

# THE DATASCIENTIA ECOSYSTEM

Fostering sharing, learning,  
research and innovation at local  
level



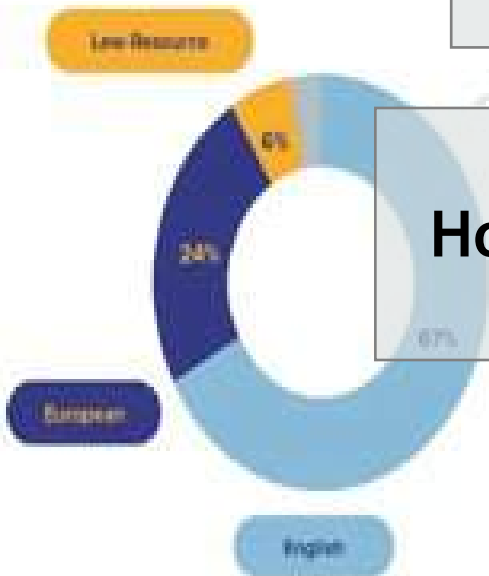
DataScientia

ICTAS – DataScientia Symposium

22.07.2025

Matteo Busso

## NLP Solutions by Language



8,000+ evolving languages  
Scientific experts and local experts  
(expertise so deep that few people know it)

## Population Size of Languages



How can we scale?  
How can we engage people in a long-lasting experience?

# INDEX



How to scale: ideas from Citizen Science



Learning experience: Community of practices



Requirements for a growing community



The DataScientia Community



The Community in action



Conclusion

# CITIZEN SCIENCE



**How to scale: ideas from Citizen Science**



Learning experience: Community of practices



Requirements for a growing community



The DataScientia Community



The Community in action



Conclusion

# WHAT IS CITIZEN SCIENCE?

- Citizen science refers to the **involvement of non-professional scientists** (ordinary members of the public) in scientific research.
- It enables **large groups** of people to contribute to **data collection**, analysis, and discovery, often through online platforms or local projects.
- This approach **helps researchers gather and process vast amounts of data** that would be impossible to handle alone.

## Design:

- Scientists design a project with clear tasks that can be completed by volunteers.

## Recruitment:

- People are recruited, often via the internet, schools, or community groups.

## Training:

- Participants are provided with tutorials or guidelines to perform tasks correctly.

## Contribution:

- Volunteers analyze data, report observations, or perform classifications (e.g., identifying species, counting objects).

## Validation:

- Data is checked for accuracy using methods like redundancy (multiple people doing the same task).

## Outcome:

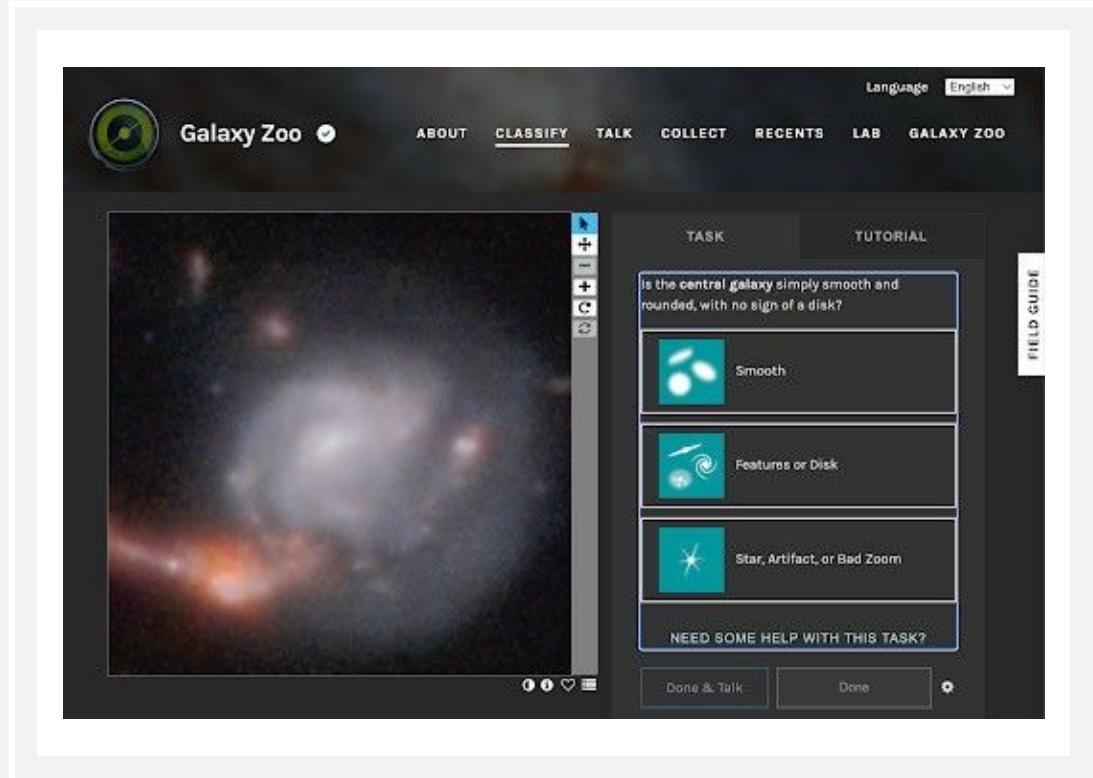
- Results contribute to scientific publications, conservation efforts, or technological advancements.

## EXAMPLE: GALAXY ZOO

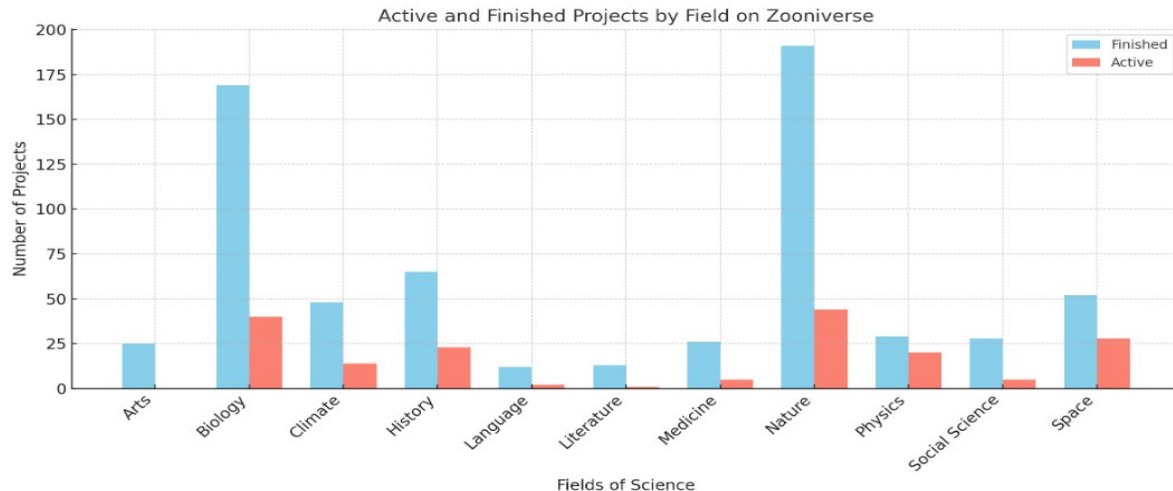
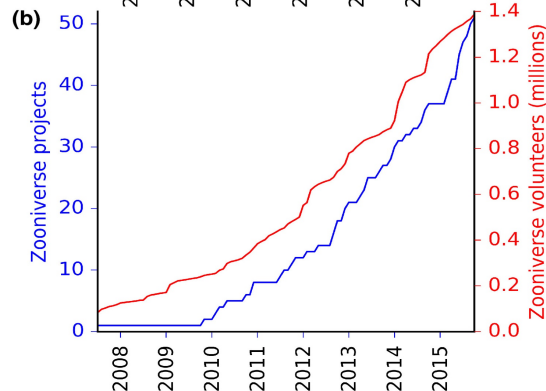
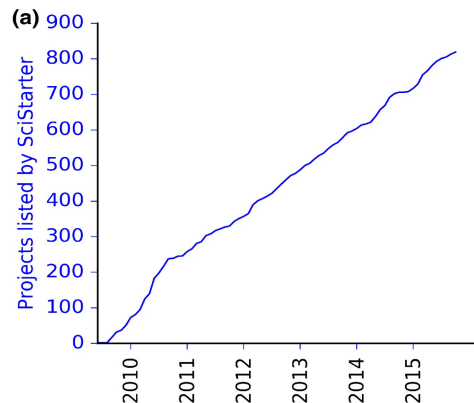
**Background:** Galaxy Zoo (launched in 2007) asked volunteers to classify images of galaxies from the Sloan Digital Sky Survey (SDSS).

**Task:** Participants determined galaxy shapes (spiral, elliptical, etc.) to help understand galactic evolution.

**Impact:** Millions of classifications were made, leading to numerous scientific papers and discoveries (including new



# EXPONENTIAL GROWTH IN POPULARITY



[iNaturalist](https://www.inaturalist.org)



[CitSci](https://citsci.org)



[eu-citizen.science](https://eu-citizen.science)



[SciStarter](https://SciStarter)

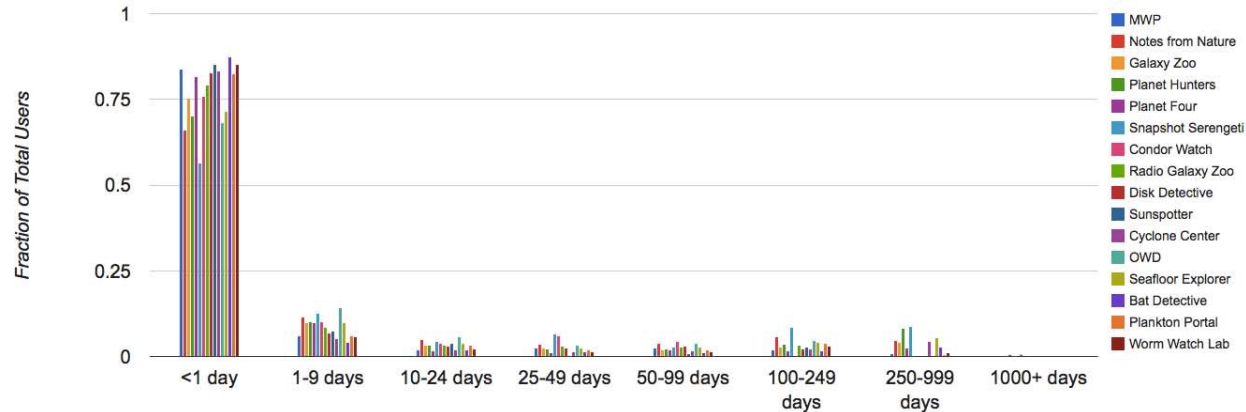


[zooniverse](https://zooniverse)

and many more ...

Increase in citizen-science projects and volunteers  
~ McKinley et al. [2015]

## DROP OUT IN 16 CITIZEN SCIENCE PROJECTS



**Source:** Segal, A., Gal, Y. A., Simpson, R. J., Victoria Homsy, V., Hartswood, M., Page, K. R., & Jirotko, M. (2015, May). Improving productivity in citizen science through controlled intervention. In *Proceedings of the 24th international conference on world wide web* (pp. 331-337).



## LIMITS OF CITIZEN SCIENCE APPROACH

● Insight	● Summary
● <b>High dropout rates</b>	● Over half of volunteers leave within a year, 60–83% don't return after the first visit
● <b>Core contributors</b>	● A small group supplies the majority of work: top 10% = ~80% of contributions
● <b>Motivation matters</b>	● Learning, to be part of something, s
● <b>Learning vs doing</b>	● Classifying objects doesn't necessi structured education is needed
● <b>Recognition</b>	● Switching from competitive leaderboards to egalitarian acknowledgment
● <b>Engagement diversity</b>	● Volunteer types vary—understanding this helps tailor retention strategies

**There is no AI community!**

**Sources:** **Improve quality** Shinbrot, H., & Shilkin, C. (2015). Group science: contribution patterns and their implications. *Proceedings of the national academy of sciences*, 112(3), 679-684.

**Enhanced data integrity** Shinbrot, X. R., et al. (2023). Why citizen scientists volunteer: the influence of motivations, barriers, and perceived project relevancy on volunteer participation and retention from a novel perspective. *Journal of Environmental Planning and Management*, 66(1), 122-142.

# TOWARDS PARTICIPATORY APPROACHES

**Contributory citizen science**  
(Majority of early online citizen science projects)



Scientist as  
project designer

Participatory  
technology  
or strategy

Citizens as  
data gatherers

VS.

**Co-creation & participatory approaches**  
(Citizen social sciences, AR, science shops)



Citizens' & CSOs'  
real-world  
problems

Scientist as  
co-designer and  
facilitator

Shared, open, and  
reflexive research  
process

# COMMUNITY OF PRACTICES



How to scale: ideas from Citizen Science



**Learning experience: Community of practices**



Requirements for a growing community



The DataScientia Community



The Community in action



Conclusion

# COMMUNITIES OF PRACTICE

- These are groups of people who meet regularly to learn, share experiences, solve problems, and grow collective knowledge. They exist everywhere—in both professional and social settings—and foster the development of tacit skills and shared practices.

A GUIDE TO MANAGING KNOWLEDGE

## CULTIVATING COMMUNITIES OF PRACTICE

ETIENNE WENGER  
RICHARD McDERMOTT  
WILLIAM M. SNYDER



HARVARD BUSINESS SCHOOL PRESS

## CHRYSLER CASE STUDY

- In the 1990s, Chrysler drastically improved its product development cycle by shifting from functional structures to product-oriented platforms.
- However, this created problems of knowledge fragmentation. The solution was the creation of **Tech Clubs**, informal cross-functional communities that facilitated knowledge sharing and led to the creation of an **Engineering Book of Knowledge (EBoK)**.



# THE DOUBLE-KNIT

## Manages living knowledge

- communities capture and update expertise even when projects end or company structures change.

## Accelerates problem solving

- employees can quickly access distributed knowledge and experience within the organization.

## Fosters innovation

- informal exchanges between experts stimulate new ideas and solutions.

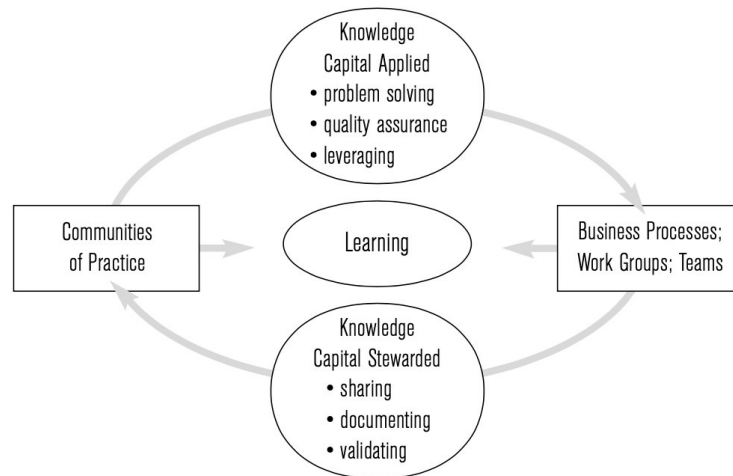
## Makes the organization more resilient

- even during reorganizations or market shifts, knowledge stays cohesive thanks to the communities.

## Strengthens identity and belonging

- employees maintain a “professional home” despite changes in organizational structures.

FIGURE 1-1 THE MULTIMEMBERSHIP LEARNING CYCLE



# FRACTAL STRUCTURE (SHELL PILOTS)

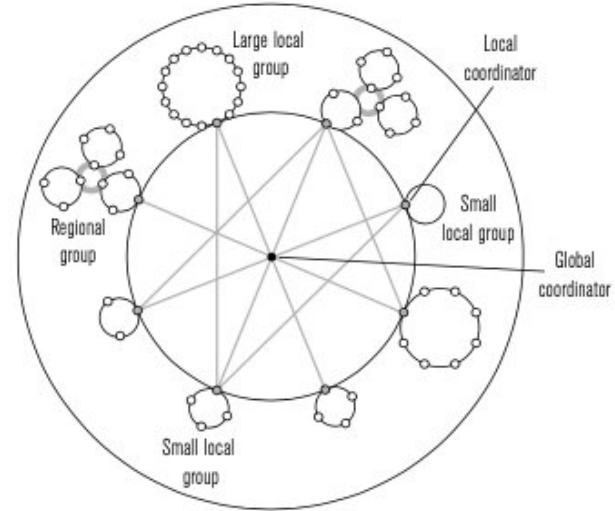
Structured as a Network of Local “Cells”:

- Local communities (cells) were formed inside business units.
- Each local cell had **autonomy** to structure itself (e.g., weekly meetings, informal chats).
- A **local coordinator** connected each cell to the global community.

A Coordinators’ Network:

- Coordinators regularly met (virtually and sometimes in-person).
- They became the backbone of the global network, facilitating both local and global knowledge sharing.

FIGURE 6-1 FRACTAL STRUCTURE FOR A GLOBAL COMMUNITY



Source: R. McDermott and J. Jackson, "Designing Global Communities"

# REQUIREMENTS



How to scale: ideas from Citizen Science



Learning experience: Community of practices



**Requirements for a growing community**



The DataScientia Community



The Community in action



Conclusion



IN A NUTSHELL:  
REQUIREMENTS FOR A  
GROWING COMMUNITY

### Citizen science

- Focus on data
- Diversity of projects
- Focus on the relationship between researcher and participant
- Platform for connecting people through data collections

### Community of practice

- Focus on knowledge and learning
- Focus on sharing
- Path to get involved and become an expert
- Path to foster local knowledge in a global setting

# THE DATASCIENTIA COMMUNITY



How to scale: ideas from Citizen Science



Learning experience: Community of practices



Requirements for a growing community



**The DataScientia Community**



The Community in action



Conclusion

# ABOUT DATASCIENTIA

A global Community shaping AI and data  
for the benefit of people and society

Combines sharing, education, research,  
and innovation in one ecosystem.

Open to all - citizens, researchers,  
developers, educators, institutions

In close collaboration with partner  
Universities



## THE DATASCIENTIA DOUBLE-KNIT

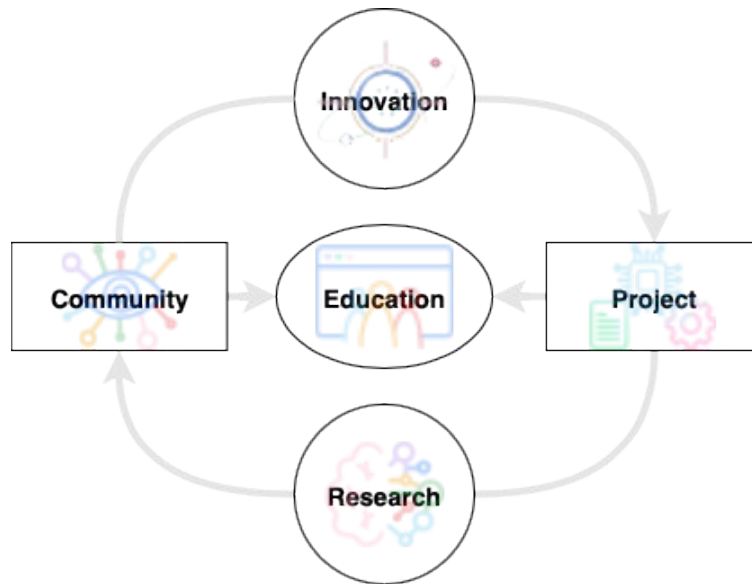
**Community:** Explore AI & data's real-life relevance with the help of others

**Projects:** Collaborate or lead ethical, inclusive scientific projects.

**Education:** Learn via multilingual courses and local learning paths.

**Research:** Develop and share new insights and knowledge

**Innovation:** Build real-world applications grounded in shared knowledge.



# The DataScientia Ecosystem



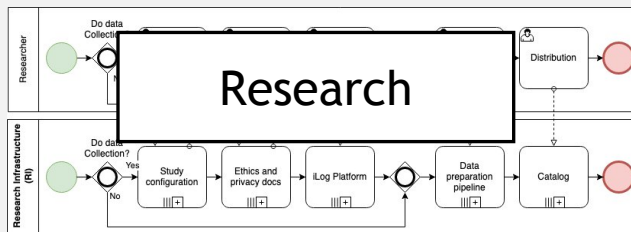
Innovation

Community

Education

Project

Research



**DataScientia COMMUNITY**  
AI made social, powered by you.

[Join Discussion Groups](#)

Learn collaboratively. Engage in moderated discussions to ask questions, share resources, and co-learn with the world.

**DIVERSITY ONE OPEN CHALLENGE**  
ACM UNICOMP 2025  
2025 October 12-13  
Espoo, Finland

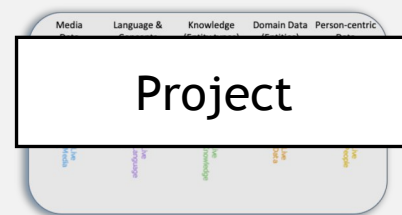
Join the first edition of the DiversityOne Open Challenge on October 12-13, 2025, in Espoo, Finland, and explore one of the most geographically diverse datasets for everyday life behavior modeling. This challenge provides a unique opportunity to work with rich mobile sensing data and receive valuable research feedback. Selected papers may be invited for an extended version in IEEE Pervasive Computing. [Learn More](#)

**22 JULY**  
**ICTAS 2025 Pre-conference Symposium**

© 10:00 am - 1:00 pm © Babes-Bolyai University of Technology, South Africa

2025 IEEE Conference on Information Communications Technology and Society (ICTAS) Pre-Conference Symposium DataScientia: Empowering African Languages in the Age of AI Despite the vast linguistic...

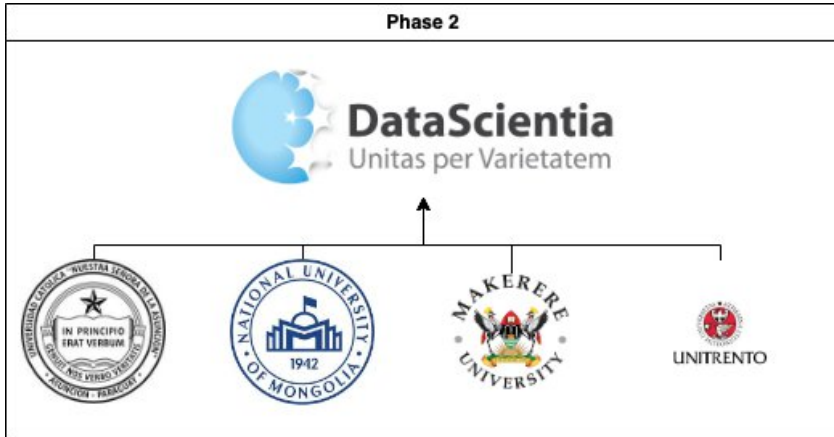
[READ MORE](#)



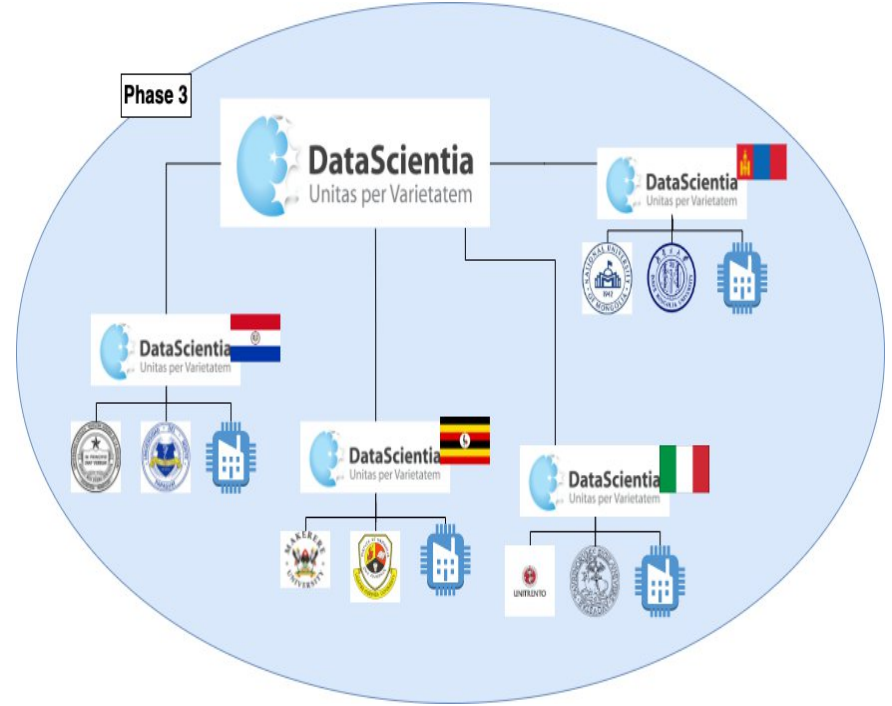
### Phase 1



### Phase 2



### Phase 3





## HOW TO GET INVOLVED?

There are many ways in which you can contribute. You can choose the one or more activities that best fit your competence, skills, and interests. Contribute to DataScientia by supporting data collection, educating others, programming AI tools, or managing local communities. You can create or participate in experiments, develop educational materials, share experiences, teach classes, code open-source AI tools, or lead local community efforts. Dive into diverse opportunities tailored to your skills and interests.



Contribute to collecting person-centric data by creating or participating in projects.



Contribute to education: materials, experiences, teaching.



Contribute in development of software and AI tools.



Help manage local DataScientia communities.



# THE COMMUNITY IN ACTION



How to scale: ideas from Citizen Science



Learning experience: Community of practices



Requirements for a growing community



The DataScientia Community



**The Community in action**



Conclusion

- Monetary
- Badges
- Recognition



Everyday AI:  
Introduzione ai dati dell'Intelligenza Artificiale

UNIVERSITÀ  
di PADOVA

Biblioteca

TRAILER

WORLDWIDE TOGETHER WITH PARTNER UNIVERSITIES

# 60+ PROJECTS

A VIBRANT COMMUNITY OF PEOPLE WITH DIVERSE BACKGROUNDS

all focused on human-aware artificial intelligence initiatives, gathering diversity aware resources about human behaviour, language, and knowledge. We create valuable datasets accessible to all, allowing us to understand everyday life activities across diverse communities.

BE PART OF OUR VIBRANT COMMUNITY

[Login](#)[Register](#)


[Join the Project](#)
[♥ Add to favourites](#)
[↗ Share](#)

No additional metadata available for the project.

# DiversityOne Open Challenge: Exploring People's Everyday Life Behavior with Mobile Data

♥ 2

Project Status

Inprogress

Participant Enrollment

Open

Start date

Sat Mar 01 2025

End date

Thu Oct 16 2025

Goal

The main goal of the workshop (co-located with Ubicomp 2025 conference on October 12 and 13, 2025) is to study the potential of the DiversityOne dataset given its size and multifaceted diversity, in terms of, e.g., sensors used, human feedback, and geographical diversity. The workshop topics are kept quite open to ensure contributions from a broader community of researchers. A non-exhaustive list of topics includes:

- studies exploiting the richness of the dataset both in size and in type of data;
- studies focusing on a diversity-aware comparison of human behaviors across cultures or profiles;
- studies focusing on the design and documentation of the dataset collection;
- studies focusing on the design affordances of the dataset;
- proposals for new types of data-driven studies;
- machine learning algorithms to improve the datasets;
- smartphone sensing for behavior modeling.

WORLDWIDE  
60+

A VIBRANT COMMUNITY

all focused on human aspects of  
behavior, language, and culture

WORLDWIDE TOGETHER WITH PARTNER UNIVERSITIES

## 60+ PROJECTS

A VIBRANT COMMUNITY OF PEOPLE WITH DIVERSE BACKGROUNDS

AI focused on human-aware artificial intelligence initiatives, gathering diversity aware resources about human behavioral, language, and technology. We create valuable diversity awareness in AI, allowing us to understand everyday life activities across diverse communities.

BE PART OF OUR VIBRANT COMMUNITY

[Login](#) [Register](#)

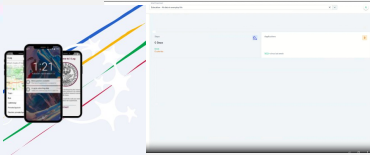


[Join the Project](#)

[Join the Community](#)

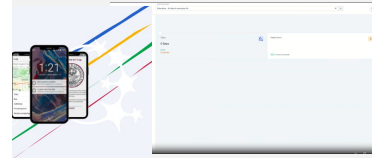
[Contact](#)

An additional resource available for the community



# Incentives

- Monetary
- Badges
- Recognition



WORLDWIDE TOGETHER WITH PARTNER UNIVERSITIES

## 60+ PROJECTS

A VIBRANT COMMUNITY OF PEOPLE WITH DIVERSE BACKGROUNDS

AI focused on human-aware artificial intelligence initiatives, gathering diversity aware resources about human behavioral, language, and knowledge. We create valuable datasets accessible to all, allowing us to understand everyday life activities across diverse communities.

BE PART OF OUR VIBRANT COMMUNITY

[Login](#) [Register](#)

**DiversityOne**  
Open People's Everyday Mobile Data

Project Name  
Project Description  
Start Date  
End Date  
Link

As additional resources available for the project



# Artificial Intelligence in Everyday Life 2.0: Educating University Students from Different Majors

Maria Kasinidou  
Open University of Cyprus  
Nicosia, Cyprus  
maria.kasinidou@ouc.ac.cy

Styliani Kleanthous  
CYENS Centre of Excellence &  
Open University of Cyprus  
Nicosia, Cyprus  
s.kleanthous@cyens.org.cy

Matteo Busso  
University of Trento  
Trento, Italy  
matteo.busso@unitn.it

Marcelo Rodas  
University of Trento &  
Fondazione Bruno Kessler

Jahna Otterbacher  
Open University of Cyprus &  
CYENS Centre of Excellence

Fausto Giunchiglia  
University of Trento  
Trento, Italy

## OPEN ONLINE COURSES

### ABSTRACT

With the surge in applications of AI, the associated advantages and disadvantages are widespread. Consequently, in the university setting, there is a crucial need to educate not only computer science majors but also students from various disciplines about AI. In this experience report, we present an overview of an introductory course that we offered to students coming from different majors. Moreover, we discuss the assignments and quizzes of the course, which provided students with a firsthand experience of AI processes and insights into their learning patterns. Additionally, we provide a summary of the course evaluation, as well as students' performance. Finally, we present insights gained from teaching this course and elaborate on our future plans.

### CCS CONCEPTS

• Social and professional topics → Computing education.

### KEYWORDS

Artificial Intelligence, AI Education, AI literacy, university students

With the surge in applications of AI, the associated advantages and disadvantages are widespread. Consequently, in the university setting, there is a crucial need to educate not only computer science majors but also students from various disciplines about AI. In this experience report, we present an overview of an introductory course that we offered to students coming from different majors. Moreover, we discuss the assignments and quizzes of the course, which provided students with a firsthand experience of AI processes and insights into their learning patterns. Additionally, we provide a summary of the course evaluation, as well as students' performance. Finally, we present insights gained from teaching this course and elaborate on our future plans.

Several initiatives have considered the design and evaluation of curriculum and courses on Machine Learning (ML) and AI [14], targeting university students in computer science (CS) [6, 22, 24], students in majors other than CS [25], younger students in middle-school, and K-12 students and teachers [15, 17, 18, 21]. A recent review found that although most initiatives focus on CS students, educating K-12 students on AI is also becoming popular [19].

Much effort by the research community, in collaboration with teachers, has been devoted to understanding how AI education at

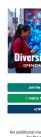
WORLDWIDE TOGETHER WITH PARTNER UNIVERSITIES  
**60+ PROJECTS**

A VIBRANT COMMUNITY OF PEOPLE WITH DIVERSE BACKGROUNDS

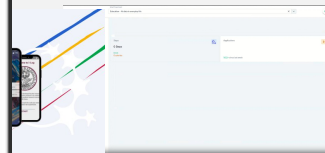
AI-based or human-assisted artificial intelligence initiatives, gathering diversity aware resources about human behaviour, language, and knowledge. We create valuable datasets, resources to AI, allowing us to understand everyday life activities across diverse communities.

BE PART OF OUR VIBRANT COMMUNITY

Log in Register



AI address book







UNIVERSITÀ  
DI TRENTO



DataScientia

# Everyday AI: Introduzione ai dati dell'Intelligenza Artificiale

TRAILER

WORLDWIDE TODAY  
60+ P

A VIBRANT COMMUNITY OF

AI focused on human-centric artificial intelligence, data science, and machine learning, for everyday life.

AI

AI

AI

AI

AI

AI

AI

AI

AI

AI

AI

AI

AI

AI

AI

AI

AI

AI

AI

AI

AI

AI

AI

AI

AI

AI

AI

AI

AI

AI

AI

AI

AI

AI

AI

AI

AI

AI

AI

AI




WORLDWIDE TOGETHER WITH PARTNER UNIVERSITIES

# 60+ PROJECTS

A VIBRANT COMMUNITY OF PEOPLE WITH DIVERSE BACKGROUNDS

AI focused on human-aware artificial intelligence initiatives, gathering diversity aware resources about human behavioral, language, and technology. We create valuable datasets accessible to all, allowing us to understand everyday life activities across diverse communities.

[Log in](#) [Register](#)



**DiversityOne**  
COMMUNITY

[Join the Project](#)

[Learn More](#)

[Contact](#)

**DiversityOne People/ Mobile**

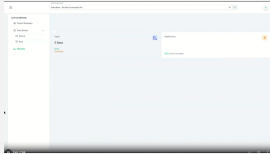
Project Details

How to use

FAQ

Team

An additional resource available for the project




☰

Select Experiment

Education - AI data in everyday life

✕

▼



ILOG DAHSBOARD

🏠 Project Summary

📅 Data Events

🕒 Future

🕒 Past

📊 My Data

Steps

0 Steps

0 km

0 calories

Applications

%52+ since last week

WORLDWIDE TOGETHER WITH PAR

60+ PRO




A VIBRANT COMMUNITY OF PEOPLE WITH

AI focused on human-aware artificial intelligence initiatives, gather behavioral, linguistic, and knowledge. We create valuable digital everyday life activities across diverse

BE PART OF OUR VIBRANT CO

Log In Register

2:41 / 7:06

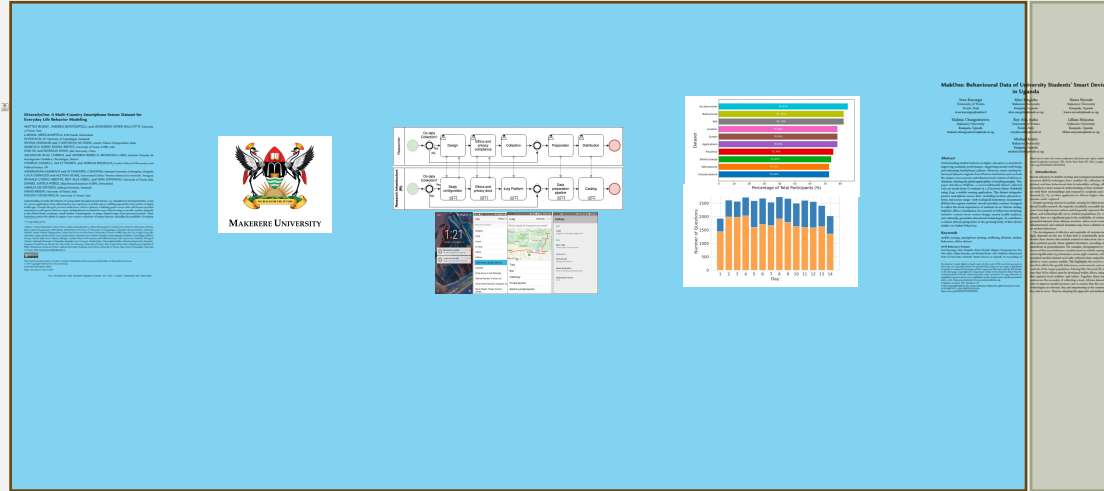
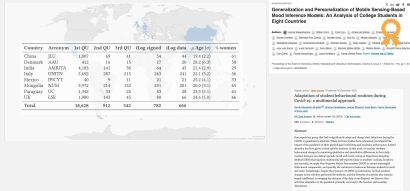


# THE MAKERERE UNIVERSITY EXPERIENCE (UGANDA)

## DataScientia Community GLOBAL

DS MAK  
LOCAL

Mak2  
Oct  
2025





Country	Acronym	1st QU	2nd QU	3rd QU	iLog signed	iLog data	$\mu$ Age ( $\sigma$ )	% women
China	JLU	1,007	69	41	54	44	19.4 (2.2)	61
Denmark	AAU	412	16	15	27	20	28.2 (6.3)	58
India	AMRITA	4,183	141	38	64	45	21.4 (2.9)	29
Italy	UNITN	5,692	287	215	263	241	22.1 (3.2)	56
Mexico	IPICYT	40	9	11	21	21	25.2 (4.1)	33
Mongolia	NUM	3,972	214	152	231	201	20.0 (3.1)	65
Paraguay	UC	1,342	33	25	43	28	23.3 (5.1)	61
UK	LSE	1,980	143	45	88	66	24.6 (5.0)	66
<b>Total</b>		<b>18,628</b>	<b>912</b>	<b>542</b>	<b>782</b>	<b>666</b>		





# Generalization and Personalization of Mobile Sensing-Based Mood Inference Models: An Analysis of College Students in Eight Countries

**Authors:**  Lakmal Meegahapola,  William Droz,  Peter Kun,  Amalia de Götzen,  Chaitanya Nutakki,  Shyam Diwakar,  Salvador Ruiz Correa,  Donglei Song,  Hao Xu,  Miriam Bidoglia,  George Gaskell,  Altangerel Chagnaa,  Amarsanaa Ganbold,  Tsolmon Zundui,  Carlo Caprini,  Daniele Miorandi,  Alethia Hume,  Jose Luis Zarza,  Luca Cernuzzi,  Ivano Bison,  Marcelo Rodas Brites,  Matteo Busso,  Ronald Chenu-Abente,  Can Günel,  Fausto Giunchiglia,  Laura Schelenz,  Daniel Gatica-Perez [\(Less\)](#) [Authors Info & Claims](#)

Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, Volume 6, Issue 4 • Article No.: 176, pp 1–32  
• <https://doi.org/10.1145/3569483>

Regular article | [Open access](#) | Published: 05 December 2023

## Adaptation of student behavioural routines during Covid-19: a multimodal approach

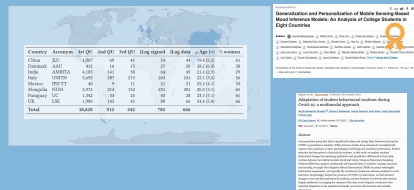
[Nicolò Alessandro Girardini](#) , [Simone Centellegher](#), [Andrea Passerini](#), [Ivano Bison](#), [Fausto Giunchiglia](#) & [Bruno Lepri](#)

[EPJ Data Science](#) 12, Article number: 55 (2023) | [Cite this article](#)

630 Accesses | 6 Altmetric | [Metrics](#)

### Abstract

One population group that had to significantly adapt and change their behaviour during the COVID-19 pandemic is students. While previous studies have extensively investigated the impact of the pandemic on their psychological well-being and academic performance, limited attention has been given to their activity routines. In this work, we analyze students' behavioural changes by examining qualitative and quantitative differences in their daily routines between two distinct periods (2018 and 2020). Using an Experience Sampling Method (ESM) that captures multimodal self-reported data on students' *activity*, *locations* and *sociality*, we apply Non-Negative Matrix Factorization (NMF) to extract meaningful behavioural components, and quantify the variations in behaviour between students in 2018 and 2020. Surprisingly, despite the presence of COVID-19 restrictions, we find minimal changes in the activities performed by students, and the diversity of activities also remains largely unaffected. Leveraging the richness of the data at our disposal, we discover that activities adaptation to the pandemic primarily occurred in the *location* and *sociality* dimensions.



## DiversityOne: A Multi-Country Smartphone Sensor Dataset for Everyday Life Behavior Modeling

MATTEO BUSSO\*, ANDREA BONTEMPELLI, and LEONARDO JAVIER MALCOTTI, *University of Turin, Italy*

LAOMAI MEDGASAPOLA, ETH Zurich, Switzerland

PETER ROUN, IT University of Copenhagen, Denmark

SHYAM DEBBAR and CHAITANYA NATH, Amrita Vishwa Vidyapeetham, India

MARCELO CARLOS RODAS BRITEZ, *University of Toronto & FIM, Italy*

HAO XU and DONGLI SONG, Nan University, China

SALVADOR RUIZ-CORREA and ANDREA-REBECA MENDOZA-LARA, Instituto Politécnico de Investigaciones Científicas y Tecnológicas, México

GEORGE GASKELL, SALLY STARES, and MIRIAM BIDOLLA, London School of Economics and Political Science, UK

ANABELLA CANELO and ALTACIRIO CHACONA, National University of Mexico, Mexico

LUCA CERNUZZI and ALETHA HUNE, *Università Cattolica "Ministero Sanzioni de la Anonimazione", Palermo*

RONALD CHENU-ASANTE, ROY ALIA AGRIU, and PIER RAMPANO, University of Trento, Italy

DANIEL GATICA-PEREZ, Mass Research Institute &amp; IITR, Switzerland

884414 IN THE CUPPIM, Salzburg University, Ehrensdorf

FRANCESCO BELLINI, University of Turin, Italy

PAULINO CUNHA-CUNHA, University of Torino, Italy

Understanding everyday life behavior of young adults through personal devices, e.g., smartphones and smartwatches, is becoming an important topic, from collecting the most capabilities in mobile apps to creating appropriate interventions in the health space. Towards this goal, previous studies have relied on datasets reflecting personal data and data with human-provided annotations or self-reports. However, many existing datasets are limited in scope, either focusing on specific countries (primarily in the United States), involving a small number of participants, or using a limited range of pre-provided sensors. These limitations restrict the ability to capture cross-country variations of human behavior, including the variability of studies.

[illegible][illegible]

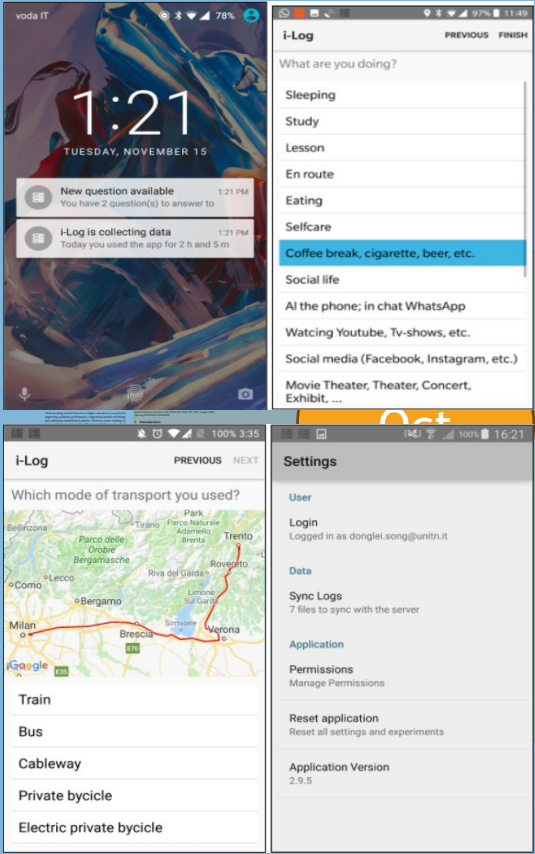
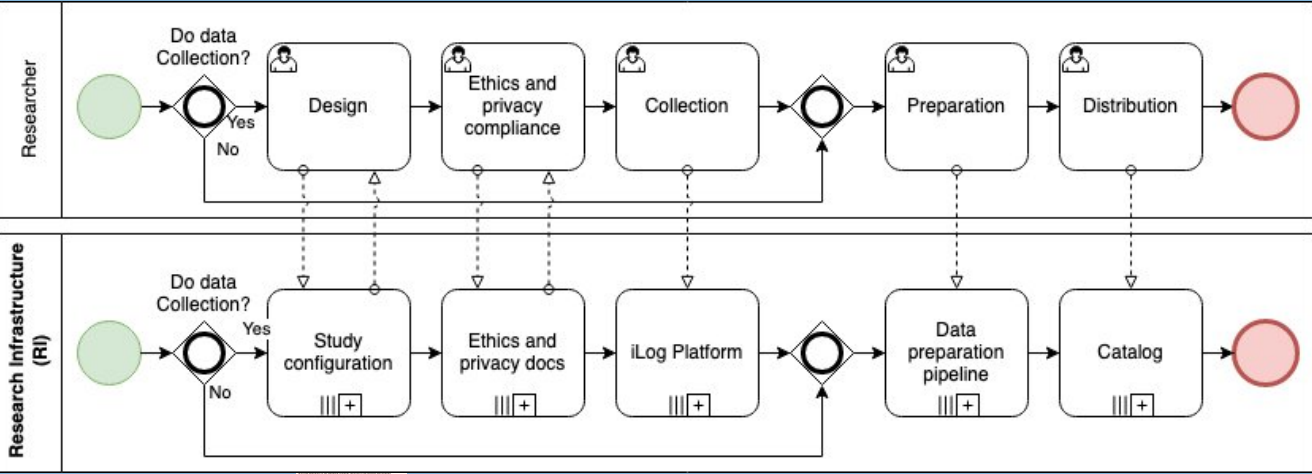
This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.  
© 2020 Copyright held by the owner/authors.  
All rights reserved. <https://doi.org/10.2196/2020.01.00000>  
<https://www.jmir.org/2020/1/e20200/>

Mak2  
Oct  
2025

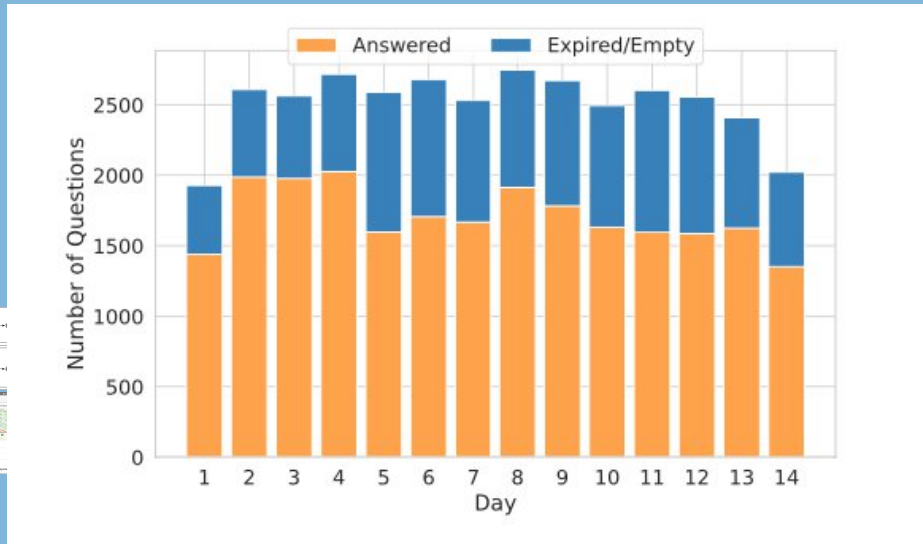
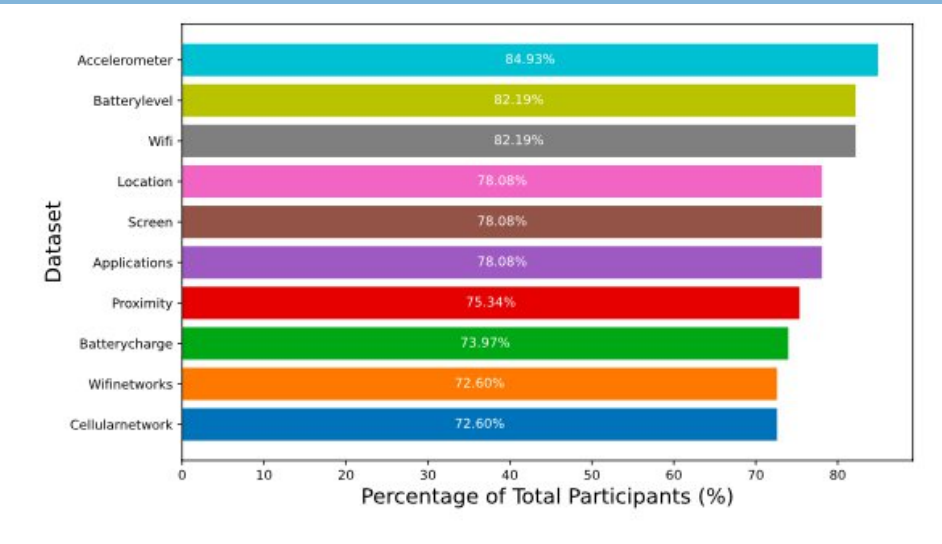


Mak2  
Oct  
2025

[illegible]

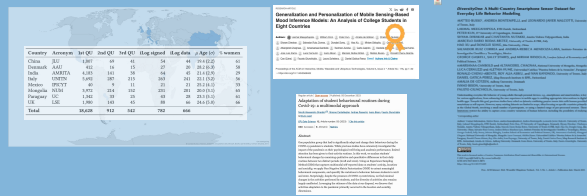








Scientific paper at ACM Ubicomp Conference



MakOne: Behavioural Data of University Students' Smart Devices in Uganda

- Ivan Kayunga  
University of Trento  
Trento, Italy  
ivan.kayunga@unitn.it
- Alice Magisha  
Makerere University  
Kampala, Uganda  
alice.magisha@mak.ac.ug
- Hawa Nyende  
Makerere University  
Kampala, Uganda  
hawa.nyende@mak.ac.ug
- Halima Chongomwezi  
Makerere University  
Kampala, Uganda  
halima.chongomwezi@mak.ac.ug
- Roy Alia Asiku  
University of Trento  
Trento, Italy  
royalia.asiku@unitn.it
- Lillian Moyama  
Makerere University  
Kampala, Uganda  
lillian.moyama@mak.ac.ug
- Michael Kizito  
Makerere University  
Kampala, Uganda  
michael.kizito@mak.ac.ug

**Abstract**  
Understanding student behavior in higher education is essential for improving academic performance, supporting mental well-being, and informing institutional policies. However, most existing behavioral datasets originate from Western institutions and overlook the unique socio-economic and infrastructural contexts of African institutions, limiting the global applicability of resulting insights. This paper introduces MakOne, a novel multimodal dataset collected over six weeks from 72 students at a (University Name Withheld) using i.i.g, a mobile sensing application. The dataset integrates passive smartphone sensor data—including location, physical activity, and screen usage—with ecological momentary assessments (EMAs) that capture students' moods and daily routines. Designed to reflect the lived experiences of students in an African setting, MakOne offers a foundation for research in behavior modeling, inclusive context-aware system design, mental health analytics, and culturally grounded educational technologies. It contributes a critical African perspective to the growing body of data-driven studies on student behavior.

**Keywords**  
mobile sensing, smartphone sensing, wellbeing, lifestyle, student behavior, Africa dataset

**ACM Reference Format**  
Ivan Kayunga, Alice Magisha, Hawa Nyende, Halima Chongomwezi, Roy Alia Asiku, Lillian Moyama, and Michael Kizito. 2025. MakOne: Behavioral Data of University Students' Smart Devices in Uganda. In *Proceedings of the ACM Ubicomp Conference*. ACM, New York, NY, USA, 5 pages. <https://doi.org/10.1145/XXXXXX.XXXXXX>

**1 Introduction**  
Recent advances in mobile sensing and ecological momentary assessment (EMA) techniques have enabled the collection of fine-grained, real-time behavioral data in naturalistic settings, offering researchers a more nuanced understanding of how students interact with their surroundings and respond to academic and social demands [5, 11], yet their applications in African higher education remain under explored.

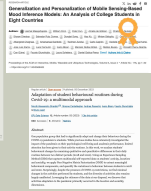
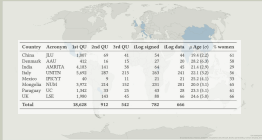
Despite growing interest in mobile sensing for behavioral and mental health research, the majority of publicly accessible datasets come from high-income nations and frequently represent Western, urban, and technologically savvy student populations [10, 11]. As a result, there is a significant gap in the availability of contextually grounded datasets from African societies, where socio-economic, infrastructural, and cultural dynamics may have a distinct impact on student behavior.

The development of effective and equitable AI systems increasingly depends on the use of data that is contextually grounded. Studies have shown that models trained on data from one region often perform poorly when applied elsewhere, revealing critical limitations in generalization. For example, Magisha et al. [8] showed that neural inference models based on mobile sensing data varied significantly in performance across eight countries, with personalized models trained on locally collected data outperforming global or cross-country models. This highlights the need to collect data that reflects the specific behaviors, environments, and cultural contexts of the target population. Following this, Shumway [14] emphasizes that AI for Africa must be developed within Africa, using data that captures local realities, and values. Together, these insights underscore the necessity of collecting a local, African dataset—not only to improve model accuracy, but to ensure that the resulting technologies are relevant, fair, and empowering to the communities they aim to serve. Thus by adapting the approach and methodology

Mak2  
Oct  
2025

Data Community GLOBAL

**Mak2**  
**Oct**  
**2025**



## OUR FIRST DATASCIENTIA JOINT LAB @ZHUHAI (CINA)

Focus on

Human Behavior Data  
Collection

Everyday AI Education



# CONCLUSION



How to scale: ideas from Citizen Science



Learning experience: Community of practices



Requirements for a growing community



The DataScientia Community



The Community in action



**Conclusion**



## CONCLUSION: DATASCIENTIA MISSION

- Empower individuals to understand, create, and govern AI.
- Promote cultural and contextual diversity in data and AI systems.
- Ensure all participants are active contributors, not passive subjects.
- Build tools, datasets, and learning paths for responsible innovation.

THANK YOU!!!

Don't forget our next DS  
Global Event

03-04 October in TRENTO  
(Italy)!!!



# COMMUNITY

News & Events

Monthly Newsletter

Data Marketplace

Interest Groups





## Introductory courses:

- Everyday AI series
  - Open University of Cyprus Interactive Annual Course
  - 4 MOOCs (EduGain CFU)

## Vertical courses:

- AI & Society
  - Knowledge Graph Engineering (3 Universities - UNITN, NUM, JLU)
  - Foundation Model
  - Studies on human behavior

We see education as a service to the community.

- Each course is available to everyone, adapted to everyone's diversity, anywhere in the world.
- Each course is accessible from the site and comprises individual modules that can be enjoyed separately.
- Each module is based on easily usable and adaptable resources depending on the context's needs.



## End-to-end process

## LingoGap - Task Creation

Design

Data collection

Data processing

Data distribution

# PROJECTS

## TYPE OF RESOURCES HANDLED

Data collection

Data curation

Data quality enhancement

Feature extraction

...

## TYPE OF ACTIVITIES (AS OF TODAY)

Media

- that depicts how the world appears to us, with all its diversity.

Data

- that describes how the world appears to us, i.e., what is true and false.

Personal

- data that describes how we are.

Languages

- that we use to describe how the world appears to us.

Knowledge

- that we use to provide a unitary view of how the world appears to us is the key to the purpose-driven composition of data.



User-oriented and context-aware applications  
For data collection,  
management, and sharing,  
e.g.:

- iLog app,
- Dashboard,
- iTelog methodology,
- DataScientia Catalogs

- Processes and services

SU2OSM: integrating personal GPS location with OpenStreetMap

SKEL: AI assistant for cleaning of your own data

...