Mastering the Secrets of the Wikimedia Research and Development Ecosystem in Africa

Data Engineering and Semantics
هندسة البيانات و دلالتها

WIKIMEDIA TN User Group

SCAN ME
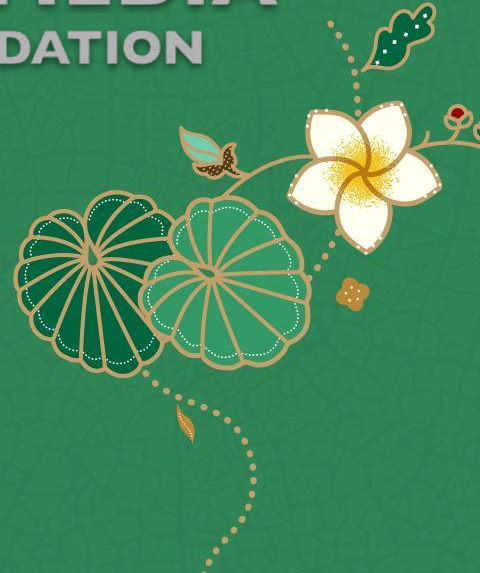
WIKIMANIA SINGAPORE | Houcemeddine Turki
User:Csisc

Disclosure

This work is a part of the "Adapting Wikidata to support clinical practice using Data Science, Semantic Web and Machine Learning" Project, funded by the Wikimedia Research Fund of Wikimedia Foundation.
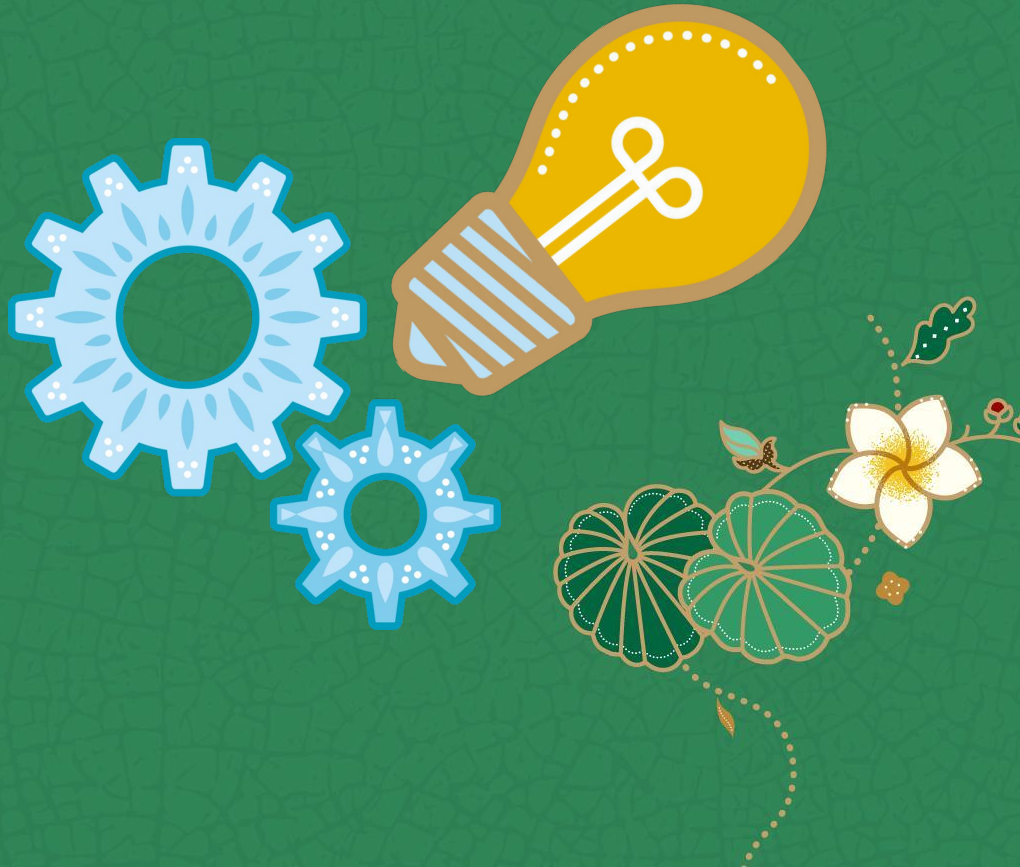
WIKIMANIA
SINGAPORE

This is an entertainment session

This should be fun

- A quiz of eight questions
- Letting the audience guess (3 minutes)
- Providing the answer
- Explaining the answer (2 minutes)

WIKIMANIA
SINGAPORE

Question 1

How many African countries are involved in Wikimedia Research and Development?

10

20

30

WIKIMANIA
SINGAPORE

Question 1

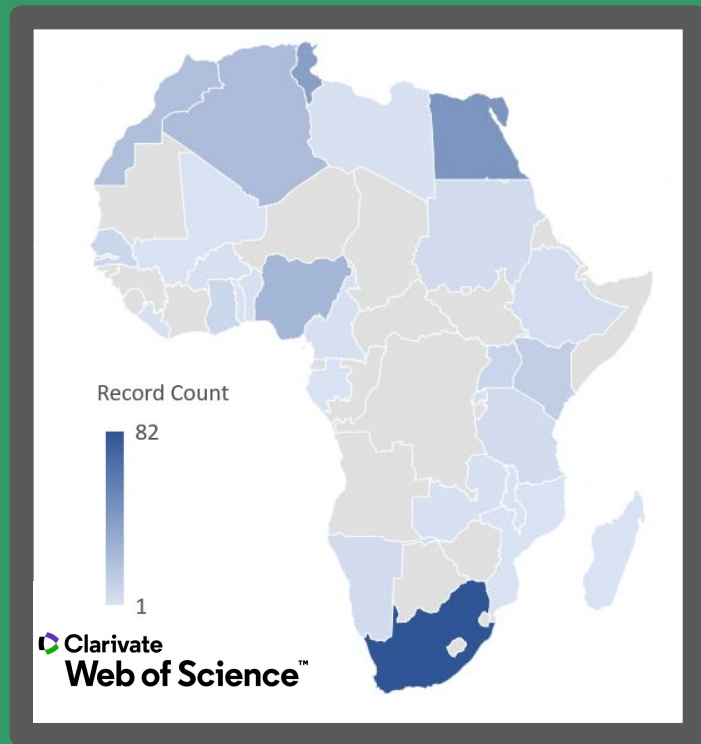How many African countries are involved in Wikimedia Research and Development?
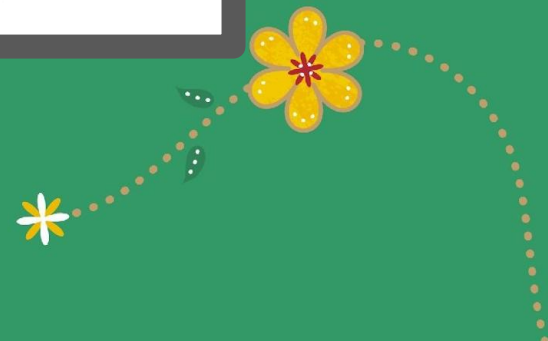
10

20

30

WIKIMANIA
SINGAPORE

Research

Development

WIKIMANIA
SINGAPORE

Question 2

When has the African Wikimedia Technical Community in Ghana been founded?

2016

2017

2018

WIKIMANIA SINGAPORE

Question 2

When has the African Wikimedia Technical Community in Ghana been founded?

2016

2017

2018

WIKIMANIA
SINGAPORE

AFRICA **WIKIMEDIA TECHNICAL** COMMUNITY

AFRICA WIKIMEDIA TECHNICAL COMMUNITY

AWMT Community. Growing technical contributors

for tomorrow

AWMT Community formerly known as **Africa Wikimedia Developers Project (AWMD)** initially started as a project that sought to start a developer community on the African continent and to encourage more developers to volunteer for the Wikimedia Foundation. Over the years, the project has broadened its scope and made opportunities for other technical contributions to be a part of the project hence the change of name from **Africa Wikimedia Developers Project (AWMD)** to **Africa Wikimedia Technical Community (AWMT)** Our **mission** is to create an enabling ecosystem for developers who are willing to learn and build new skills in the open space through volunteering as technical contributions for the Wikimedia Foundation.
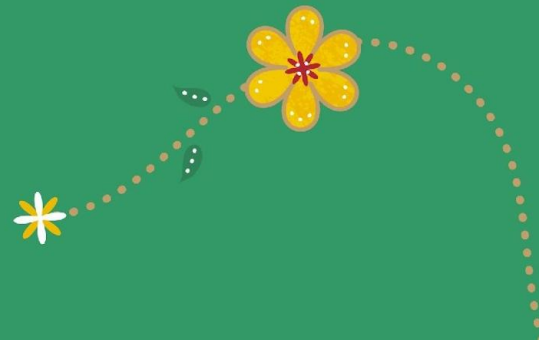
Felix Nartey — Project Management

Perside Rosalie — Community Engagement/Technical Advice

WIKIMANIA SINGAPORE

Question 3

Can you name five active Wikimedia Developers from Africa?

Yes

No

WIKIMANIA
SINGAPORE

Active Developers

Mahuton Possoupe
User:MPossoupe (WMF)
Benin

Alangi Derick
User:X-Savitar
Cameroon

Egbe Eugene Agbor
User:Eugene233
Cameroon

Sam Oyeyele
User:Jamie Tubers
Nigeria

Abel Lifaefi Mbula
User:BamLifa
DR Congo

WIKIMANIA
SINGAPORE

Question 4

Where is the most productive research institutions working on Wikimedia Projects?

Tunisia

Egypt

South Africa

WIKIMANIA
SINGAPORE

Question 4

Where is the most productive research institutions working on Wikimedia Projects?

Tunisia

Egypt

South Africa

WIKIMANIA SINGAPORE

Data Engineering and Semantics Research Unit

Data Engineering and Semantics
هندسة البيانات و دلالتها

University of Sfax, Tunisia

Informally working since 2011

Recognized as a nationwide research unit in 2012

Robust Infrastructure: Headquarters and Data Center

29 research publications as per Scopus

URL

YouTube

WIKIMANIA SINGAPORE

Question 5

Do African countries integrate the forefront of Wikimedia Research and Development?

Yes

No

WIKIMANIA
SINGAPORE

Question 5

Do African countries integrate the forefront of Wikimedia Research and Development?
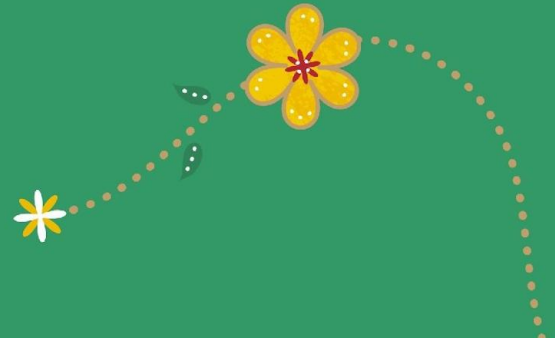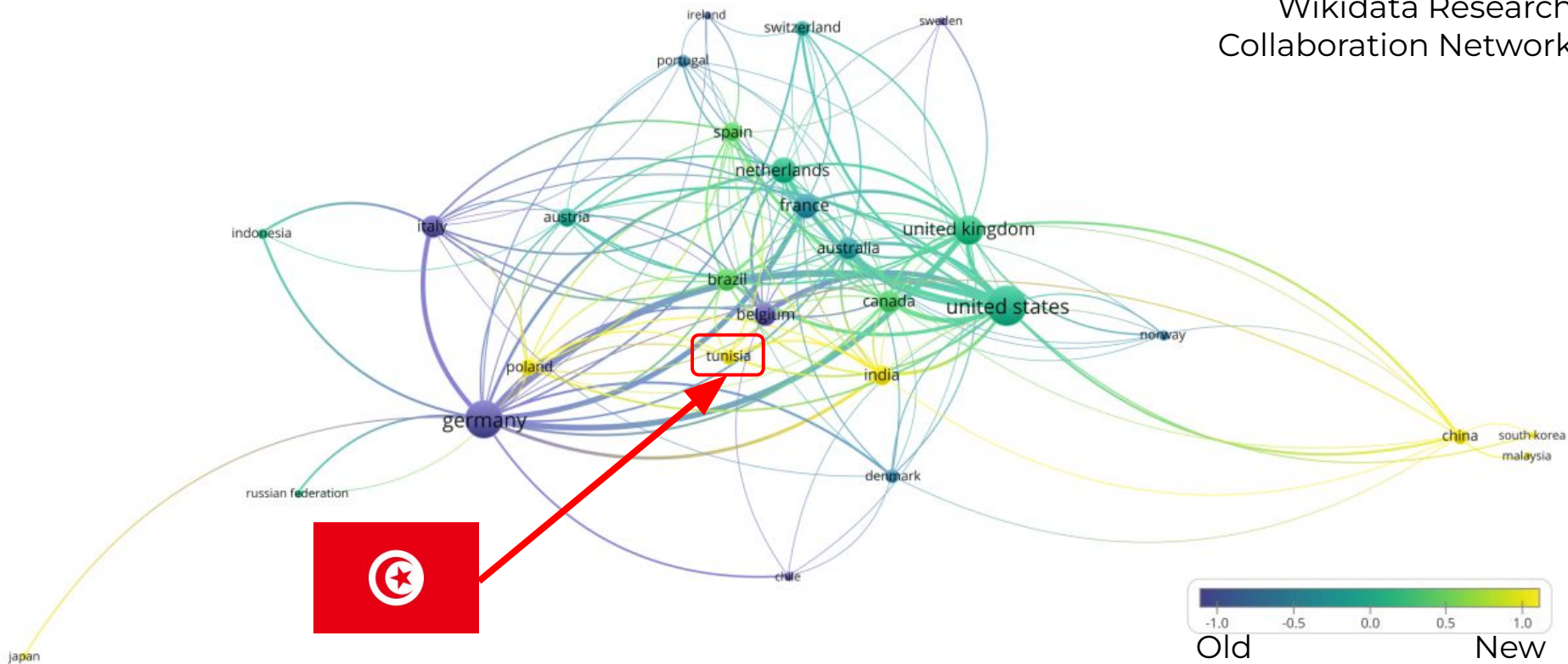
Yes

No

Not really...

- Recent integration of African countries
- Dependency on International Collaboration
- Mainly occurred thanks to collaboration with BRICS countries
- Poor productivity of African countries.

Source: Turki, H., Hadj Taieb, M. A., Ben Aouicha, M., Rasberry, L., & Mietchen, D. (2023). Ten years of Wikidata: A bibliometric study. In *Wikidata Workshop 2023* (forthcoming).

WIKIMANIA
SINGAPORE

Old   New

WIKIMANIA SINGAPORE

Question 6

What is the main target Wikimedia
Project for research and development
efforts?

Commons	Wikidata
Wikipedia	Wiktionary

**WIKIMANIA
SINGAPORE**

Question 6

What is the main target Wikimedia Project for research and development efforts?

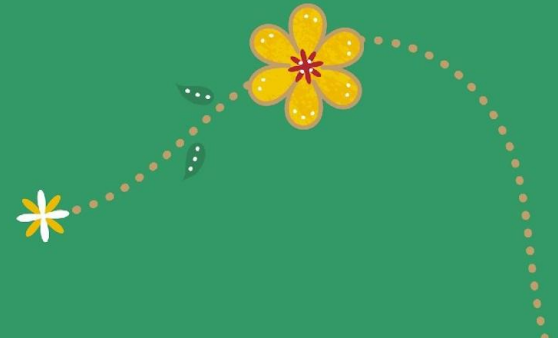|	
Commons	Wikidata
Wikipedia	Wiktionary

WIKIMANIA SINGAPORE

Wikipedia...

- +2.9k GitHub repositories
- +12,1k scholarly publications according to Scopus
- Tech giants like Google and IBM develop Wikipedia tools
- A huge interest from research institutions all over the world
- A very small productivity of Wikimedia volunteers

Source: Turki, H., Hadj Taieb, M. A., Ben Aouicha, M., Rasberry, L., & Mietchen, D. (2023). Ten years of Wikidata: A bibliometric study. In *Wikidata Workshop 2023* (forthcoming).

Source: Turki, H., Hadj Taieb, M. A., Ben Aouicha, M., Rasberry, L., & Mietchen, D. (2023). Preregistration: Comparing the use of Wikidata and Wikipedia by open-source software programmers on GitHub repositories. In *Wikidata Workshop 2023* (forthcoming).
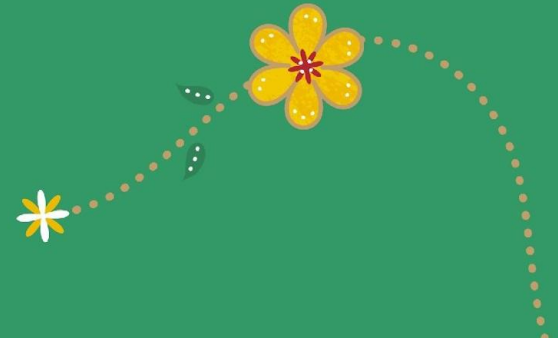
WIKIMANIA
SINGAPORE

However, Wikidata…

- +36.2k GitHub repositories
- +945 scholarly publications according to Scopus
- Easier to parse and reuse due to its structured format in the form of RDF triples
- Better engagement of Wikimedia volunteers and affiliates in research and development related to Wikidata

Source: Turki, H., Hadj Taieb, M. A., Ben Aouicha, M., Rasberry, L., & Mietchen, D. (2023). Ten years of Wikidata: A bibliometric study. In *Wikidata Workshop 2023* (forthcoming).

Source: Turki, H., Hadj Taieb, M. A., Ben Aouicha, M., Rasberry, L., & Mietchen, D. (2023). Preregistration: Comparing the use of Wikidata and Wikipedia by open-source software programmers on GitHub repositories. In *Wikidata Workshop 2023* (forthcoming).

WIKIDATA

WIKIMANIA
SINGAPORE

Question 7

What is the most efficient governance of non-profit Wikimedia research and development communities in Africa ?

Top-down governance
(Scientific society)

Self-organized distributed governance
(Grassroot)

Affiliate governance
(Wikimedia affiliate)

WIKIMANIA
SINGAPORE

Question 7

What is the most efficient governance of non-profit Wikimedia research and development communities in Africa ?

Top-down governance
(Scientific society)

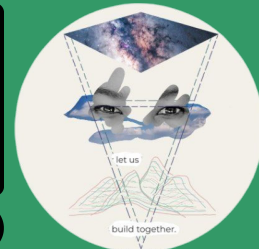**Self-organized distributed governance
(Grassroot)**

Affiliate governance
(Wikimedia affiliate)

Masakhane

The African Grassroots Community for NLP research

Wikimedia Foundation Research Award of the Year

The Wikimedia Foundation Research team established the Wikimedia Foundation Research Award of the Year in 2021 to recognize recent research that has the potential to have significant impact on the Wikimedia projects or research in this space.

Participatory Research for Low-resourced Machine Translation: A Case Study in African Languages (paper) and the Masakhane Community (learn more)

This paper and the Masakhane community have attempted to fundamentally change how we approach the challenge of "low-resourced languages" in Africa. The research describes a novel approach for participatory research around machine translation for African languages. The authors show how this approach can overcome the challenges these languages face to join the Web and some of the technologies other languages benefit from today.

Source: Currin, C., Asiedu, M. N., Fourie, C., Rosman, B., Turki, H., et al. (2023). A Framework for Grassroots Research Collaboration in Machine Learning and Global Health. In *MLGH@ICLR 2023* (Non-Archival).

WIKIMANIA SINGAPORE

status **Funded**

Wikimedia Research Fund

Understanding how Editors Use Machine Translation in Wikipedia: A Case Study in African Languages

start and end dates
July 2023 - July 2024

budget (USD)
50,000 USD

fiscal year
2022-23

applicant(s)
• Eleftheria Briakou, Tajuddeen Gwadabe and Marine Carpuat

give feedback

SisonkeBiotik

The African Grassroots Community for NLP research

SISONKE BIOTIK

AFRICA MACHINE LEARNING AND HEALTH WORKSHOP DEEP LEARNING INDABA 2022

@ Deep Learning Indaba, Tunis, Tunisia | 25-26 August 2022

Partners

WIKIMEDIA FOUNDATION

Data Engineering and Semantics
هندسة البيانات و دلالاتها

Wikimedia Foundation

Data Engineering and Semantics Research Unit

status **funded**

Wikimedia Research Fund

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

start and end dates
1 August 2022

end date
31 July 2023

budget (USD)
49,867.21 USD

applicant(s)
• Csisc

give feedback

Source: Currin, C., Asiedu, M. N., Fourie, C., Rosman, B., Turki, H., et al. (2023). A Framework for Grassroots Research Collaboration in Machine Learning and Global Health. In *MLGH@ICLR 2023* (Non-Archival).

WIKIMANIA SINGAPORE

Question 8

Is the African Artificial Intelligence community interested in Wikimedia research and development?

Yes

No

WIKIMANIA SINGAPORE

Question 8

Is the African Artificial Intelligence community interested in Wikimedia research and development?
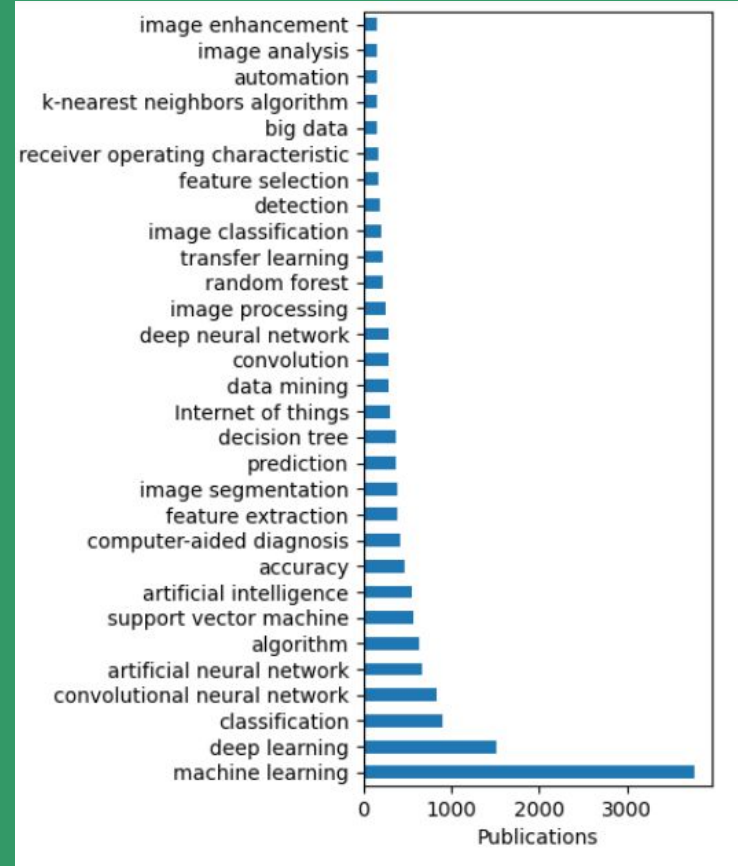
Yes

No

WIKIMANIA
SINGAPORE

Nope

Minor interest in using Wikimedia Projects for training machine learning applications

- Fancy techniques like Artificial Neural Networks
- Few-shot learning rather than knowledge-based systems
- Using benchmarks rather than open resources requiring pre-processing
- Interest in Image Analysis and Internet of Things rather than Semantic Analysis

Source: Turki, H., et al. (2023). Machine Learning for Healthcare: A Bibliometric Study of Contributions from Africa. *Expert Systems with Applications* (Forthcoming). https://www.preprints.org/manuscript/202302.0010/v2.

WIKIMANIA
SINGAPORE

Any question

Houcemeddine Turki
User:Csisc

E-mail: turkiabdelwaheb@hotmail.fr
Phone: +21629499418
Twitter: @Csisc1994
LinkedIn: Houcemeddine Turki

**WIKIMANIA
SINGAPORE**

WIKIMANIA
SINGAPORE