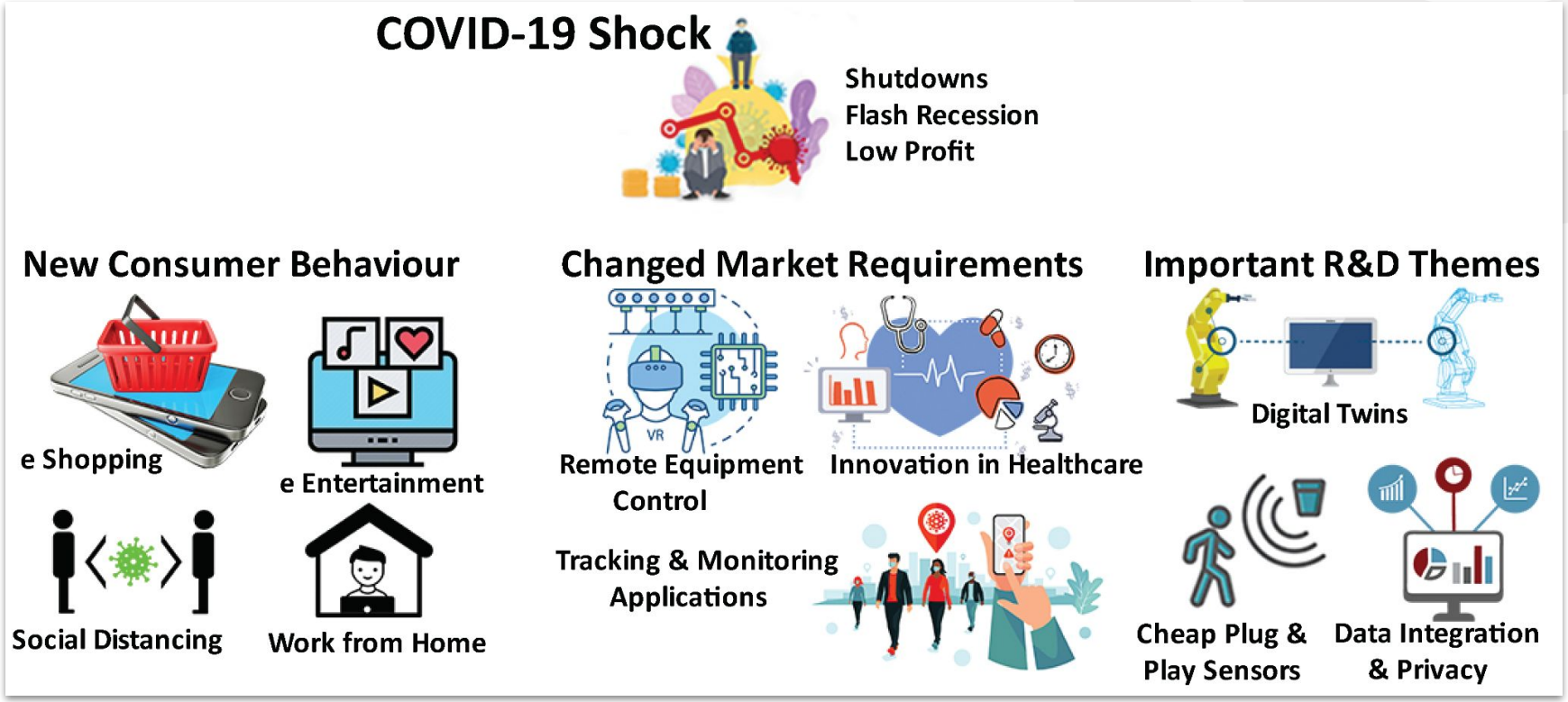
AI: ‘general purpose technology’ that can leverage digital data to address challenges across the 4Ps Medicine:



DEEP LEARNING INDABA

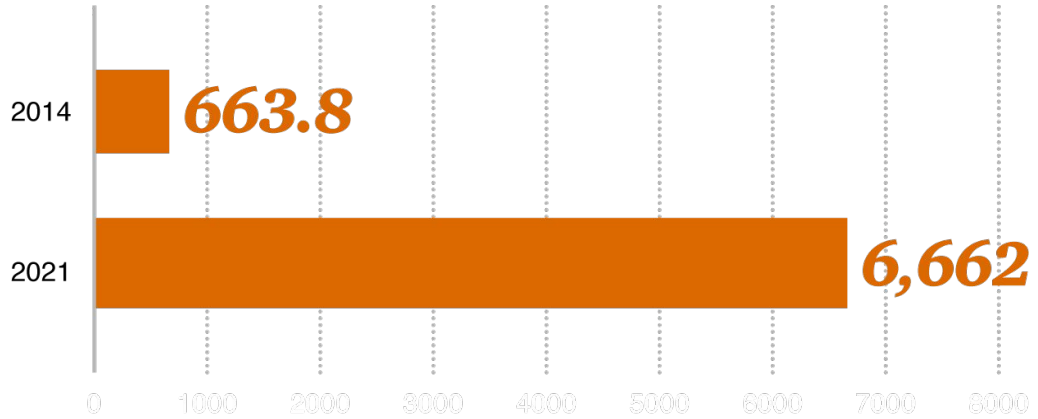
GROWING AI FOR HEALTHCARE IN AFRICA

- **COVID-19 boosted** the development of **AI use** cases globally.



GROWING AI FOR HEALTHCARE IN AFRICA

Figure 1: Artificial Intelligence Market for Healthcare Applications, World, 2014, 2021 (in Millions)



Source: Frost & Sullivan 2016 Transforming healthcare through artificial intelligence systems

New AI-powered algorithm to predict third wave of COVID-19 in South Africa

April 12, 2021 Faculty of Science; Artificial Intelligence; South Africa; Wits University sandramc

TORONTO, April 12, 2021 – An Artificial Intelligence (AI)-based algorithm, designed by the University of the Witwatersrand (Wits University) in partnership with York University, iThemba LABS, and the Provincial Government of Gauteng, shows there is a low risk for a third wave of COVID-19 infection in all provinces of South Africa.

- **COVID-19 boosted** the development of **AI use** cases in the **medical sector**.

- **In other cases, research shows that COVID-19 also created a compounded historical & digital divides.**

DEEP LEARNING INDABA

SOME PROBLEMS...

Lack of **accurate**, **up to date**, complete, representative **datasets** and technical **skills** network.

Lower accuracy of algorithms trained on **foreign data not representative** of African data

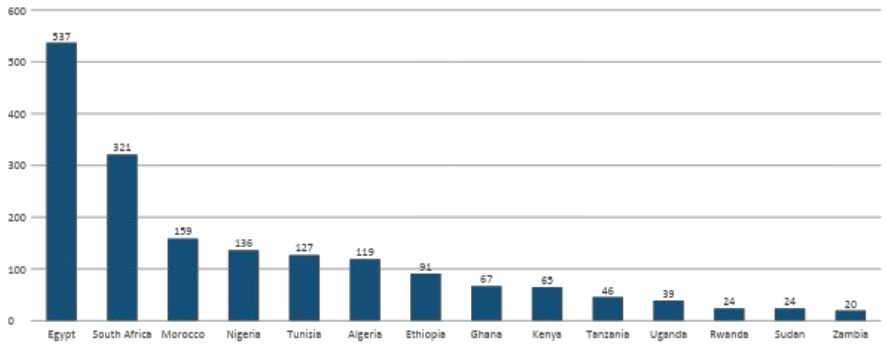
Gaps in **digital infrastructure** and inclusion (including internet connectivity and uptake)

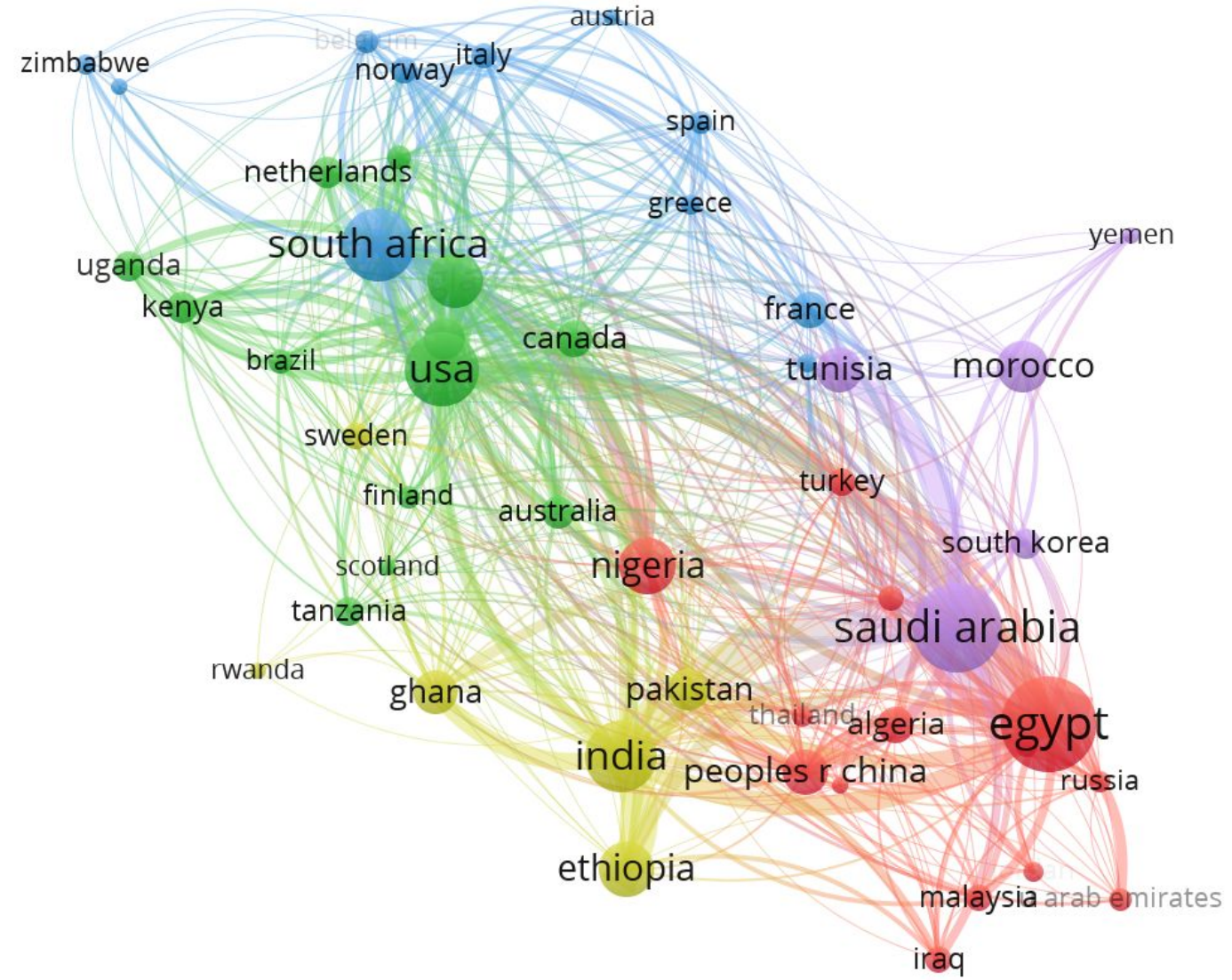
And in some cases, lack of **trust** and data **governance**

DEEP LEARNING INDABA

Not all African countries are represented

South Africa and Egypt are at the forefront, with other African countries having little contribution.



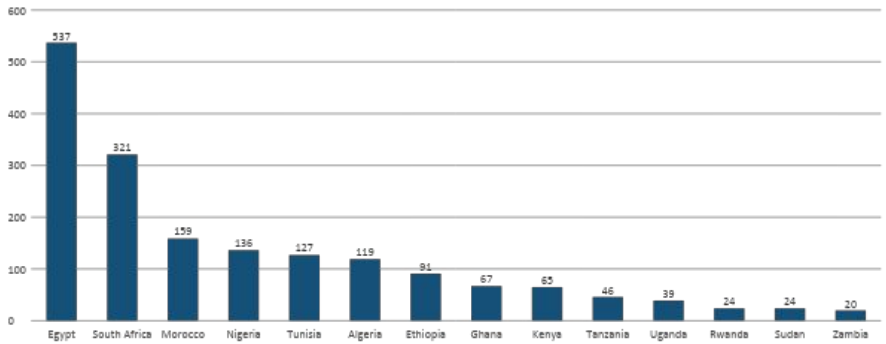


DEEP LEARNING INDABA

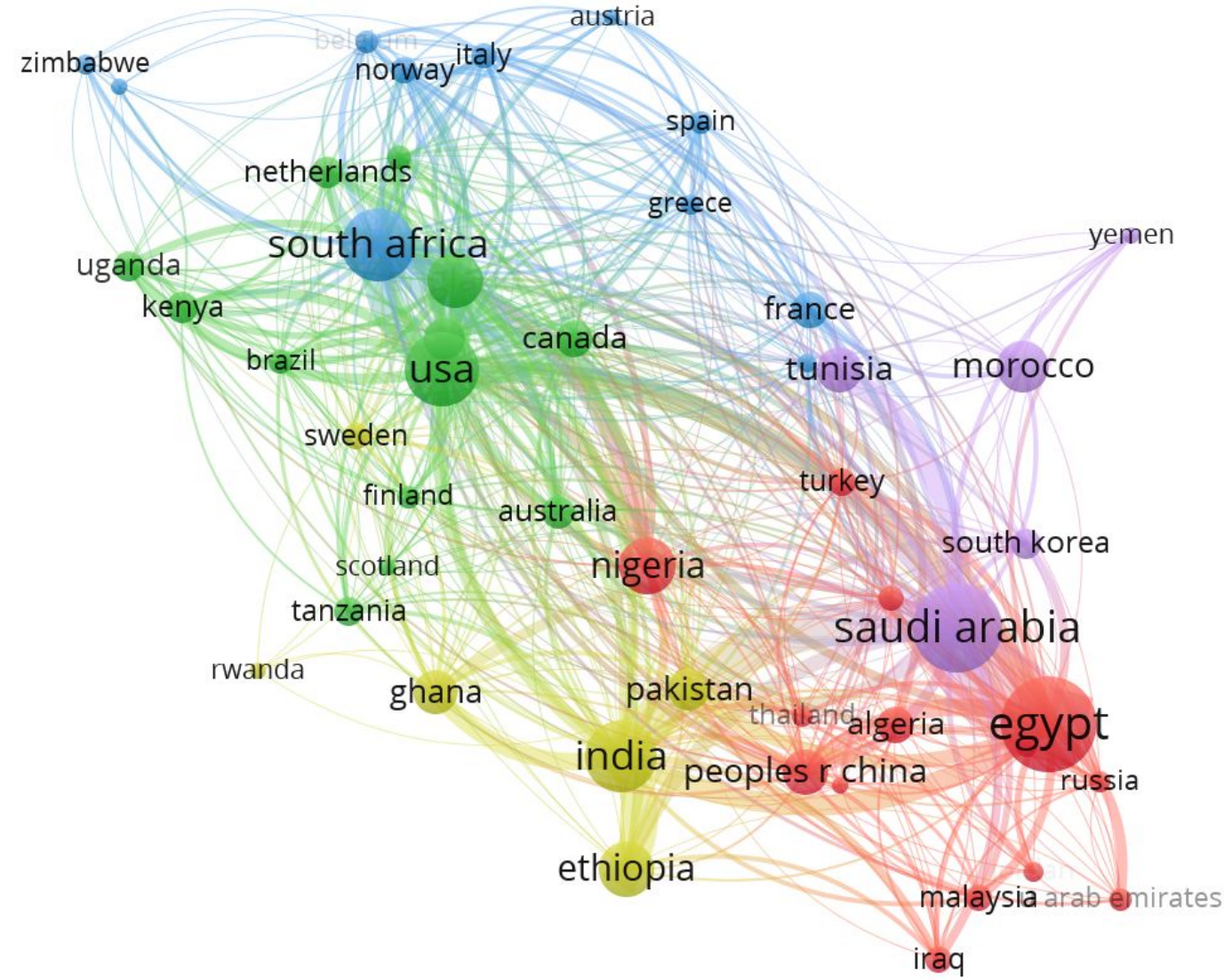
SISONKE BIOTIK

Not all African countries are represented

South Africa and Egypt are at the forefront, with other African countries having little contribution.



We need to represent many more African countries

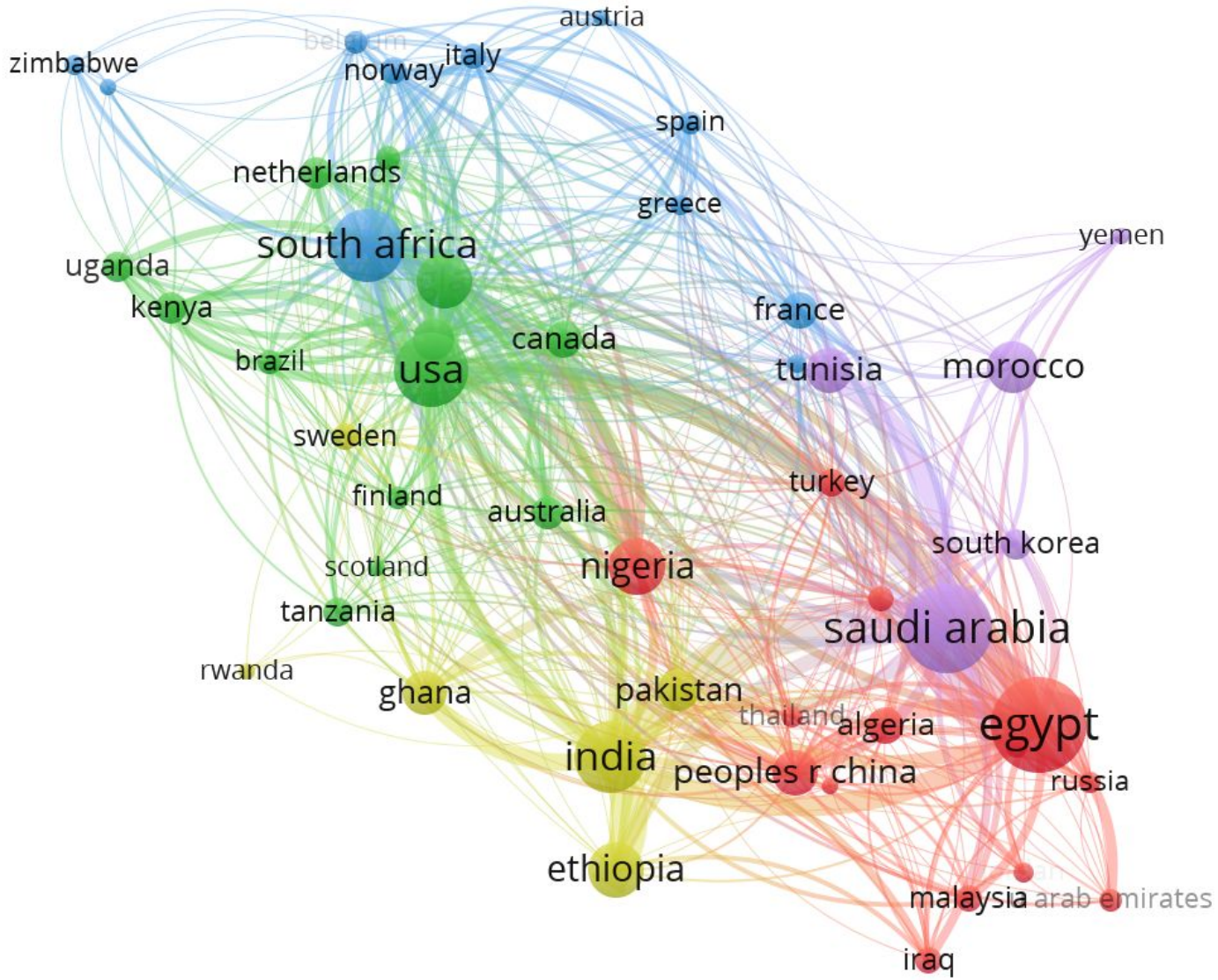


DEEP LEARNING INDABA

SISONKE BIOTIK

A larger Influence from external stakeholders

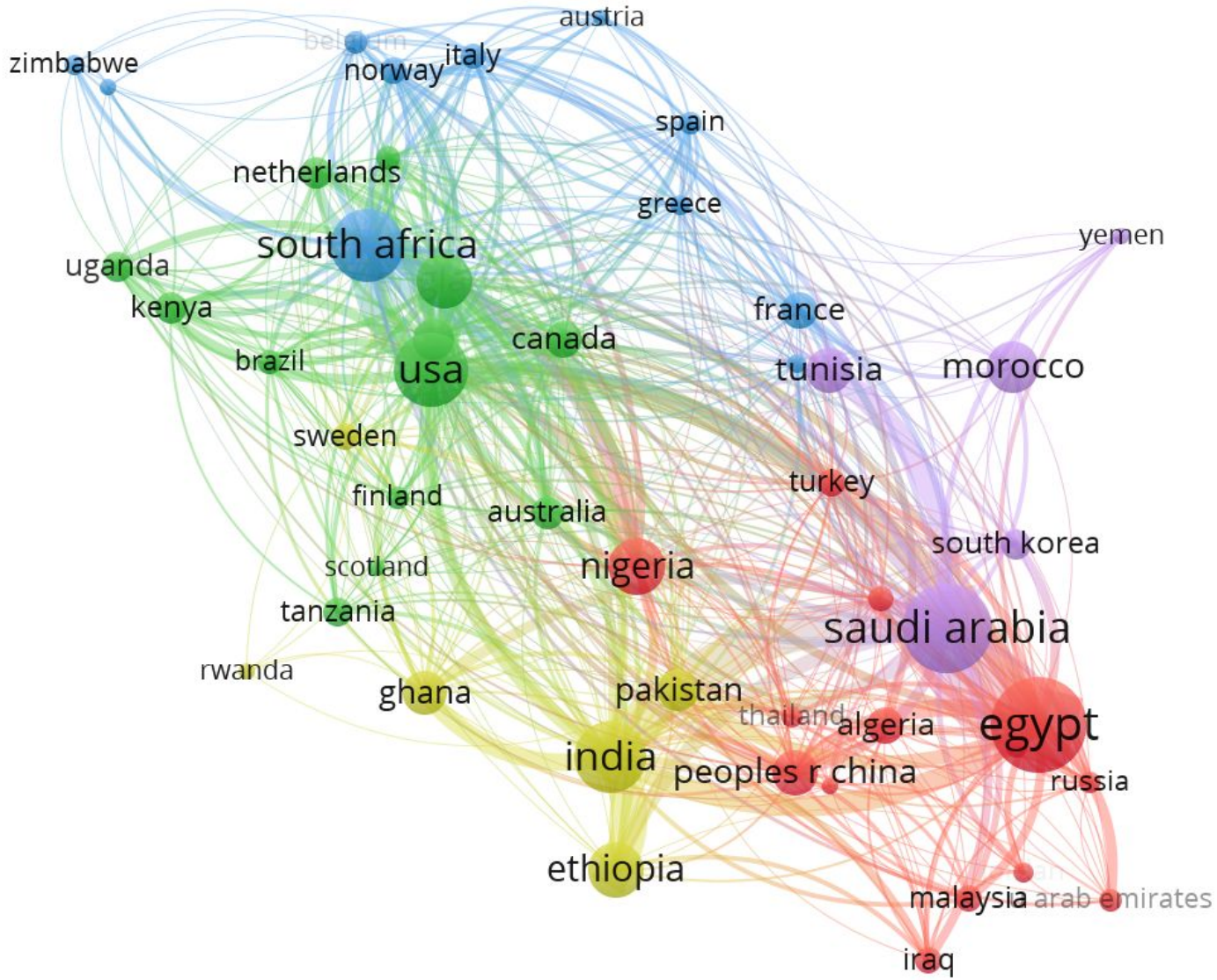
External, non-African funders, collaborators (like the US and Saudi Arabia).



A larger Influence from external stakeholders

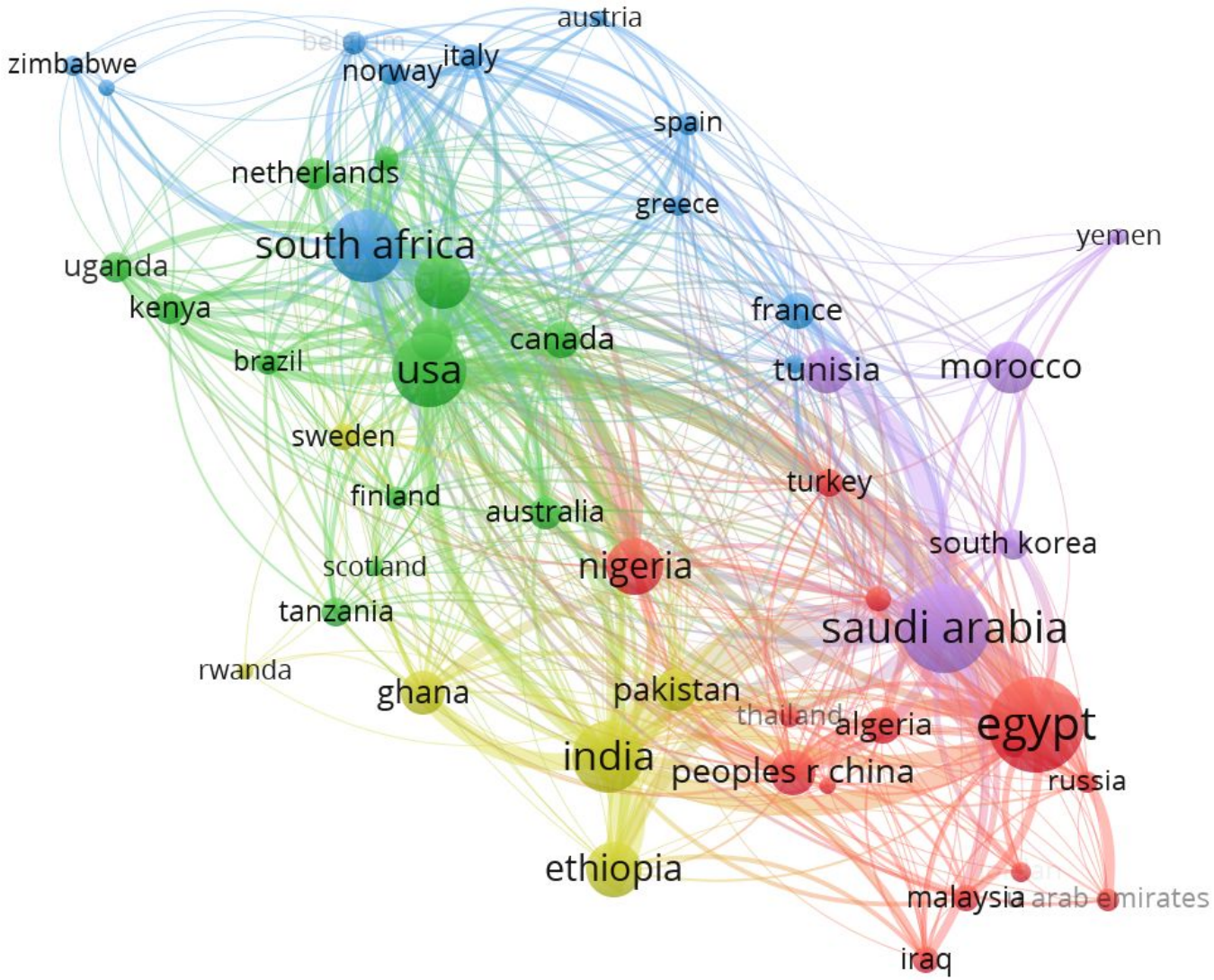
External, non-African funders, collaborators (like the US and Saudi Arabia).

We need more african-centric funders, organizations and collaborations.



Uneven collaborative network of AI for health in African

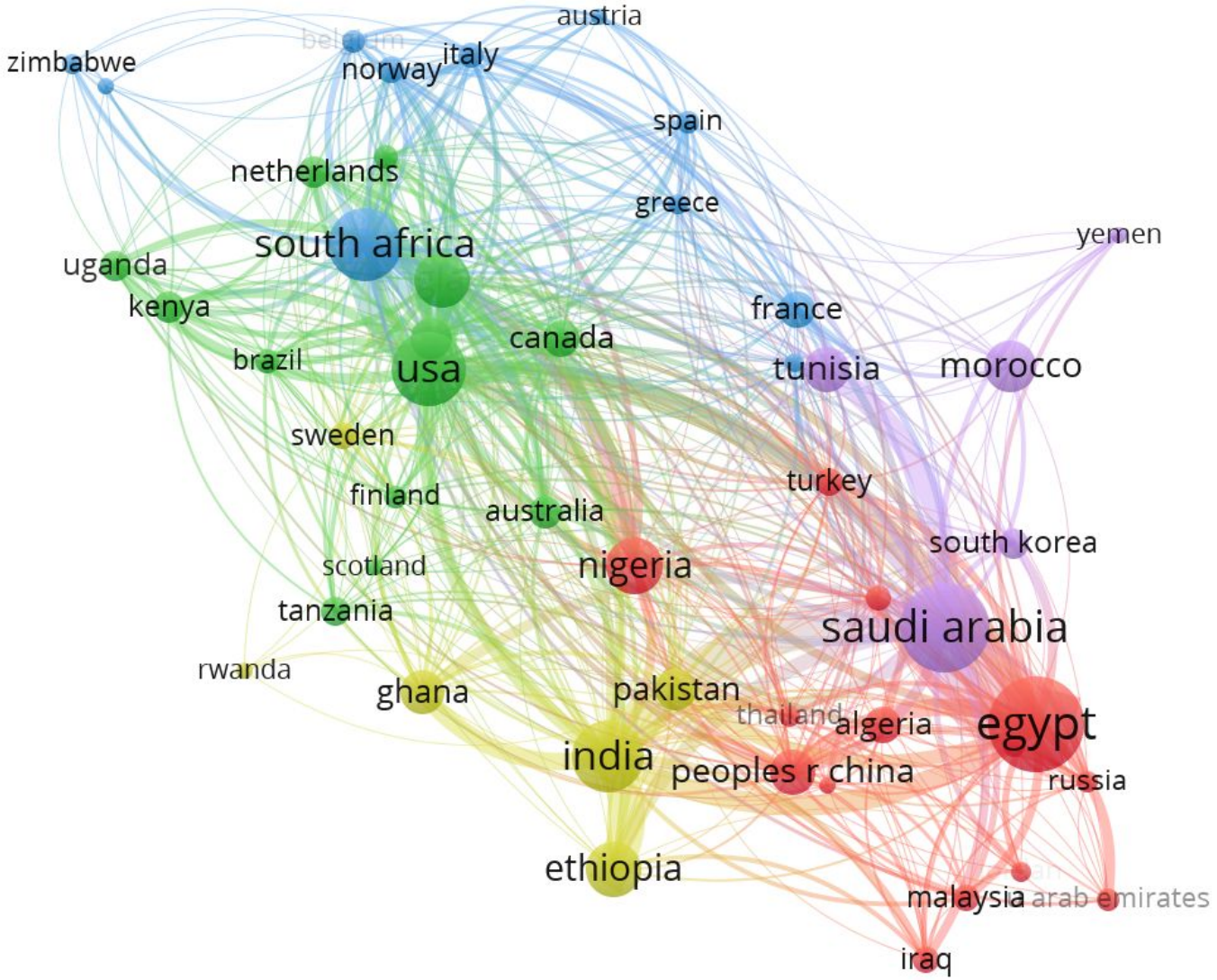
French-speaking countries not represented enough.



Uneven collaborative network of AI for health in African

French-speaking countries not represented enough.

We need a more diverse community to get involved in AI for health research.



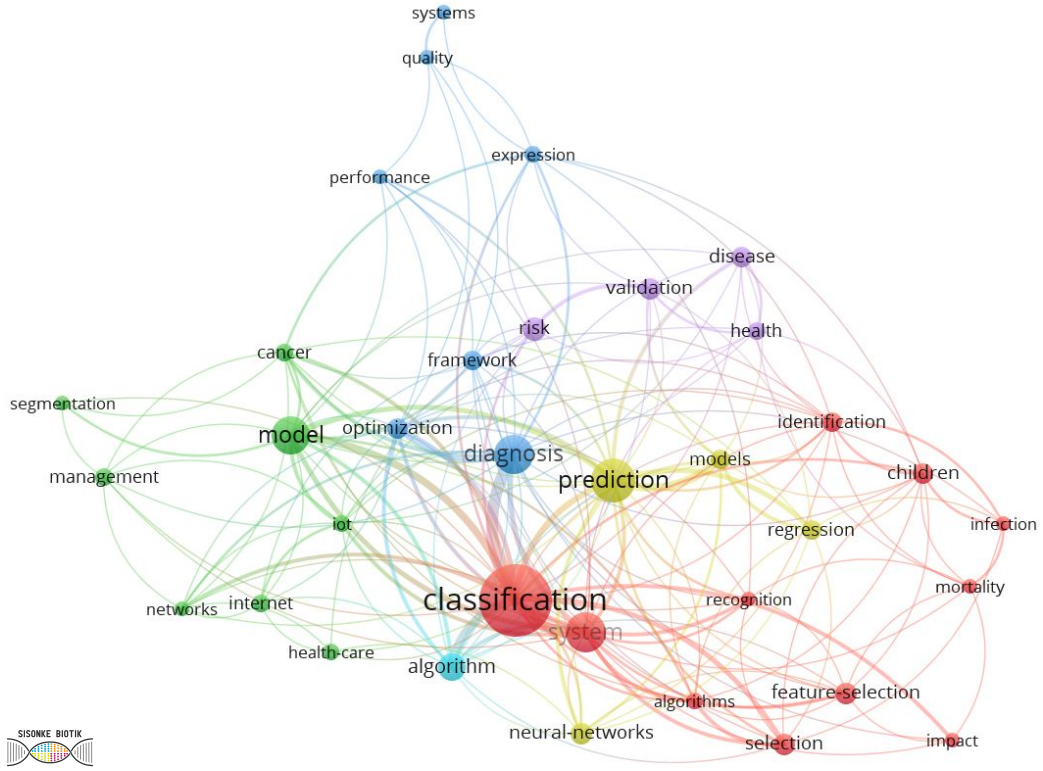
DEEP LEARNING INDABA

SISONKE BIOTIK

Research focus of Africa over time

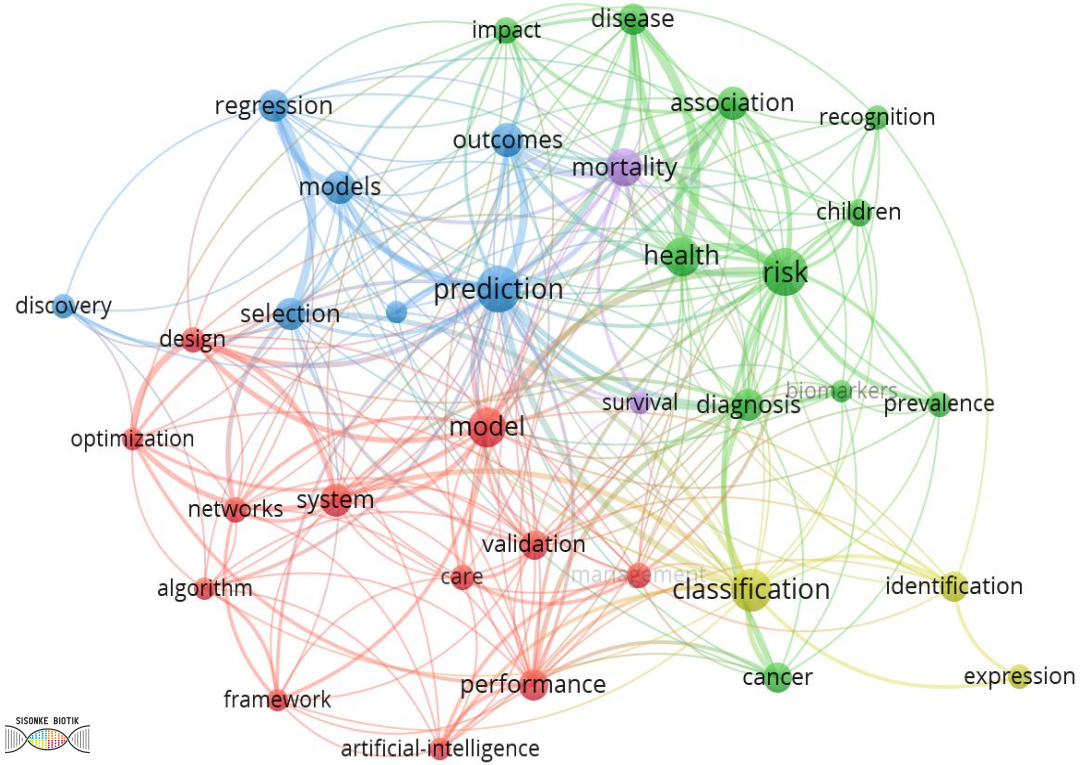
AFRICA

A less complex network for a continent



USA

A more complex network for just a single country

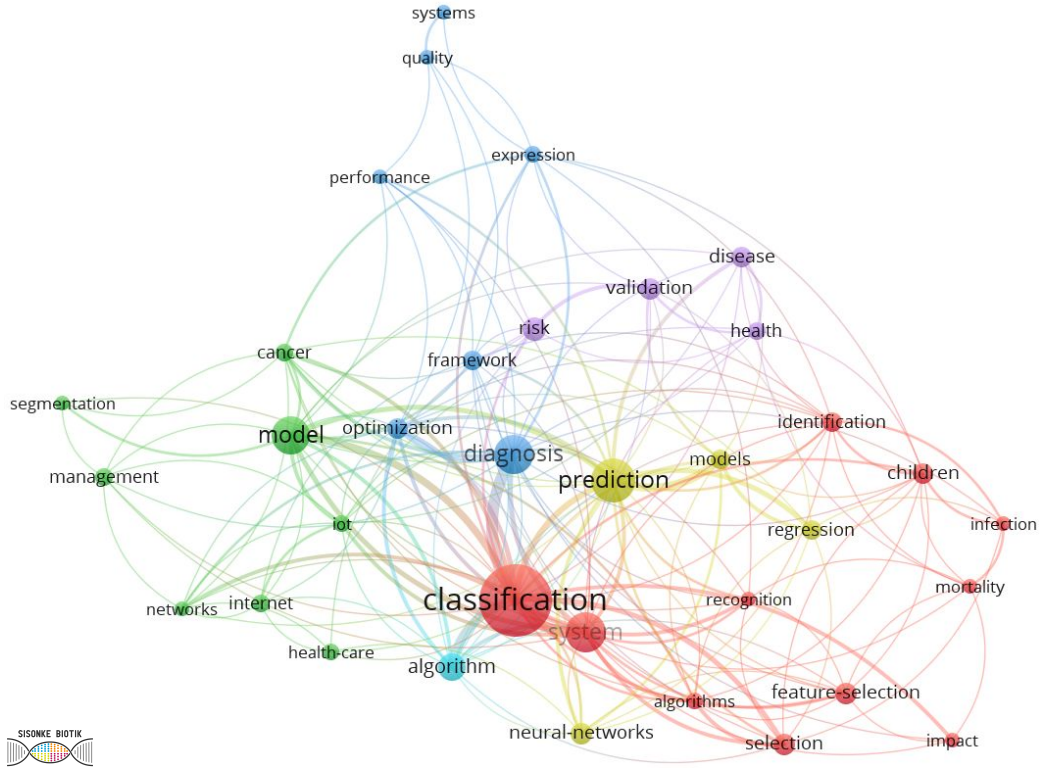


DEEP LEARNING INDABA

Research focus of Africa over time

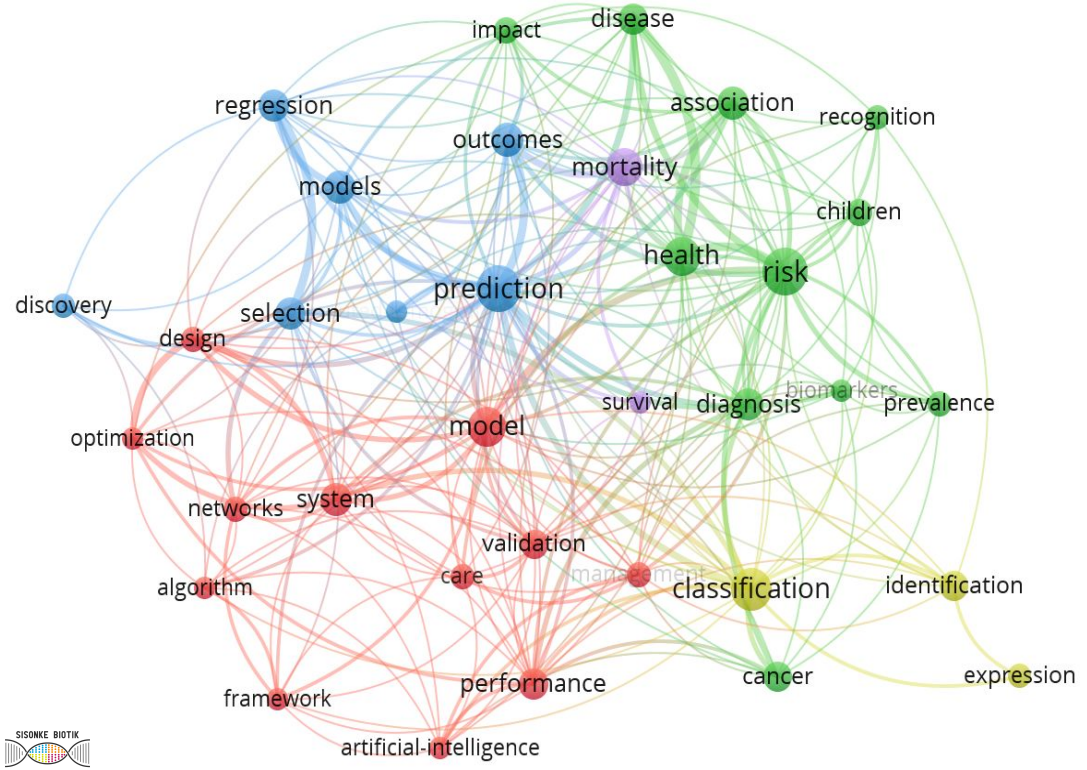
AFRICA

A less complex network for a continent



USA

A more complex network for just a single country

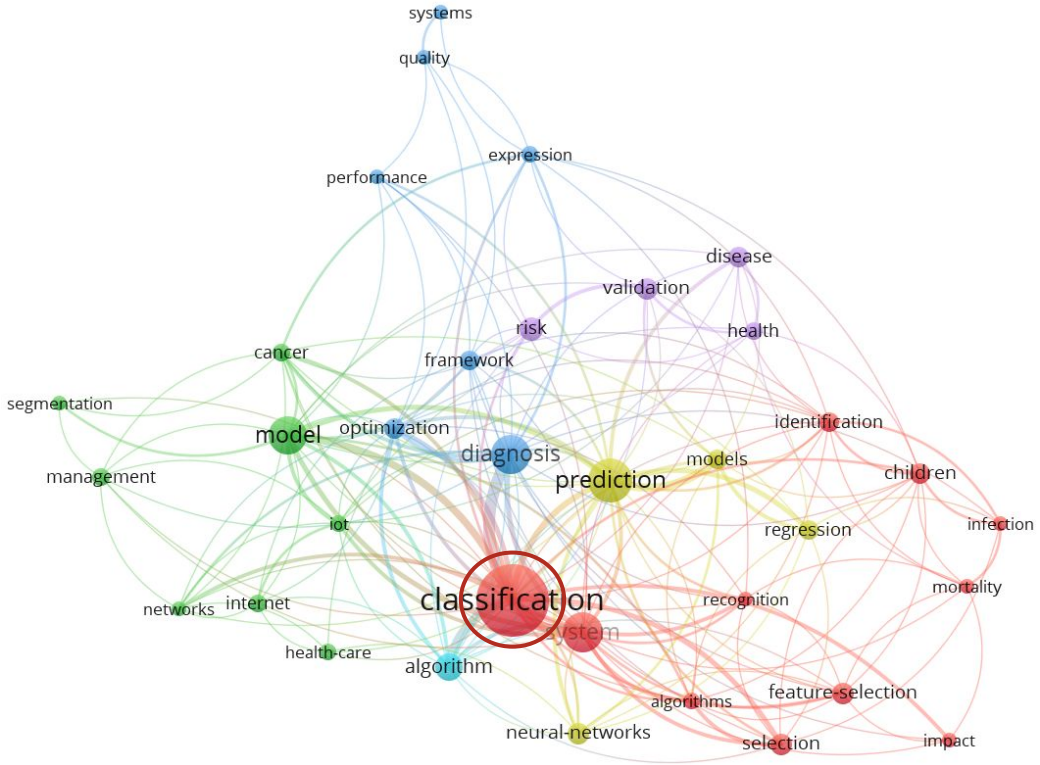


We need more hands on deck

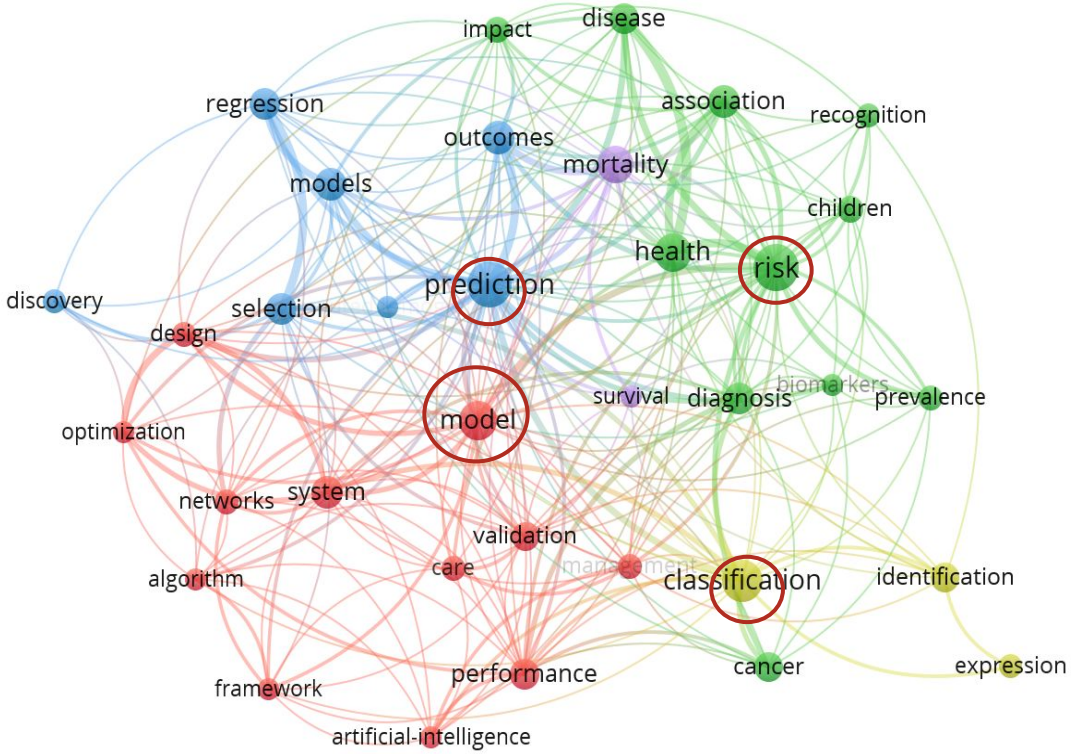
DEEP LEARNING INDABA

Research focus of Africa over time

AFRICA Works on “classification” dominates



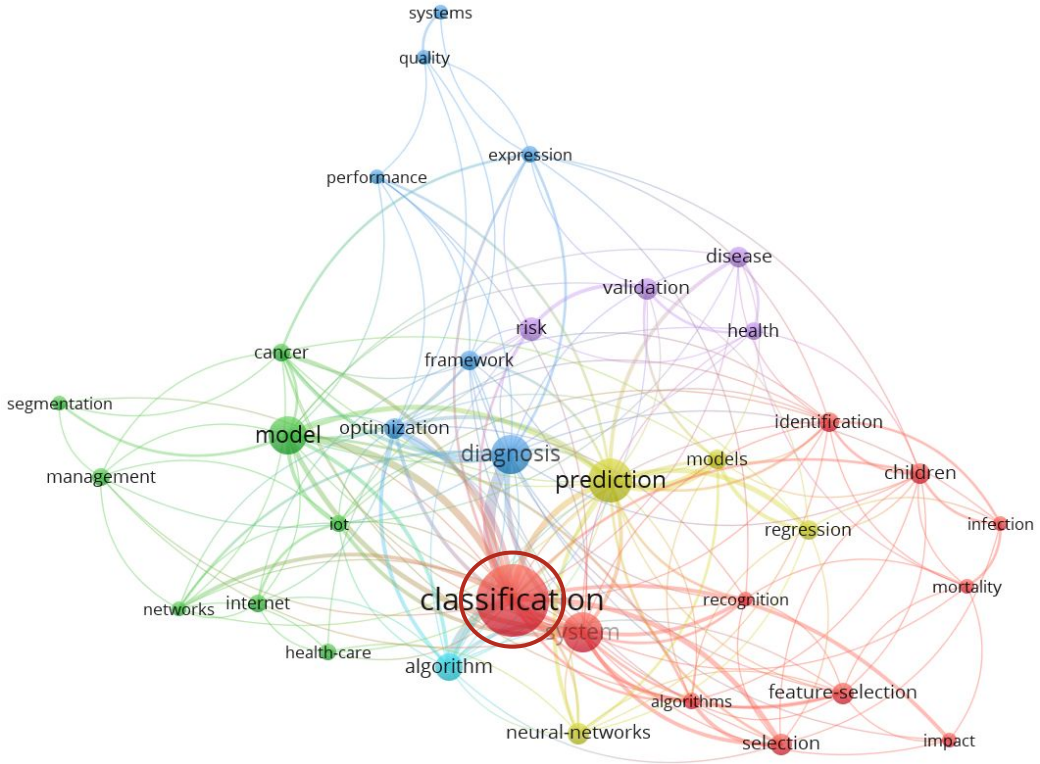
USA The focus is significantly distributed across classification, prediction, risk, model, selection, prediction



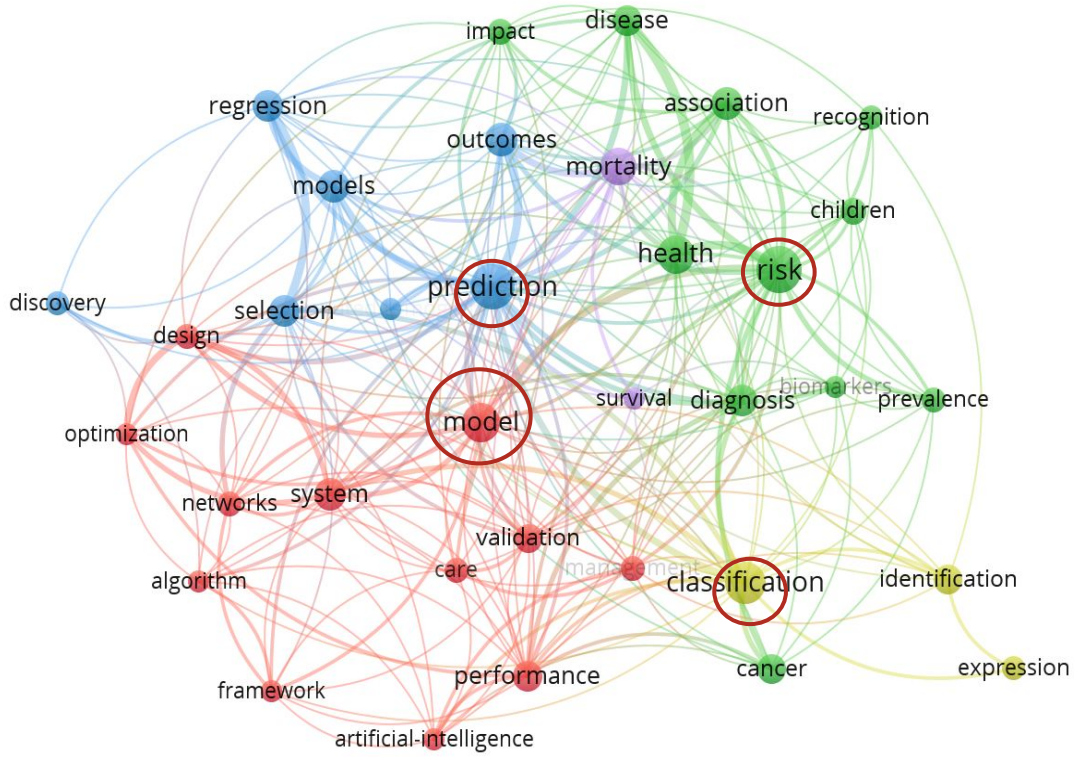
DEEP LEARNING INDABA

Research focus of Africa over time

AFRICA Works on "classification" dominates



USA The focus is significantly distributed across classification, prediction, risk, model, selection, prediction

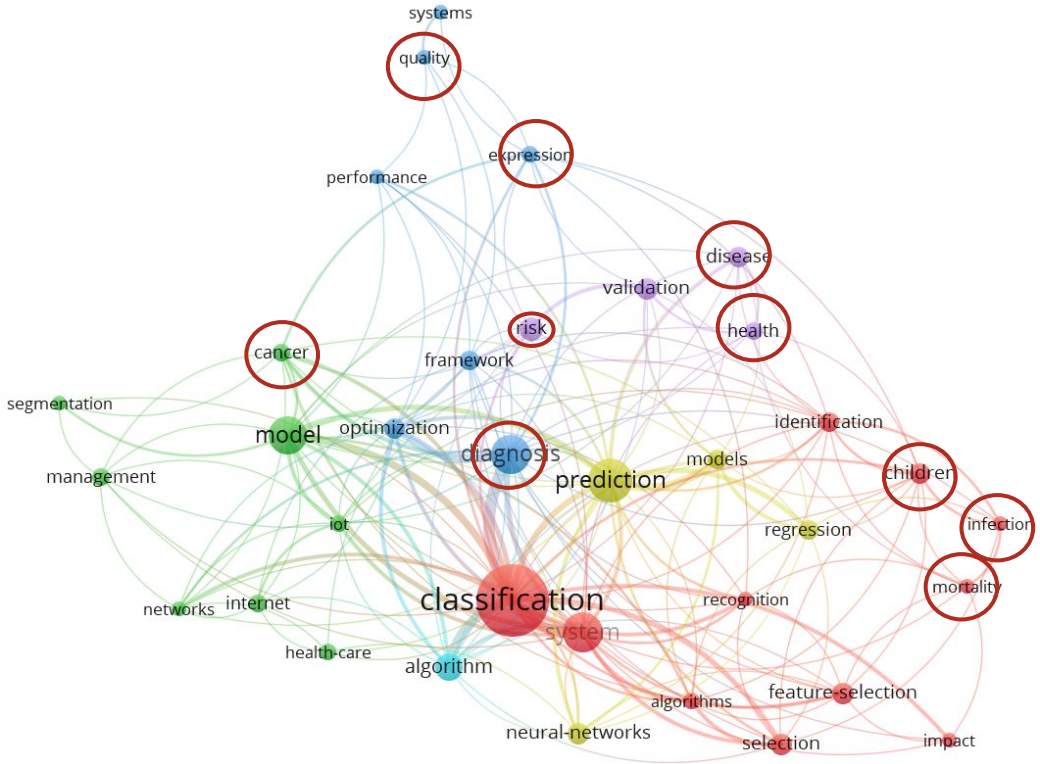


We need more than just tackling "classification" problems

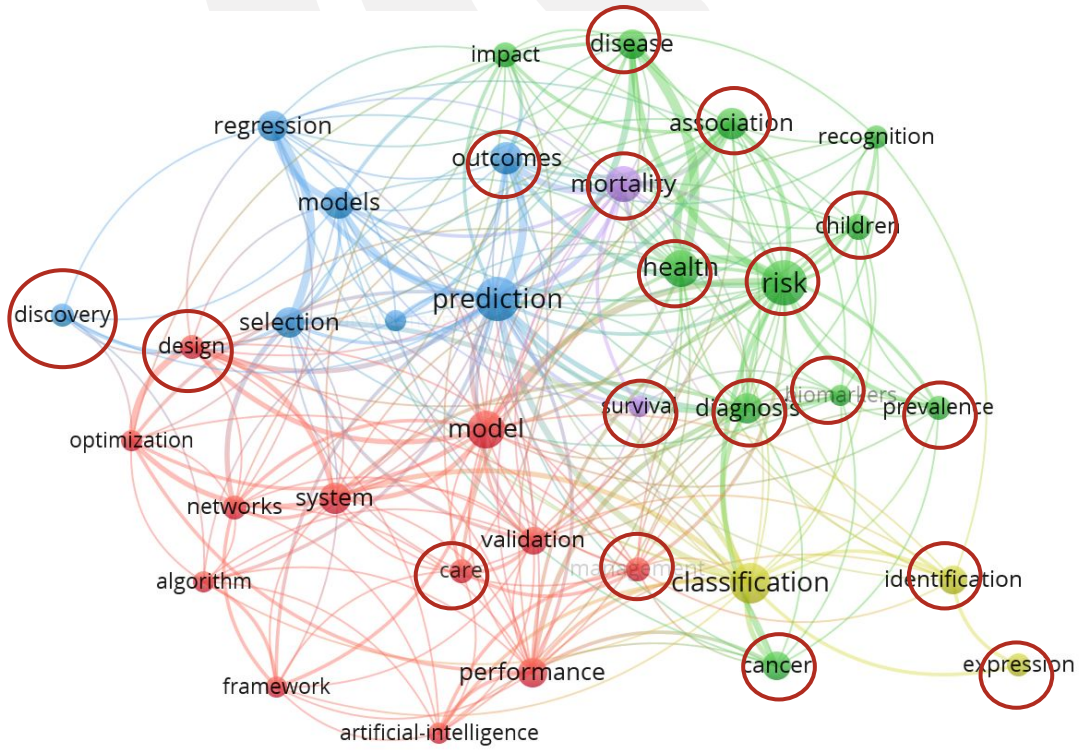
DEEP LEARNING INDABA

Research focus of Africa over time

AFRICA More machine learning keywords than health



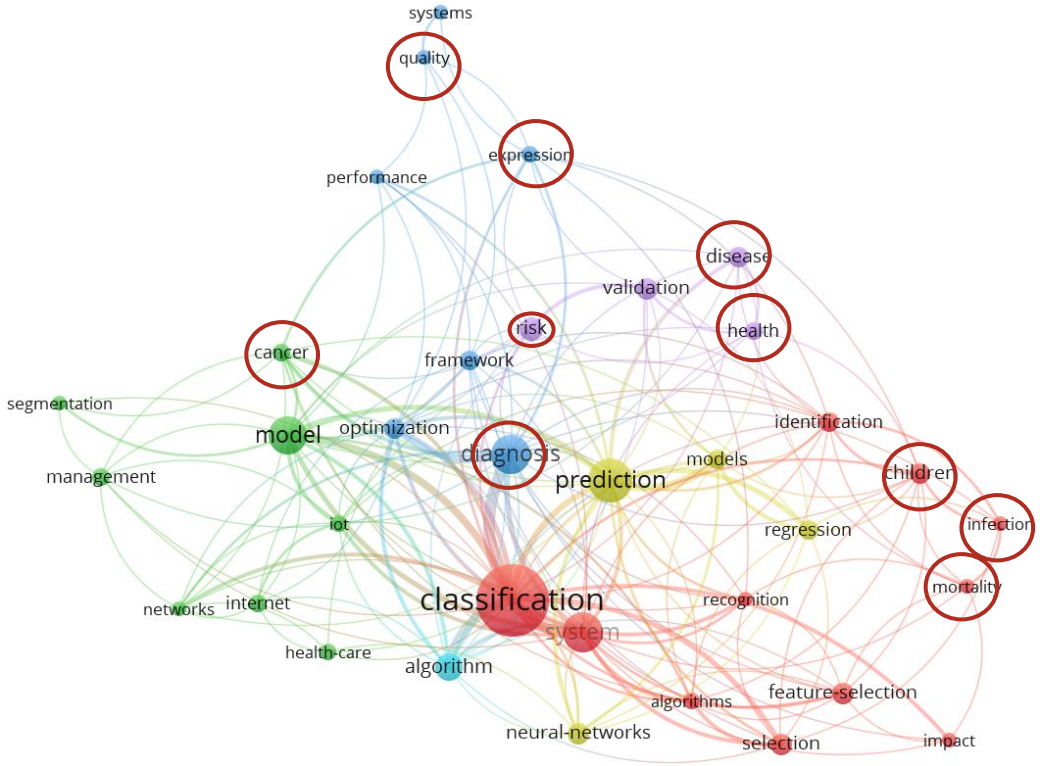
USA A fair proportion of health related keywords when compare to the machine learning keywords.



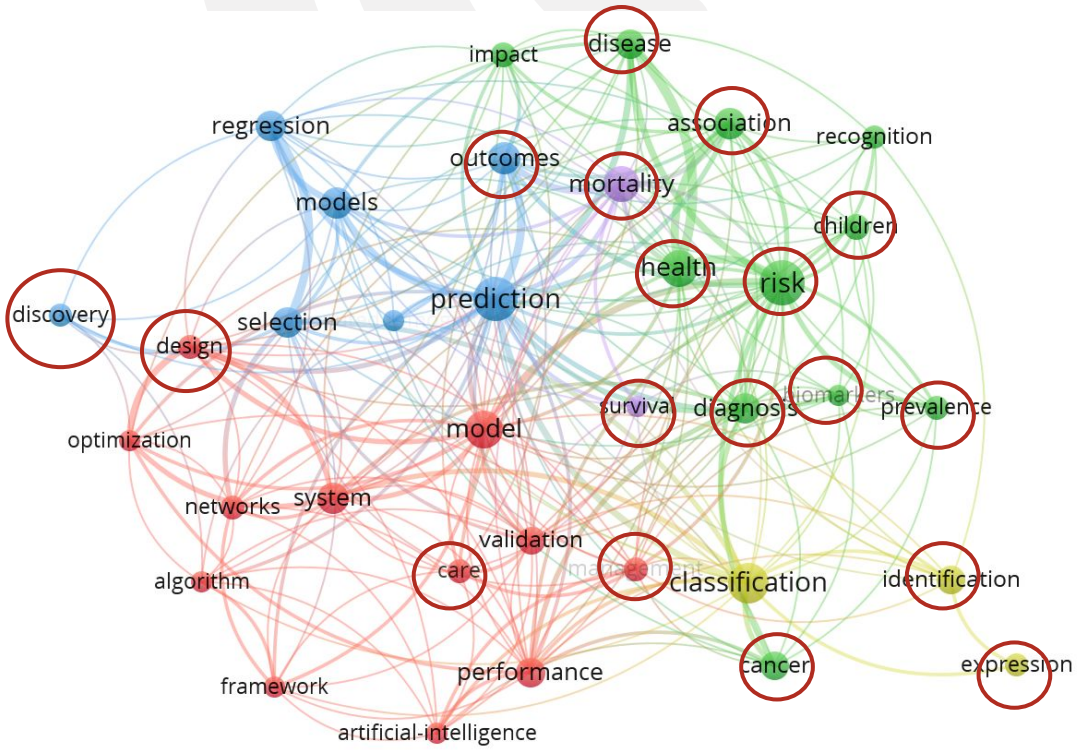
DEEP LEARNING INDABA

Research focus of Africa over time

AFRICA More machine learning keywords than health



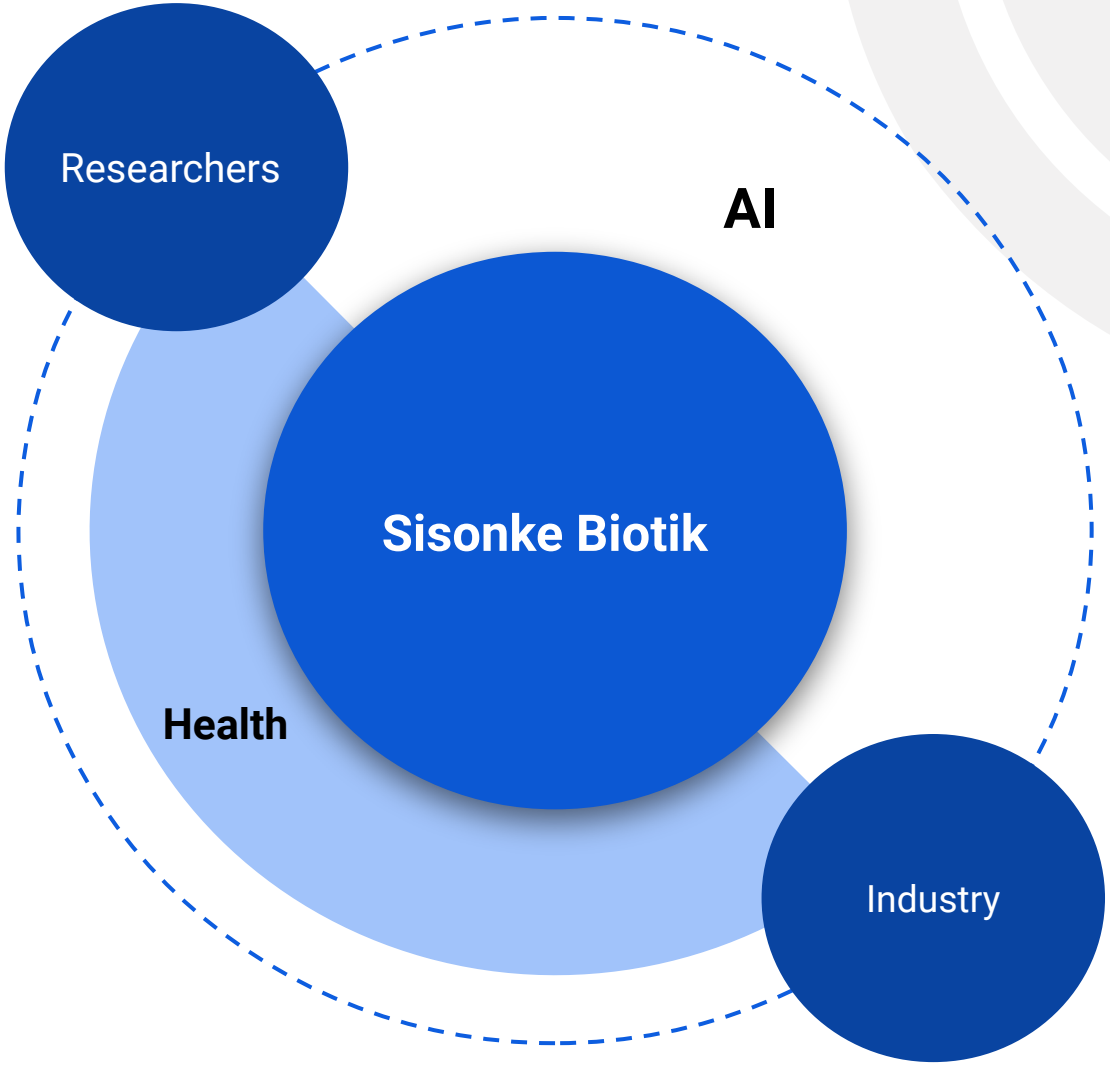
USA A fair proportion of health related keywords when compare to the machine learning keywords.



We need an increased diverse research in the health sector

DEEP LEARNING INDABA

OUR SOLUTION:



DEEP LEARNING INDABA

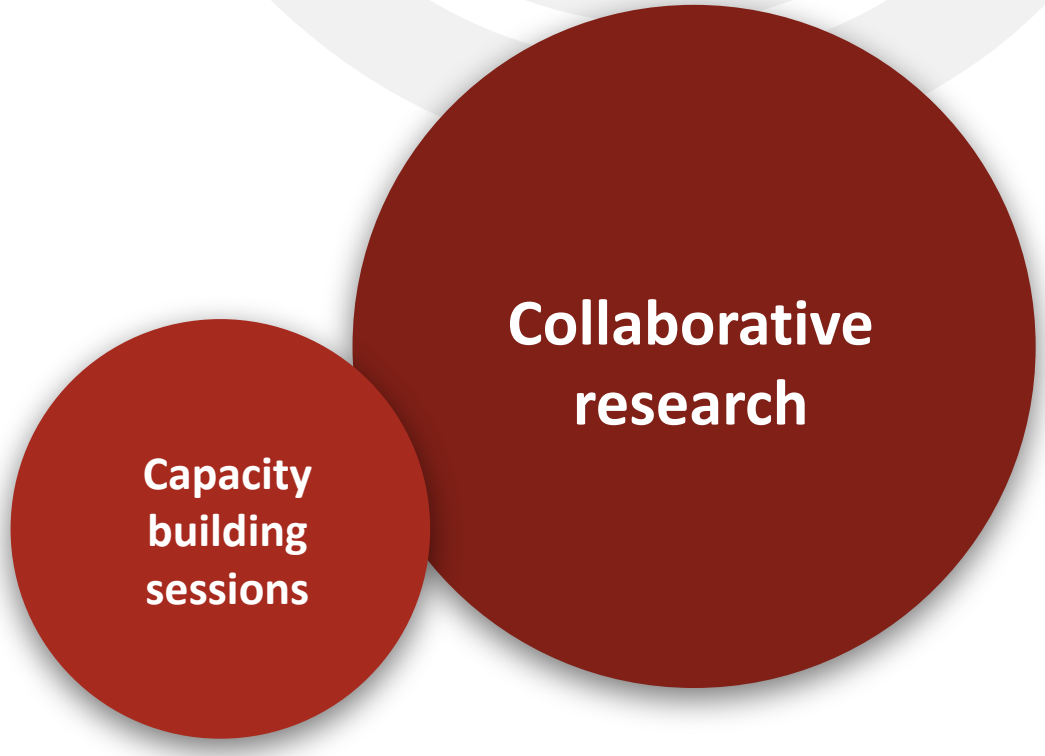
SISONKE BIOTIK

As a community we offer:

DEEP LEARNING INDABA

SISONKE BIOTIK

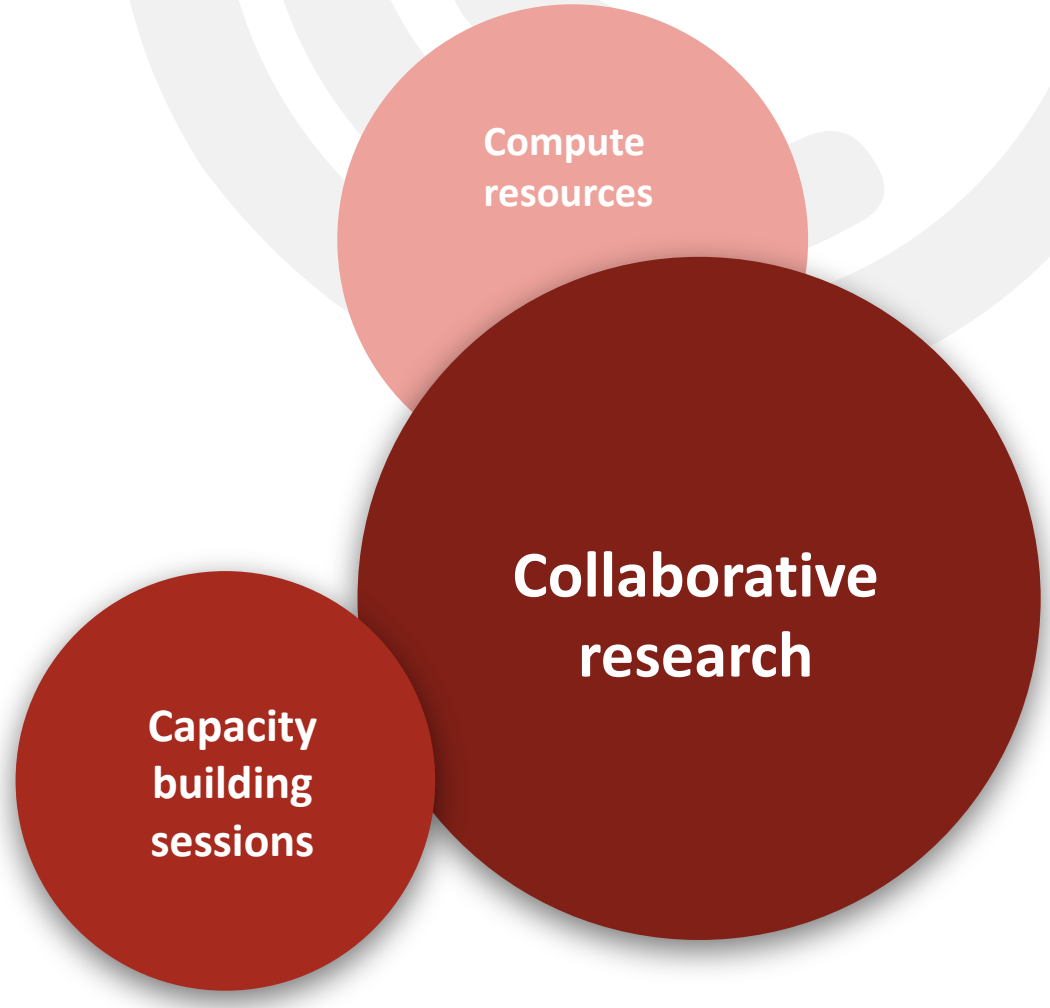
As a community we offer:



DEEP LEARNING INDABA

SISONKE BIOTIK

As a community we offer:



DEEP LEARNING INDABA

SISONKE BIOTIK

As a community we offer:

DEEP LEARNING INDABA

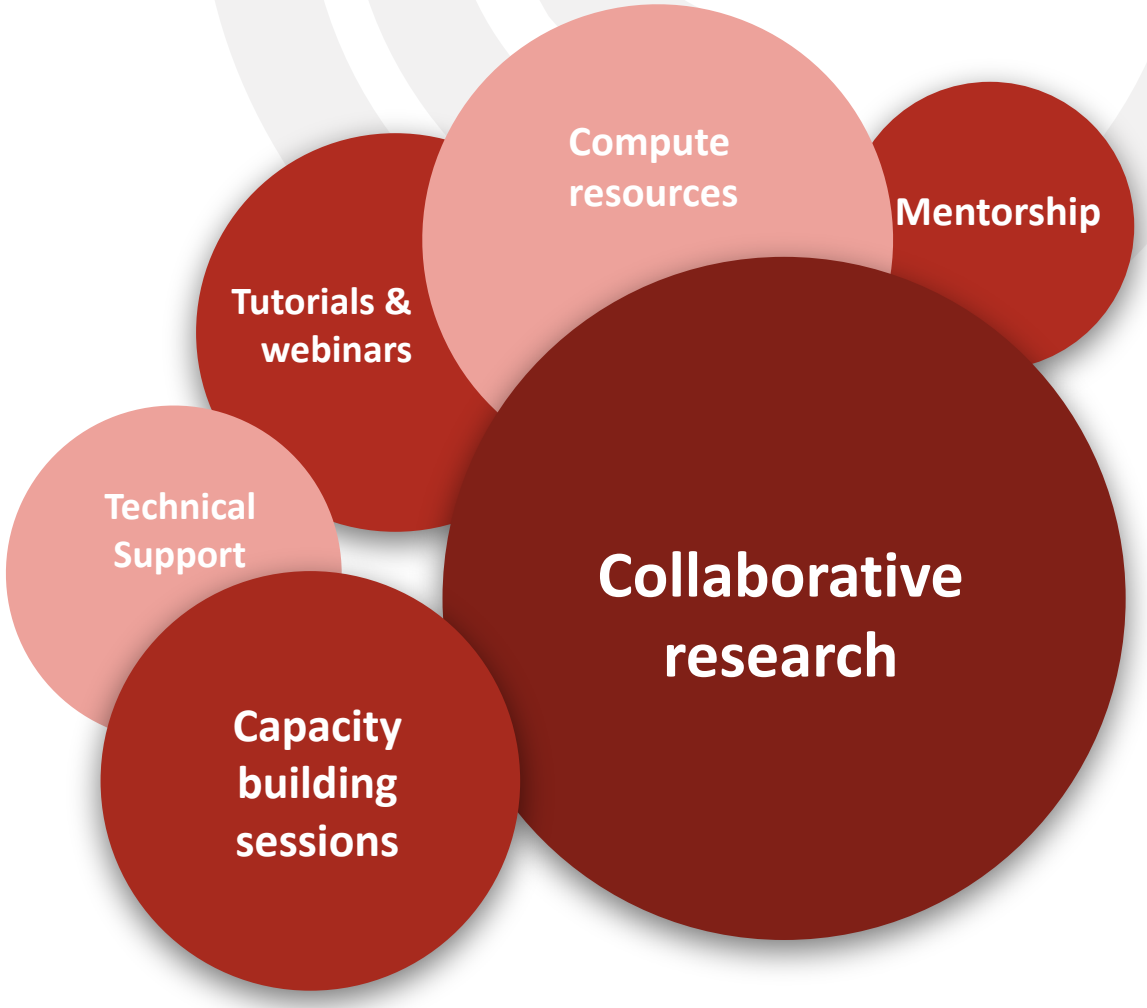
SISONKE BIOTIK

As a community we offer:

DEEP LEARNING INDABA

SISONKE BIOTIK

As a community we offer:



DEEP LEARNING INDABA

SISONKE BIOTIK

As a community we offer:

DEEP LEARNING INDABA

SISONKE BIOTIK

As a community we offer:

DEEP LEARNING INDABA

PROJECT HIGHLIGHT

Biomedical knowledge engineering

Drug discovery

DEEP LEARNING INDABA

KNOWLEDGE ENGINEERING

AN INTERESTING FIELD OF BIOMEDICAL MACHINE LEARNING

DATA ENGINEERING AND SEMANTICS

- Created in 2021
- First in Tunisia specialized in Wikimedia projects
- Affiliated with Faculty of Sciences, University of Sfax, Tunisia
- **Main objective:** develop novel applications of Wikimedia Projects based on Knowledge Engineering, Machine Learning, and Big Data Technologies

DEEP LEARNING INDABA

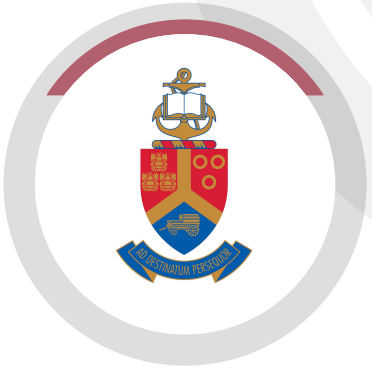
MAIN COLLABORATORS

WIKIMEDIA TUNISIA

TUNISIA

SISONKEBIOTIK

SOUTH AFRICA



UNIVERSITY OF PRETORIA

SOUTH AFRICA

WIKIMEDIA MEDICINE

UNITED STATES OF AMERICA

NEW RESEARCH PROJECT LAUNCHED 🤩

"Adapting Wikidata to support clinical practice using Data Science, Semantic Web and Machine Learning"

WIKIMEDIA FOUNDATION

- Part of a project funded by the Wikimedia Research Fund
- Launching in August 2022 - duration 1 year

DEEP LEARNING INDABA

BIOMEDICAL KNOWLEDGE ENGINEERING

Information Retrieval

BIOMEDICAL KNOWLEDGE ENGINEERING

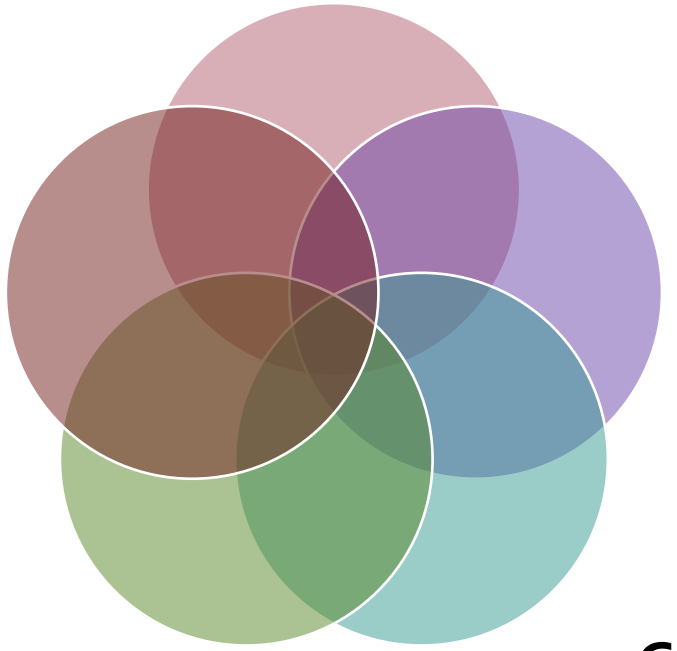
BIOMEDICAL KNOWLEDGE ENGINEERING

BIOMEDICAL KNOWLEDGE ENGINEERING

BIOMEDICAL KNOWLEDGE ENGINEERING

Named Entity Recognition

Reference Identification



Relation Extraction

Relation Validation

Relation Classification

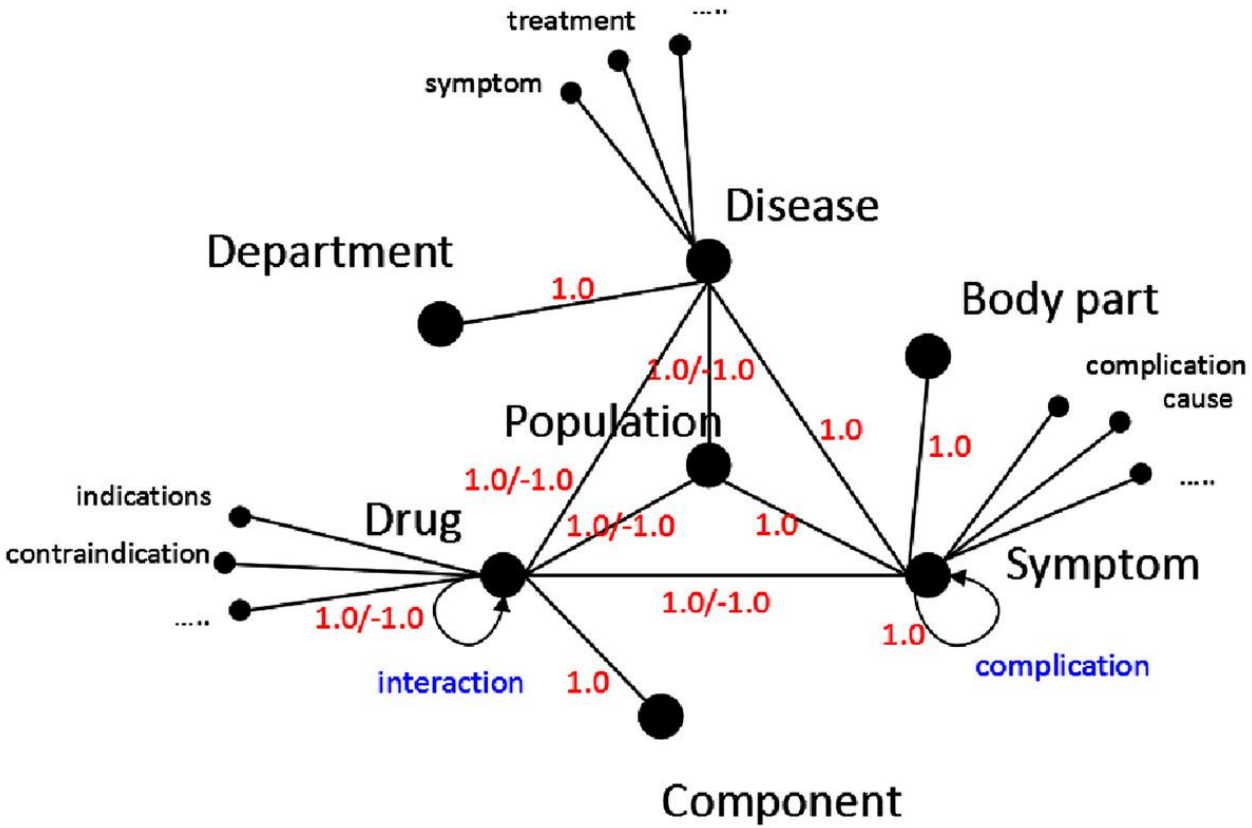
MAIN TASKS

KNOWLEDGE GRAPHS

- **Concepts represented as items**:
  Diseases, drugs, treatments, muscles..

- **Statements in the form of triples**:
  Subject – Predicate – Object

- **Human-Friendly and Machine-Readable Data Serialization Formats**:
  - *Web Standards*: RDF, FHIR, OWL
  - *File Formats*: XML, JSON, TTL



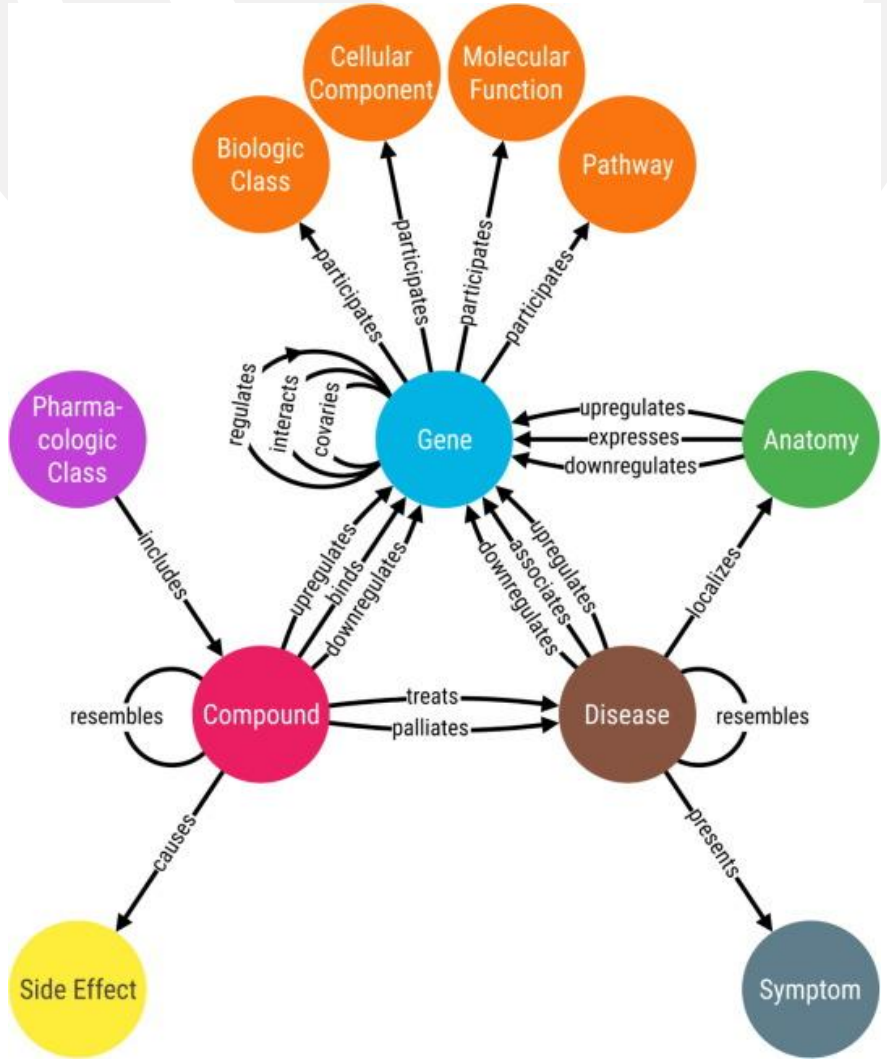
WHY WE NEED BIOMEDICAL KNOWLEDGE GRAPHS

DRIVING KNOWLEDGE-BASED SYSTEMS

- Recommender Systems
- Question Answering
- Biomedical Entity Classification
- Clinical Decision Support

- Preventing cold start
- Validating machine learning
- Supporting medical practitioners

Automation of knowledge engineering is needed because of the lack of support of African Medicine in such resources



MeSH2Matrix: Machine learning-driven biomedical relation classification based on the MeSH keywords of PubMed scholarly publications

Houcemeddine Turki, **Bonaventure F. P. Dossou**, **Chris Chinenye Emezue**, Mohamed Ali Hadj Taieb, Mohamed Ben Aouicha, Hanen Ben Hassen, Afif Masmoudi

Proceedings of the 12th International Workshop on Bibliometric-enhanced Information Retrieval co-located with 44th European Conference on Information Retrieval (ECIR 2022)

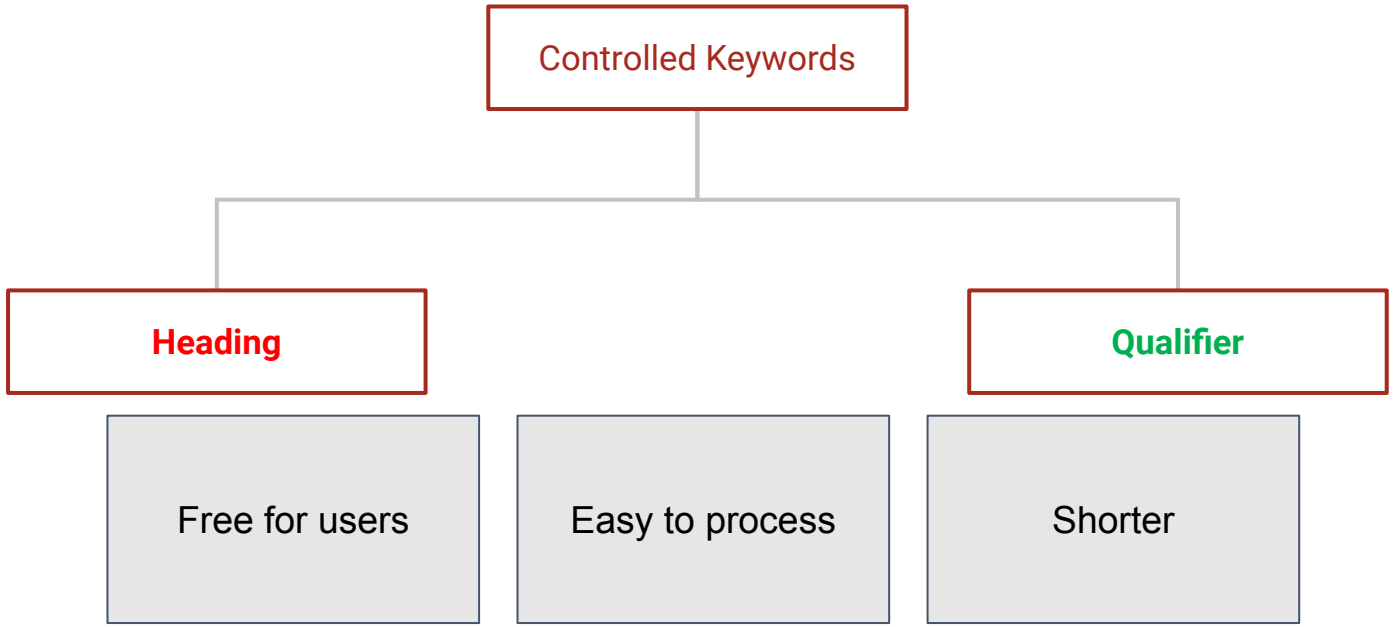
Data Engineering and Semantics
هندسة البيانات و دلالاتها

JACOBS UNIVERSITY

Technische Universität München

University of Sfax

MESH KEYWORDS



PubMed

+150k items
89 qualifiers

Ledipasvir/Sofosbuvir: a review of its use in chronic hepatitis C

Gillian M Keating [1]

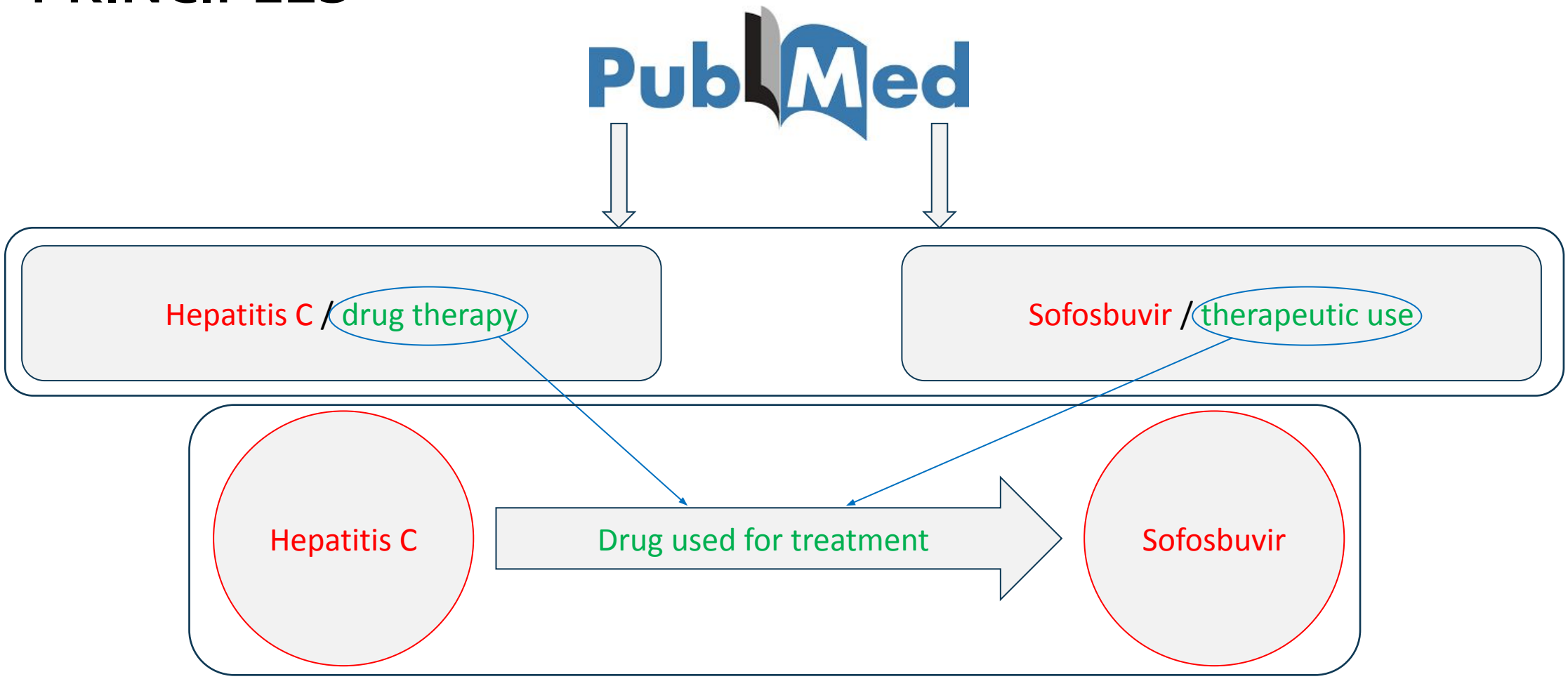
1 Springer, Private Bag 65901, Mairangi Bay 0754, Auckland, New Zealand, demail@springer.com.

MeSH terms

- Antiviral Agents / administration & dosage
- Antiviral Agents / pharmacokinetics
- Antiviral Agents / therapeutic use*
- Benzimidazoles / administration & dosage
- Benzimidazoles / pharmacokinetics
- Benzimidazoles / therapeutic use*
- Fluorenes / administration & dosage

DEEP LEARNING INDABA

PRINCIPLES

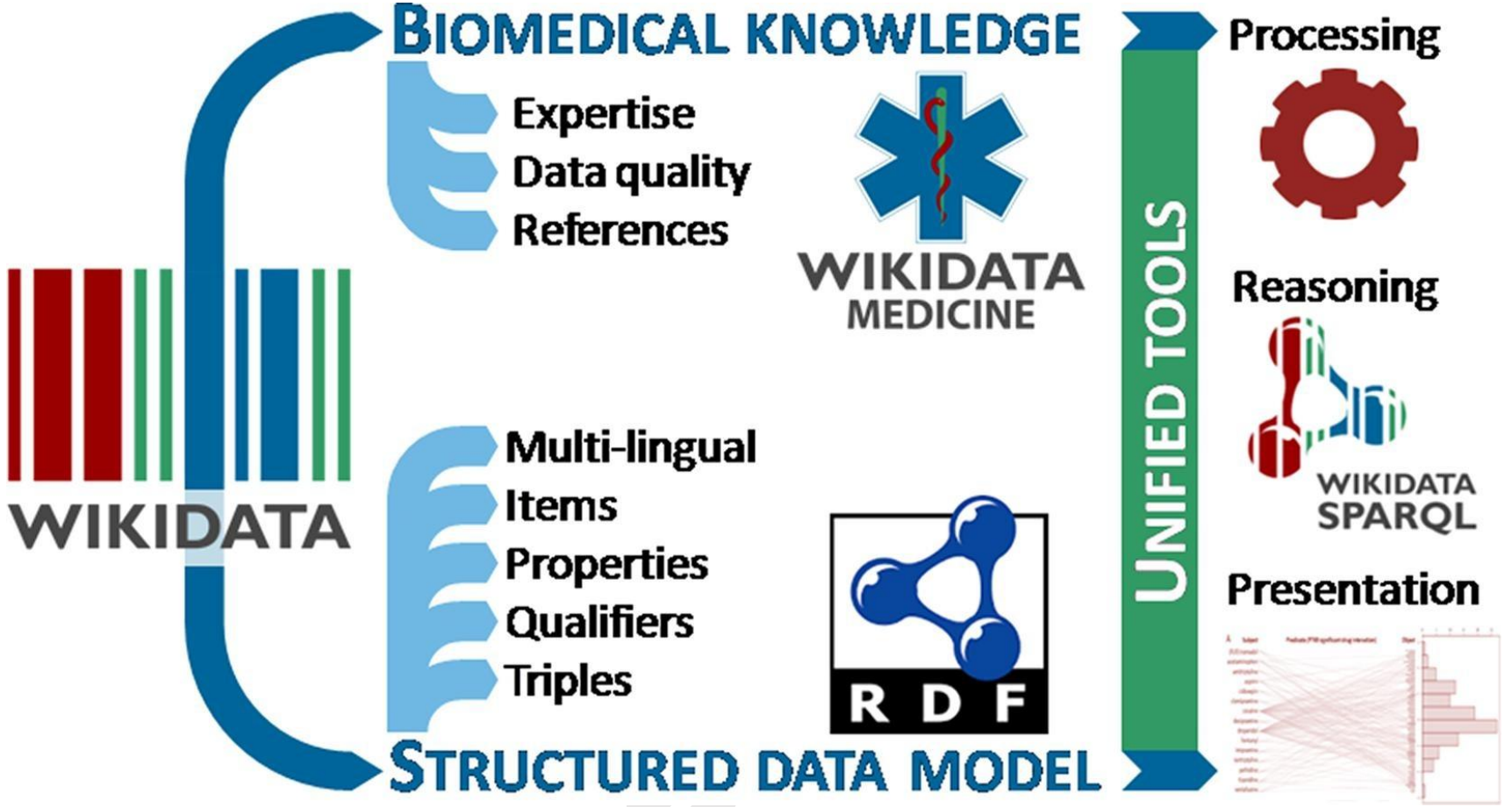

PubMed
Hepatitis C / drug therapy
Sofosbuvir / therapeutic use
Hepatitis C
Drug used for treatment
Sofosbuvir
WIKIDATA

WE NEED A DATASET OF BIOMEDICAL RELATIONS

« Wikidata can provide such relations as a multidisciplinary open knowledge graph »

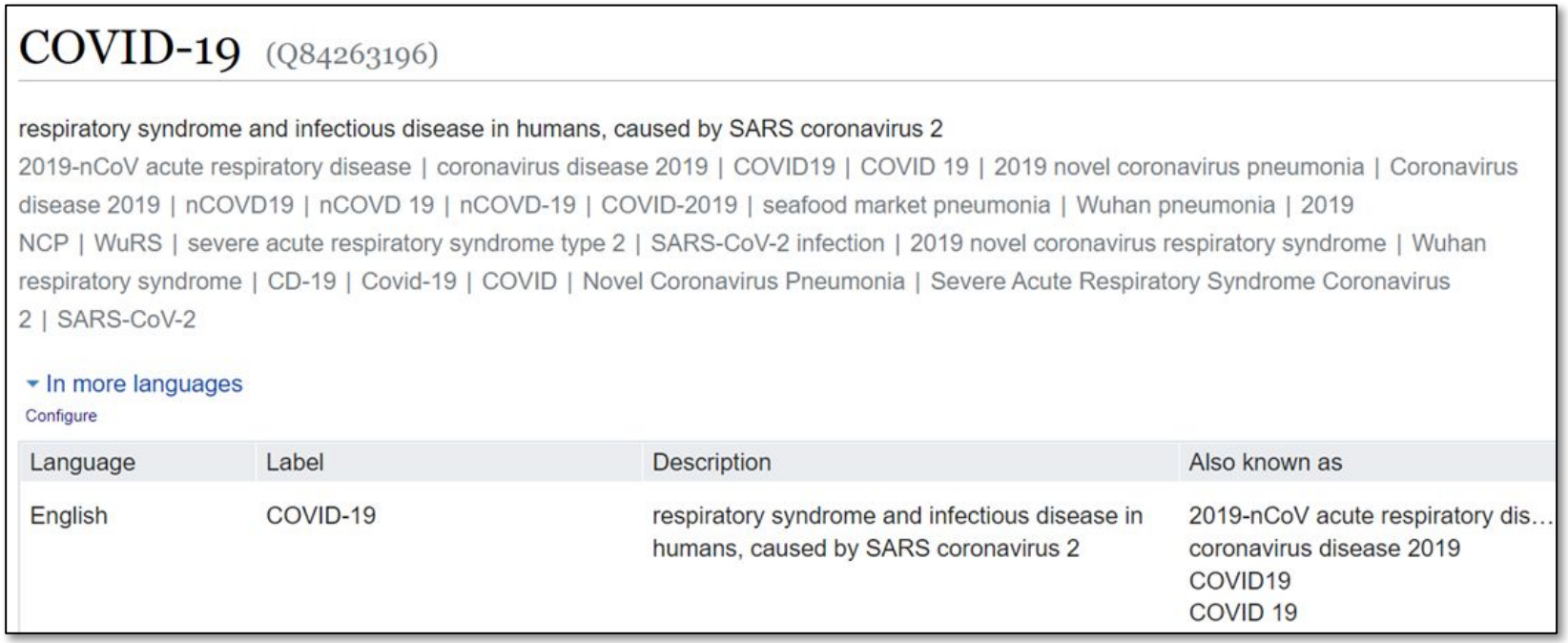
WIKIDATA

- **F**indable
- **A**ccessible
- **I**nteroperable
- **R**eusable



Easily available at https://www.wikidata.org.

WIKIDATA



COVID-19 (Q84263196)

respiratory syndrome and infectious disease in humans, caused by SARS coronavirus 2

2019-nCoV acute respiratory disease | coronavirus disease 2019 | COVID19 | COVID 19 | 2019 novel coronavirus pneumonia | Coronavirus disease 2019 | nCOVD19 | nCOVD 19 | nCOVD-19 | COVID-2019 | seafood market pneumonia | Wuhan pneumonia | 2019 NCP | WuRS | severe acute respiratory syndrome type 2 | SARS-CoV-2 infection | 2019 novel coronavirus respiratory syndrome | Wuhan respiratory syndrome | CD-19 | Covid-19 | COVID | Novel Coronavirus Pneumonia | Severe Acute Respiratory Syndrome Coronavirus 2 | SARS-CoV-2

▾ In more languages

Configure

Language	Label	Description	Also known as
English	COVID-19	respiratory syndrome and infectious disease in humans, caused by SARS coronavirus 2	2019-nCoV acute respiratory dis… coronavirus disease 2019 COVID19 COVID 19

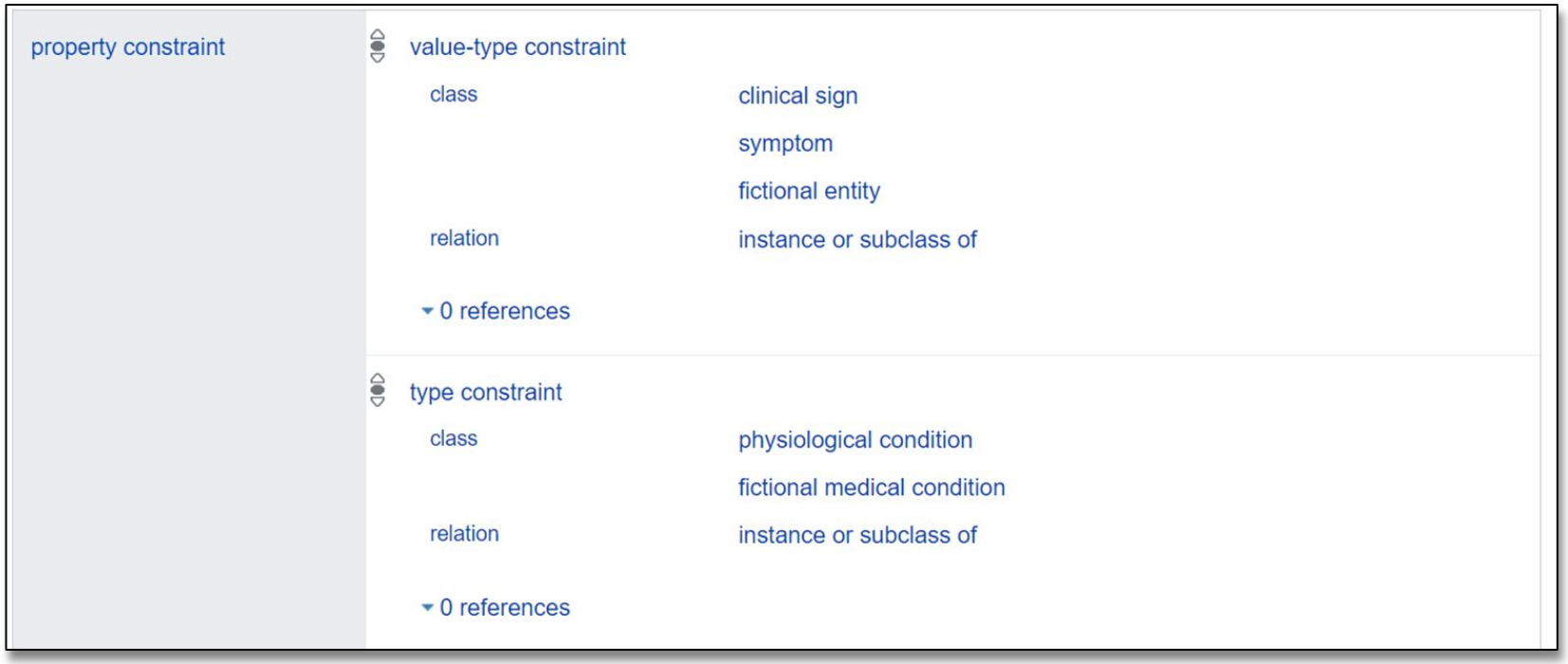
- **Concepts assigned labels, descriptions and aliases in multiple languages**
- Taxonomic relations (e.g., instance of)
- Non-Taxonomic relations (e.g., Symptoms and signs)
- Property constraints
- Aligned to MeSH Terms

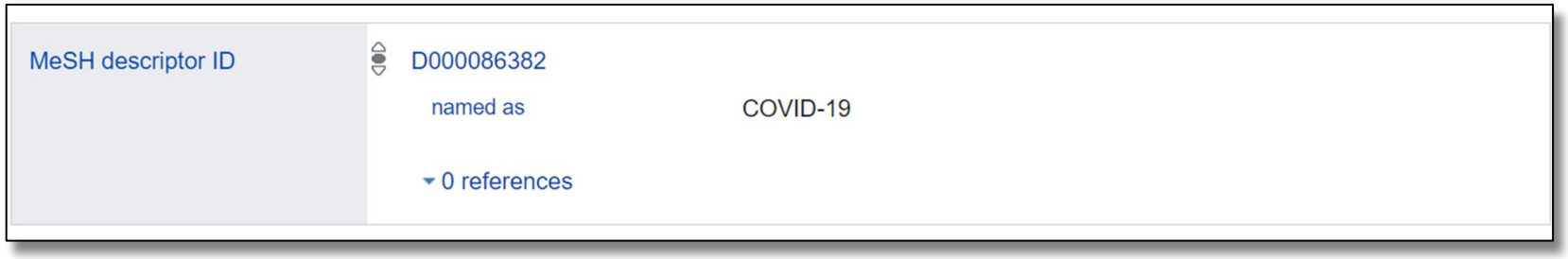
DEEP LEARNING INDABA

WIKIDATA

- Concepts assigned labels, descriptions and aliases in multiple languages
- **Taxonomic relations (e.g., instance of)**
- **Non-Taxonomic relations (e.g., Symptoms and signs)**
- Property constraints
- Aligned to MeSH Terms

WIKIDATA



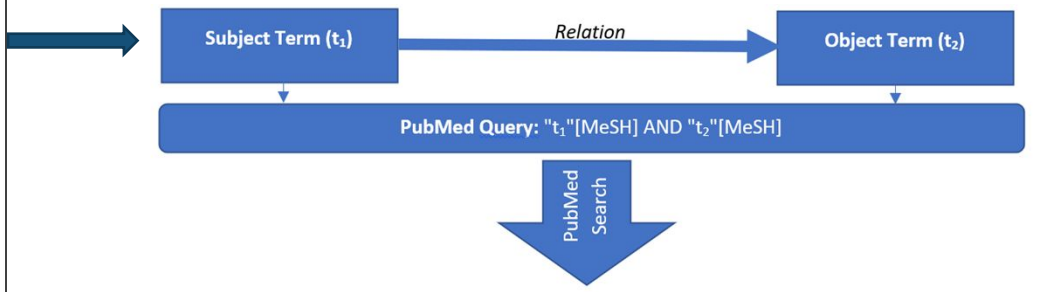


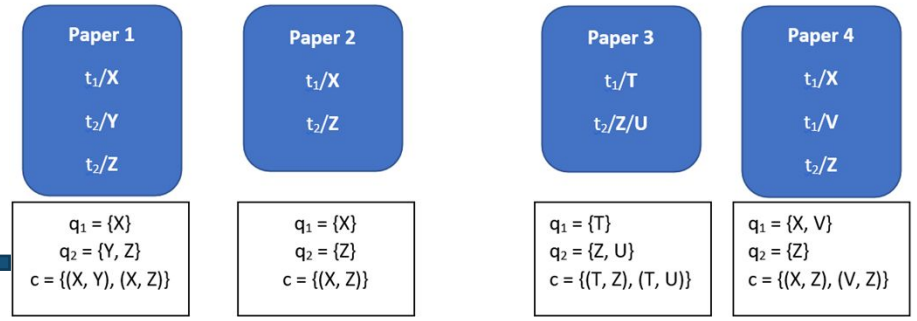
- Concepts assigned labels, descriptions and aliases in multiple languages
- Taxonomic relations (e.g., instance of)
- Non-Taxonomic relations (e.g., Symptoms and signs)
- **Property constraints**
- **Aligned to MeSH Terms**

DEEP LEARNING INDABA

```
SELECT ?subject ?reltype ?object WITH {
  SELECT * WHERE {
    ?item wdt:P486 ?subject.
    }
  }
AS %item
WHERE {
  INCLUDE %item.
  ?item ?reltype ?item1.
  ?item1 wdt:P486 ?object.
}
LIMIT 81000
```

WIKIDATA





Up to 100 Publications

	T	U	V	X	Y	Z
T	0	0	0	0	0	0
U	0.25	0	0	0	0	0
V	0	0	0	0	0	0
X	0	0	0	0	0	0
Y	0	0	0	0.25	0	0
Z	0.5	0	0	0.75	0	0

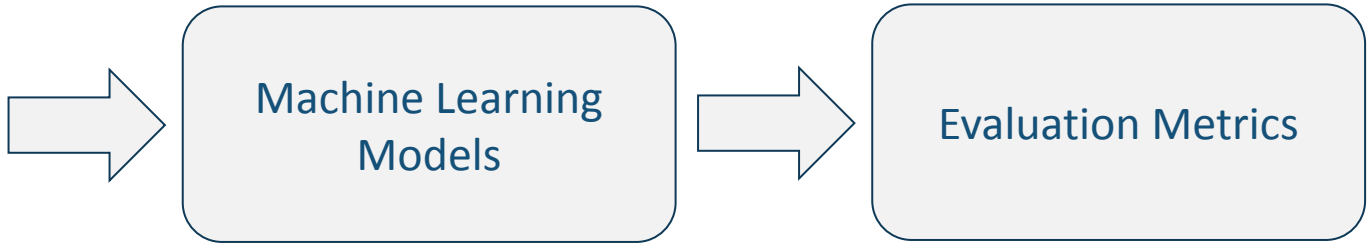
Relation

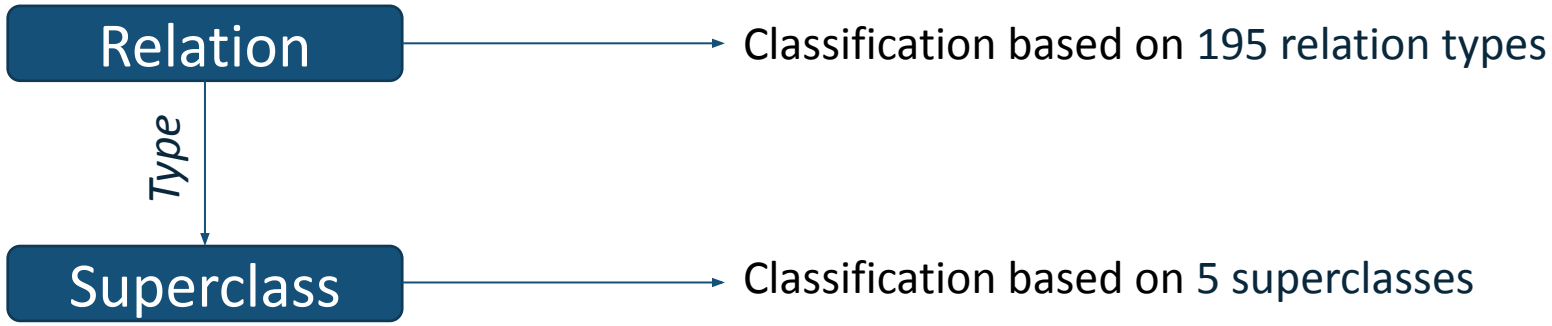
Storage in *MeSH2Matrix* Dataset

DEEP LEARNING INDABA

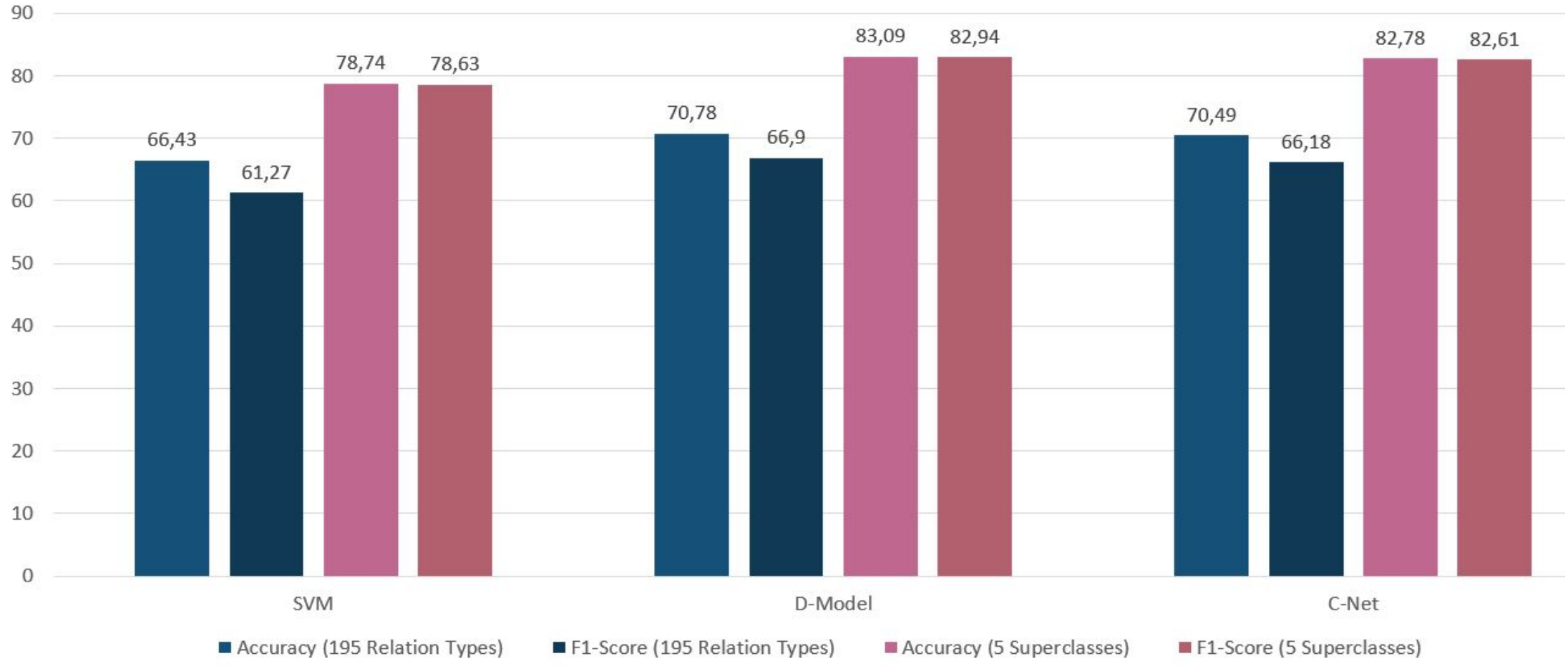
BIOMEDICAL RELATION CLASSIFICATION

	T	U	V	X	Y	Z
T	0	0	0	0	0	0
U	0.25	0	0	0	0	0
V	0	0	0	0	0	0
X	0	0	0	0	0	0
Y	0	0	0	0.25	0	0
Z	0.5	0	0	0.75	0	0





BIOMEDICAL RELATION CLASSIFICATION



DEEP LEARNING INDABA

https://github.com/SisonkeBiotik-Africa/MeSH2Matrix

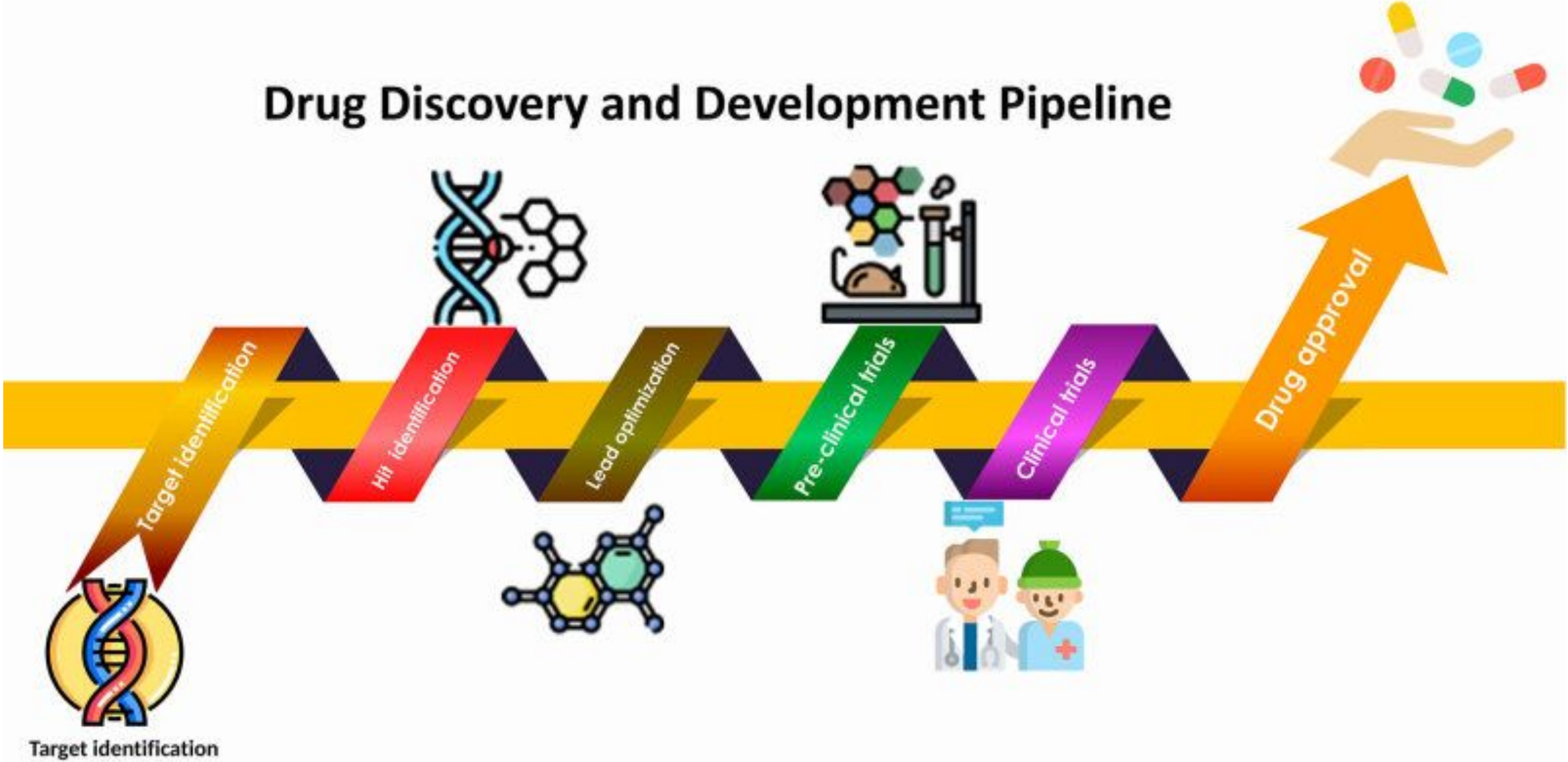
DEEP LEARNING INDABA

Machine Learning for Drug Discovery

A PATH FOR MOLECULAR HEALTH SUPPORT

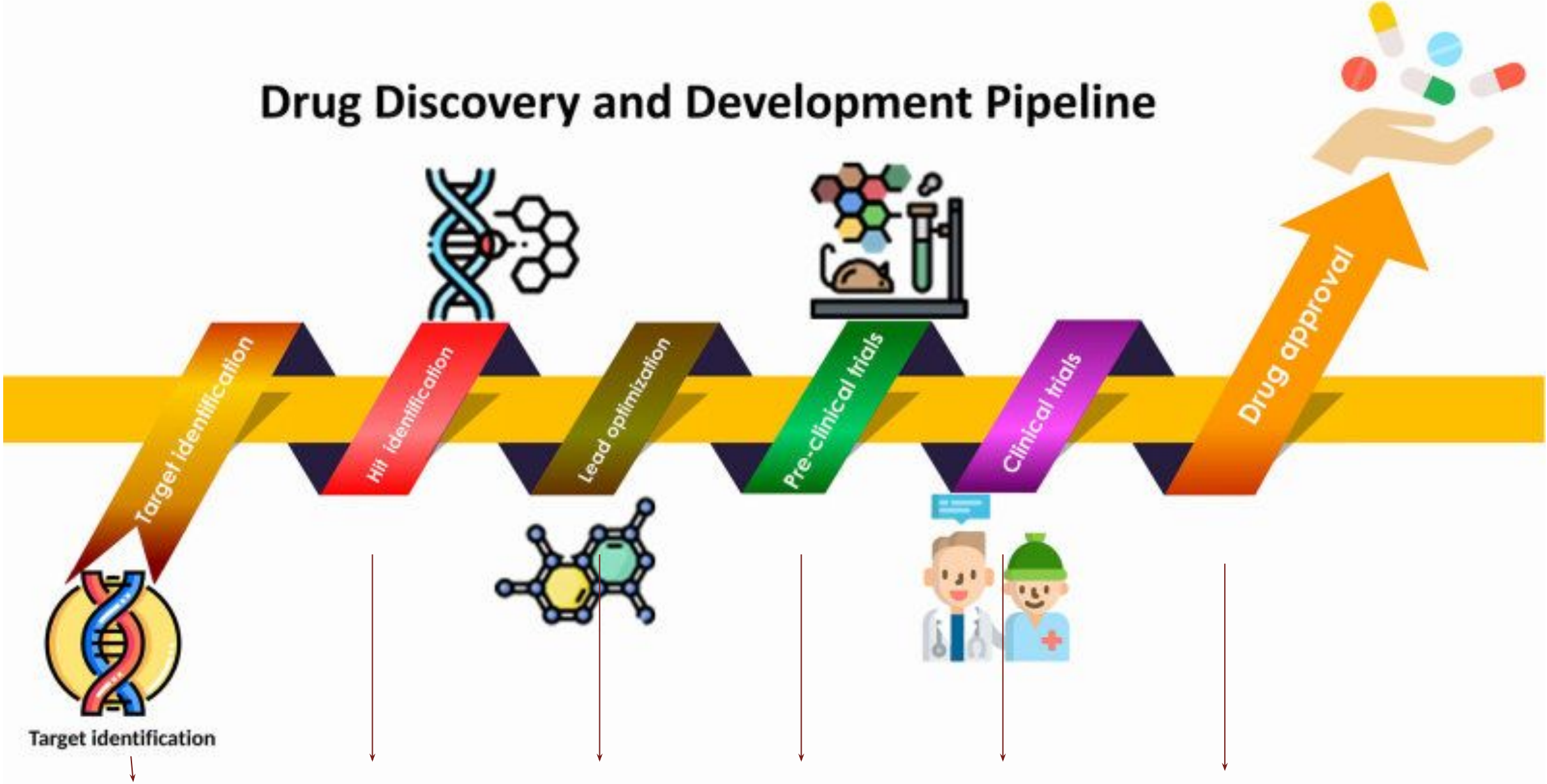
SPEAKER: Bonaventure Dossou, Researcher at Mila

Drug Discovery and Development Pipeline



Target identification

Drug Discovery and Development Pipeline



Drug Discovery and Development Pipeline

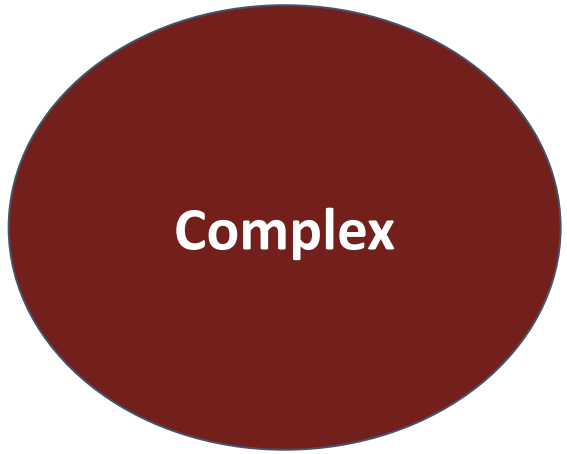
Target identification

Guessing Time: What is **the problem** with this pipeline?

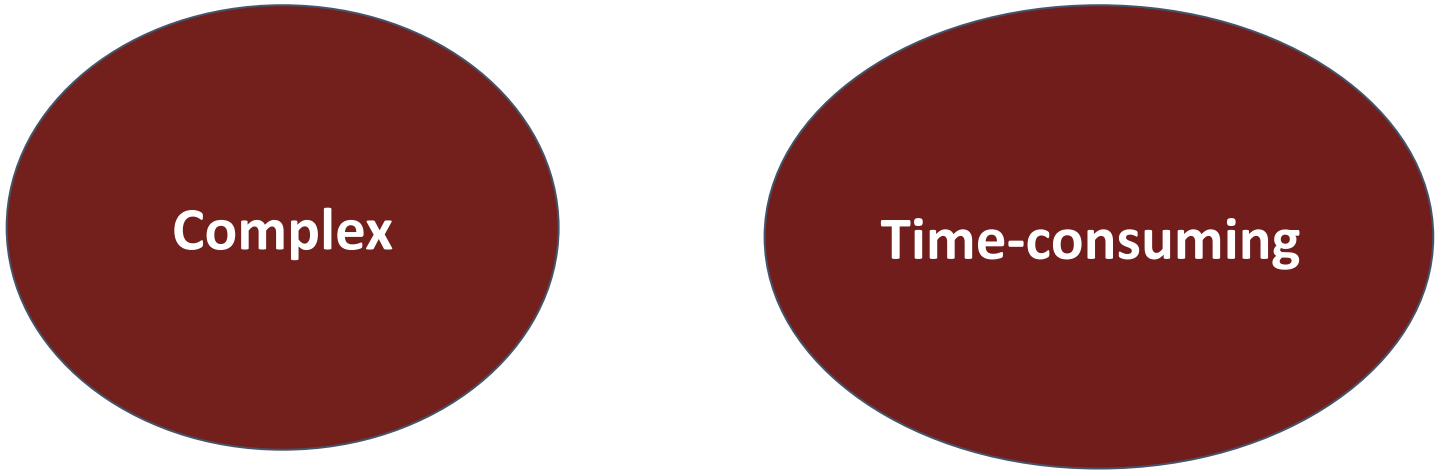
2-5yrs	0.5-1yr	1-3yr	1-2yrs	5-6yrs	1-2yrs

Modern Drug Discovery Limitations

Modern Drug Discovery Limitations

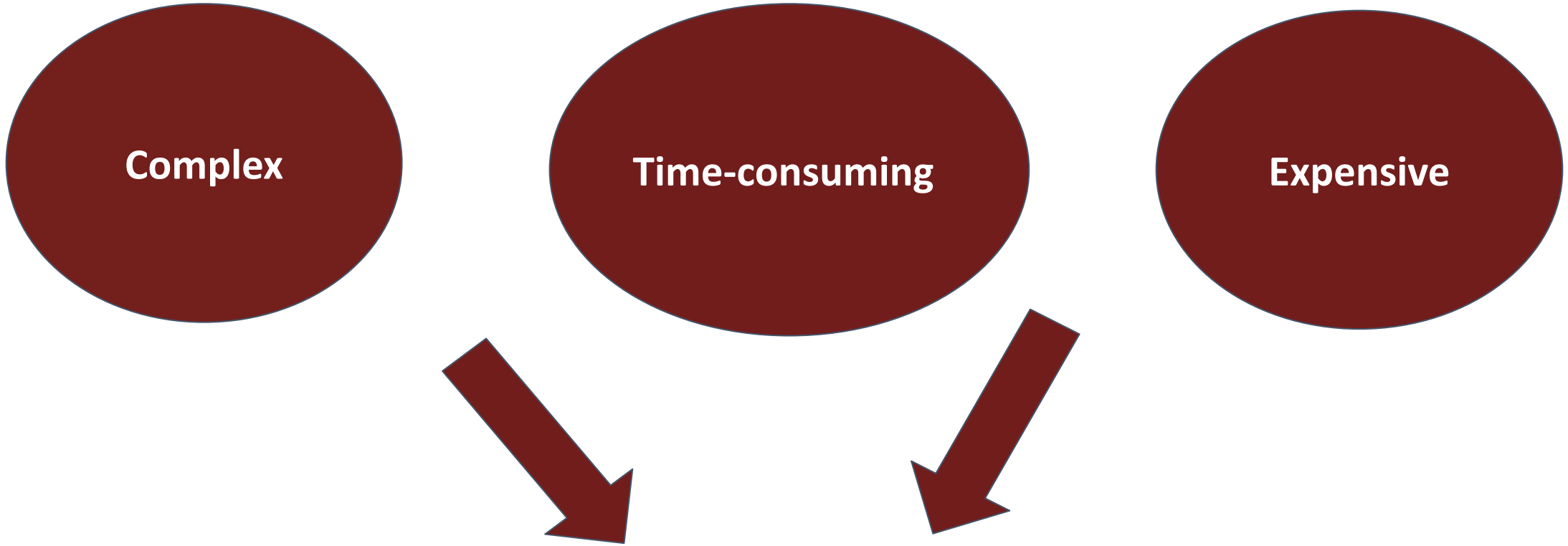


Modern Drug Discovery Limitations

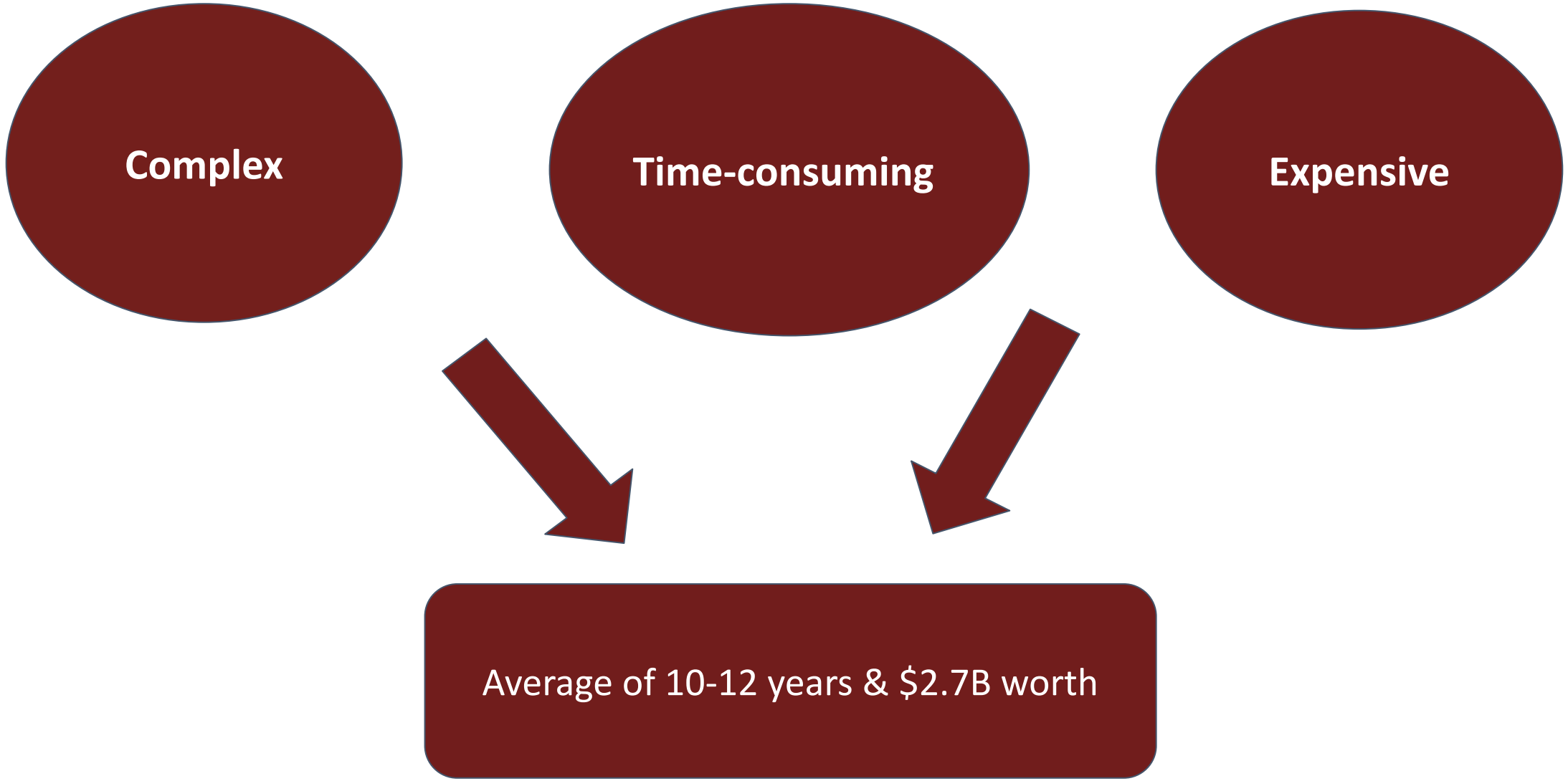


Modern Drug Discovery Limitations

Modern Drug Discovery Limitations



Modern Drug Discovery Limitations



Modern Health: AntiMicrobial Resistance

- What is AMR? Antimicrobial Resistance (AMR) occurs when bacteria, viruses, fungi and parasites **change over time** and **no longer respond to medicines** making **infections harder to treat and increasing the risk of disease spread, severe illness and death**.

- Why is it a global concern? The emergence and fast spread of drug-resistant pathogens threaten our ability to treat common infections.
  - **Consequence?** If nothing is done, even pretty common and usual diseases could be easily become FATAL
  - **State of Emergency** and Need to Action launched by the World Health Organization (WHO)

Source: WHO - Antimicrobial Resistance

How can ML help?

- ML helped achieve many breakthrough in Biology
  - AlphaFold, AlphaFold2, Breast Cancer Detection, just to name but a few

- Modern Computational Biology community thinks and agree that ML-based optimized models could help speed up the pipeline

  - “We believe that useful predictions generated by computational models combined with experimental validations could further speed up anticancer drug development” - Dr. Cui Wenqiang

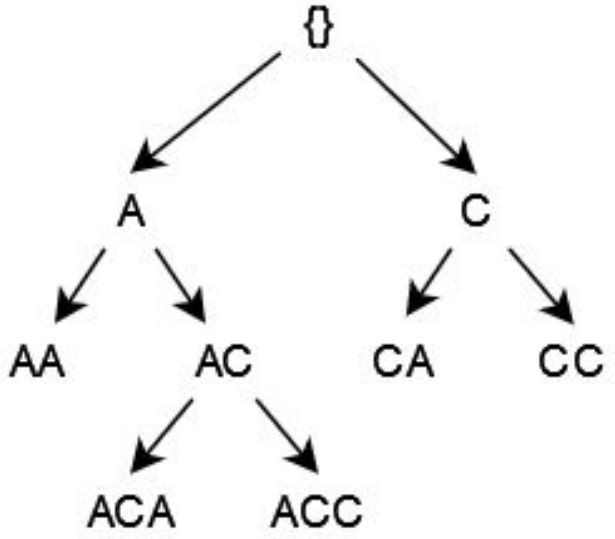
Biological Sequence Design with GFlowNets

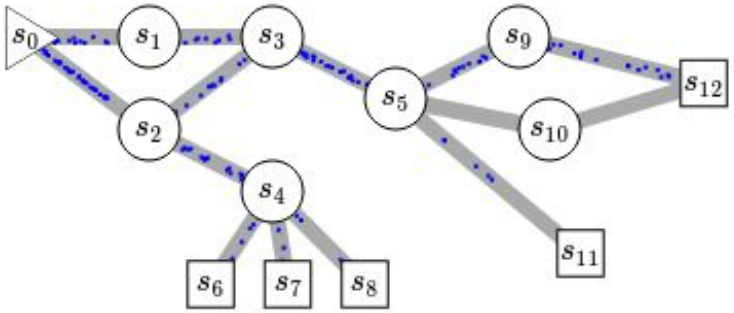
Moksh Jain, Emmanuel Bengio, Alex Hernandez-Garcia, Jarrid Rector-Brooks, **Bonaventure F. P. Dossou**, Chanakya Ekbote, Jie Fu, Tianyu Zhang, Micheal Kilgour, Dinghuai Zhang, Yelena Simine, Payel Das, and Yoshua Bengio.

International Conference in Machine Learning (ICML 2022)

Briefly about GFlowNets

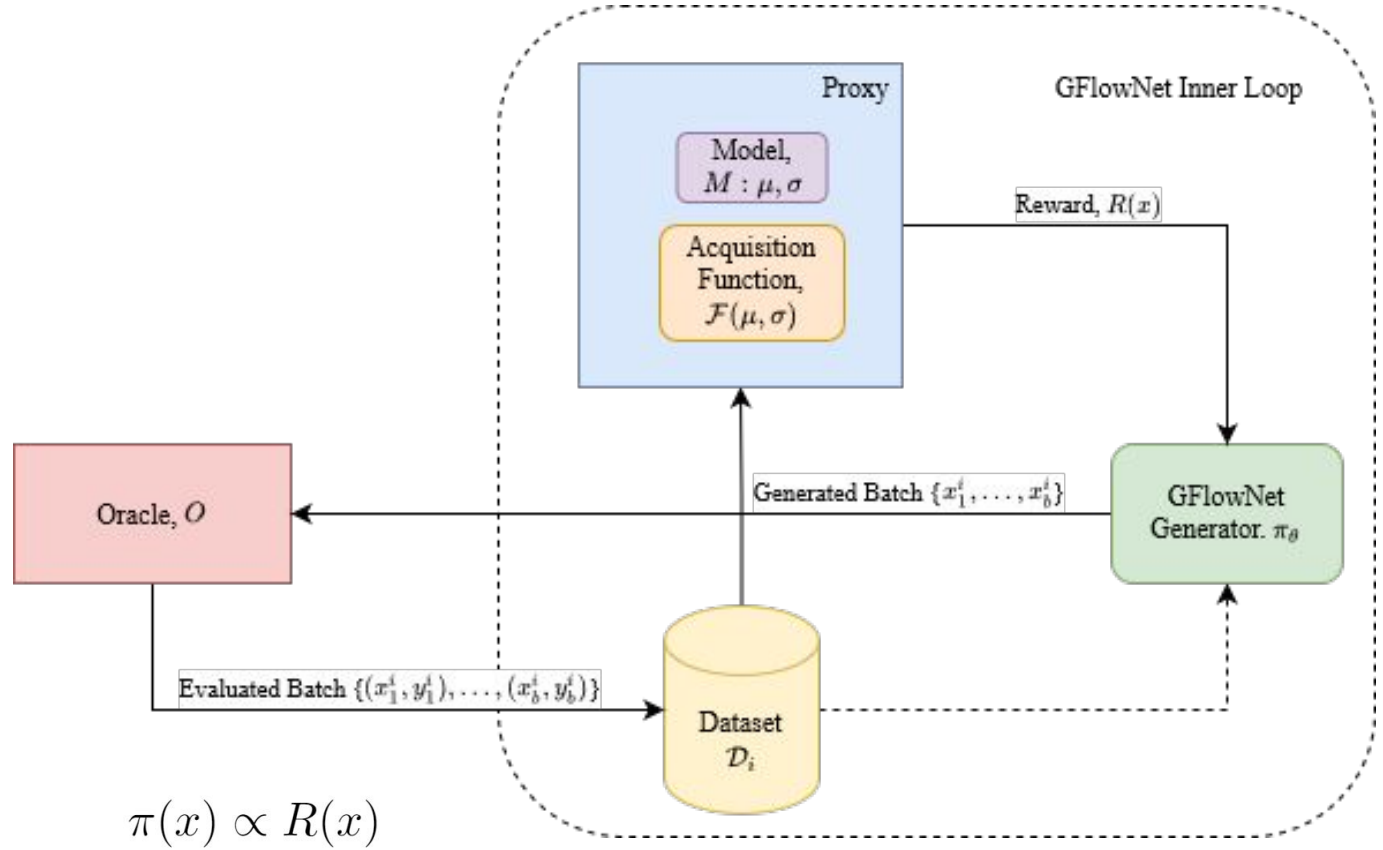
- Learn stochastic policies to sample compositional objects proportionally to reward, $\pi(x) \propto R(x)$
- Construct objects in forms of directed graphs, with edge weights corresponding to the **flow** (analogous to flow of particles) through the edge.
- Estimate flows with an offline and off-policy learning objective
- Use the learned policy to generate diverse and high scoring candidates
- Sample cost effectively





Active Learning with GFlowNets

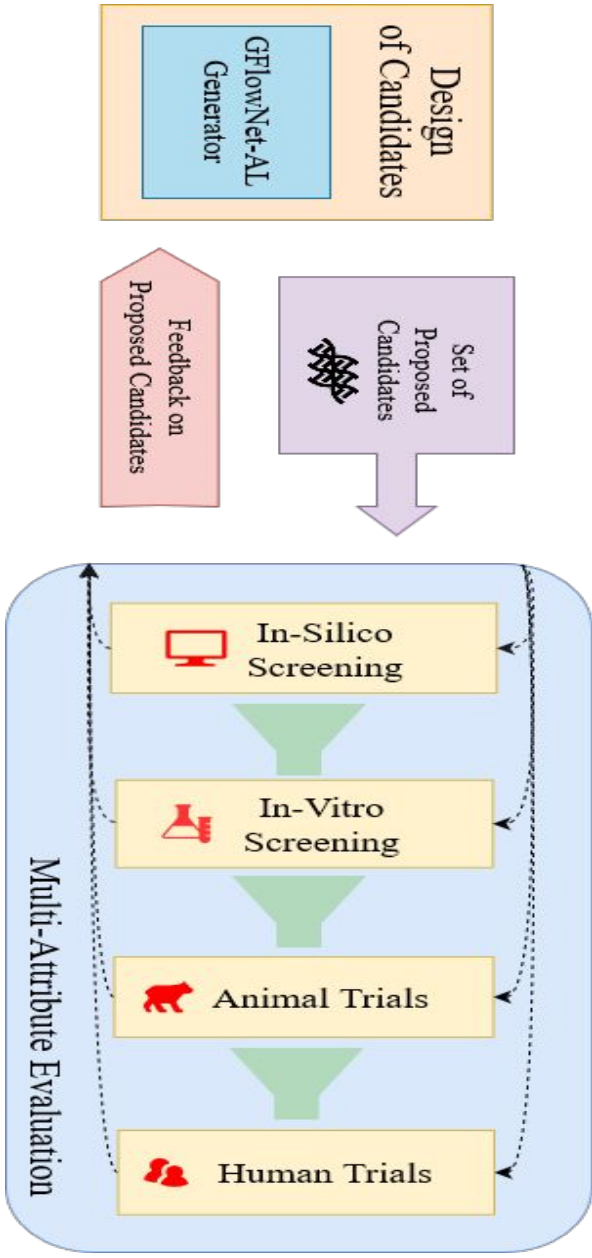
- Dataset:
  - Stores candidates generated so far, available for generating the next batch
- Proxy:
  - Model: Learned model of the Oracle
  - Acquisition Function: Combines the prediction and uncertainty of the model into a single scalar
- Generator:
  - Generates candidate batch, with policy



$$\pi(x) \propto R(x)$$

How to measure efficiency of generated candidates?

- Performance:
  - A measure of the overall quality of generated candidates
- Diversity:
  - A measure of the “modes” for the distribution captured by the method
- Novelty:
  - A measure of how different the generated candidates are from the initial dataset



Tasks

- TF-Bind-8: In the space of short DNA sequences, the goal is to search for DNA sequences with high binding activity with human transcription factors with a single round of active learning

- GFP: Similar to TF-Bind-8, in the protein's space, the goal is to search and discover proteins that are highly fluorescent with a single round of active learning

- AntiMicrobial Peptides Design: The goal of this task is to design peptide sequences with antimicrobial properties. We considered protein sequences with >=12 and <= 60 amino acids, with 10 rounds of active learning

Experimental Results (1)

Table 2. Results on TF-Bind-8 task with $K = 128$

	Performance	Diversity	Novelty
GFlowNet-AL	$\mathbf{0.84 \pm 0.05}$	4.53 ± 0.46	$\mathbf{2.12 \pm 0.04}$
DynaPPO	0.58 ± 0.02	5.18 ± 0.04	0.83 ± 0.03
COMs	0.74 ± 0.04	4.36 ± 0.24	1.16 ± 0.11
BO-qEI	0.44 ± 0.05	4.78 ± 0.17	0.62 ± 0.23
CbAS	0.45 ± 0.14	5.35 ± 0.16	0.46 ± 0.04
MINs	0.40 ± 0.14	$\mathbf{5.57 \pm 0.15}$	0.36 ± 0.00
CMA-ES	0.47 ± 0.12	4.89 ± 0.01	0.64 ± 0.21
AmortizedBO	0.62 ± 0.01	4.97 ± 0.06	1.00 ± 0.57

Table 3. Results on GFP task with $K = 128$

	Performance	Diversity	Novelty
GFlowNet-AL	$\mathbf{0.853 \pm 0.004}$	$\mathbf{211.51 \pm 0.73}$	$\mathbf{210.56 \pm 0.82}$
DynaPPO	0.794 ± 0.002	206.19 ± 0.19	203.20 ± 0.47
COMs	0.831 ± 0.003	204.14 ± 0.14	201.64 ± 0.42
BO-qEI	0.045 ± 0.003	139.89 ± 0.18	203.60 ± 0.06
CbAS	0.817 ± 0.012	5.42 ± 0.18	1.81 ± 0.16
MINs	0.761 ± 0.007	5.39 ± 0.00	2.42 ± 0.00
CMA-ES	0.063 ± 0.003	201.43 ± 0.12	203.82 ± 0.09
AmortizedBO	0.051 ± 0.001	205.32 ± 0.12	202.34 ± 0.25

Table 1. Results on the AMP Discovery Task with $K = 100$.

	Performance	Diversity	Novelty
GFlowNet-AL	$\mathbf{0.932 \pm 0.002}$	$\mathbf{22.34 \pm 1.24}$	$\mathbf{28.44 \pm 1.32}$
DynaPPO	$\mathbf{0.938 \pm 0.009}$	12.12 ± 1.71	9.31 ± 0.69
COMs	0.761 ± 0.009	19.38 ± 0.14	26.47 ± 1.3

GFlowNets achieves SOTA on Biological Sequence Design

Experimental Results (2)

- Use of epistemic uncertainty with the proxy leads to improvements across all metrics
- Better quality uncertainty estimates result in better performance

Table 4. Results on the AMP Task with $K = 100$ for GFlowNet-AL with different methods for uncertainty estimation, with UCB as the acquisition function.

	Performance	Diversity	Novelty
GFlowNet-AL Ensemble	$\mathbf{0.932 \pm 0.002}$	$\mathbf{22.34 \pm 1.24}$	$\mathbf{28.44 \pm 1.32}$
GFlowNet-AL MC Dropout	0.921 ± 0.004	18.58 ± 1.78	19.58 ± 1.12
GFlowNet-AL None	0.909 ± 0.008	16.42 ± 0.74	17.24 ± 1.44

Table 10. Results with GFlowNet-AL-None, where the uncertainty from the proxy is not used.

	Performance	Diversity	Novelty
AMP	0.909 ± 0.008	16.42 ± 0.74	17.24 ± 1.44
TF-Bind-8	0.81 ± 0.04	3.96 ± 0.32	1.73 ± 0.18
GFP	0.786 ± 0.001	205.28 ± 1.68	207.65 ± 1.19

Table 11. Results on AMP Generation Task with UCB and EI as acquisition functions and different methods for uncertainty estimation.

	UCB			EI		
	Performance	**Diversity**	**Novelty**	**Performance**	**Diversity**	**Novelty**
GFlowNet-AL-Ensemble	$\mathbf{0.932 \pm 0.002}$	22.34 ± 1.24	$\mathbf{28.44 \pm 1.32}$	0.928 ± 0.002	$\mathbf{23.61 \pm 1.05}$	26.52 ± 1.56
GFlowNet-AL-MCDropout	0.921 ± 0.004	18.58 ± 1.78	19.58 ± 1.12	0.917 ± 0.002	17.38 ± 0.64	18.34 ± 1.42

Table 12. Results on TF-Bind-8 Task with UCB and EI as acquisition functions and different methods for uncertainty estimation.

	UCB			EI		
	Performance	**Diversity**	**Novelty**	**Performance**	**Diversity**	**Novelty**
GFlowNet-AL-Ensemble	$\mathbf{0.84 \pm 0.05}$	$\mathbf{4.53 \pm 0.46}$	$\mathbf{2.12 \pm 0.04}$	0.84 ± 0.01	4.46 ± 0.58	2.02 ± 0.13
GFlowNet-AL-MCDropout	0.81 ± 0.03	3.89 ± 0.85	1.76 ± 0.15	0.81 ± 0.02	4.10 ± 0.43	1.92 ± 0.16

Table 13. Results on GFP Task with UCB and EI as acquisition functions and different methods for uncertainty estimation.

	UCB			EI		
	Performance	**Diversity**	**Novelty**	**Performance**	**Diversity**	**Novelty**
GFlowNet-AL-Ensemble	$\mathbf{0.853 \pm 0.004}$	211.51 ± 0.73	$\mathbf{210.56 \pm 0.82}$	0.851 ± 0.003	$\mathbf{212.03 \pm 0.64}$	208.31 ± 0.94
GFlowNet-AL-MCDropout	0.825 ± 0.007	204.76 ± 1.75	200.93 ± 0.46	0.838 ± 0.001	207.42 ± 1.24	208.31 ± 1.60

Conclusion

- GFlowNets achieves SOTA in the Biological Sequence Design task
- GFlowNets in Active Learning setup generates **large, diverse and novel batches**

Future works

- Improving the computation efficiency of retraining the proxy
- Better estimators of information gain
- Multi-fidelity and multi-objective extensions

Read more about our work

Biological Sequence Design with GFlowNets

Moksh Jain, Emmanuel Bengio, Alex Hernandez-Garcia, Jarrid Rector-Brooks, **Bonaventure F. P. Dossou**, Chanakya Ekbote, Jie Fu, Tianyu Zhang, Micheal Kilgour, Dinghuai Zhang, Yelena Simine, Payel Das, and Yoshua Bengio.

International Conference in Machine Learning (ICML 2022)

Mila Université de Montréal McGill UNIVERSITY JACOBS UNIVERSITY IBM

MORE PROJECTS

MeSH2Matrix

Outcomes

- Academic paper on Machine learning-driven biomedical relation classification based on the MeSH keywords of PubMed scholarly publications (Complete accepted to BIR@ECIR)

Project leads

Houcemeddine Turki (contact on Discord @ csisc#7682)
Bonaventure Dossou (contact on Discord @bona.dossou#3457
Chris Emezue (contact @ Chris Emezue#8673)

MORE PROJECTS

MeSH2Matrix

Outcomes

- Academic paper on Machine learning-driven biomedical relation classification based on the MeSH keywords of PubMed scholarly publications (Complete accepted to BIR@ECIR)

Project leads

Houcemeddine Turki (contact on Discord @ csi
Bonaventure Dossou (contact on Discord @bo
Chris Emezue (contact @ Chri

Privacy Preserving AI for African Health datasets

Outcome

Use privacy preserving AI tools to make an initial African health dataset accessible

Project leads

Archie Arakkal (contact on Discord @ Archie#9168)
Chris Fourie (contact on Discord @ Chris Fourie#5230)

Description

Approach African health data custodians (hospitals, universities, research groups) to help make their data more accessible using privacy preserving AI.

MORE PROJECTS

MeSH2Matrix

Outcomes

- Academic paper on Machine learning-driven biomedical relation classification based on the MeSH keywords of PubMed scholarly publications (Complete accepted to BIR@ECIR)

Project leads

Houcemeddine Turki (contact on Discord @ csi
Bonaventure Dossou (contact on Discord @bo
Chris Emezue (contact @ Chri

Privacy Preserving AI for African Health datasets

Outcome

Use privacy preserving AI tools to make an initial African health dataset accessible

Project leads

Archie Arakkal (contact on Discord @ Archie#9168)
Chris Fourie (contact on Discord @ Chris Fourie#5230)

Description

Approach African health data cust
more accessible using privacy pres

Survey paper on Machine Learning and Health in Africa

Outcome

Survey paper

Project leads

Chris Fourie (contact on Discord @ Chris Fourie#5230)
Houcemeddine Turki (contact on Discord @ csisc#7682)
Chris Emezue (contact @ Chris Emezue#8673)

Description

To address problems relating to machine learning and health in Africa, we first have to understand what problems exist and who is already working on them.

DEEP LEARNING INDABA

MORE PROJECTS

MeSH2Matrix

Outcomes

- Academic paper on Machine learning-driven biomedical relation classification based on the MeSH keywords of PubMed scholarly publications (Complete accepted to BIR@ECIR)

Project leads

Houcemeddine Turki (contact on Discord @ csi
Bonaventure Dossou (contact on Discord @bo
Chris Emezue (contact @ Chris

Privacy Preserving AI for African Health datasets

Outcome

Use privacy preserving AI tools to make an initial African health dataset accessible

Project leads

Archie Arakkal (contact on Discord @ Archie#9168)
Chris Fourie (contact on Discord @ Chris Fourie#5230)

Description

Approach African health data cust
more accessible using privacy pres

Survey paper on Machine Learning and Health in Africa

Outcome

Survey paper

Project leads

Ch
Ho
Ch

De
To
pr

DEEP LEARNING INDABA

CONCLUSION

Many significant **GAPS** in African Health Machine Learning Research.

We can find a place for **YOU** in this research field.

Come and join **SISONKEBIOTIK**, the African ML for Health Community.

We are here to let your **DREAMS** come true.

DEEP LEARNING INDABA

JOIN US ON

SISONKE BIOTIK

THANK YOU

SISONKEBIOTIK@GMAIL.COM

HTTPS://SISONKEBIOTIK.AFRICA