

THE DATASCIENTIA ECOSYSTEM

Fostering sharing, learning,
research and innovation at local
level



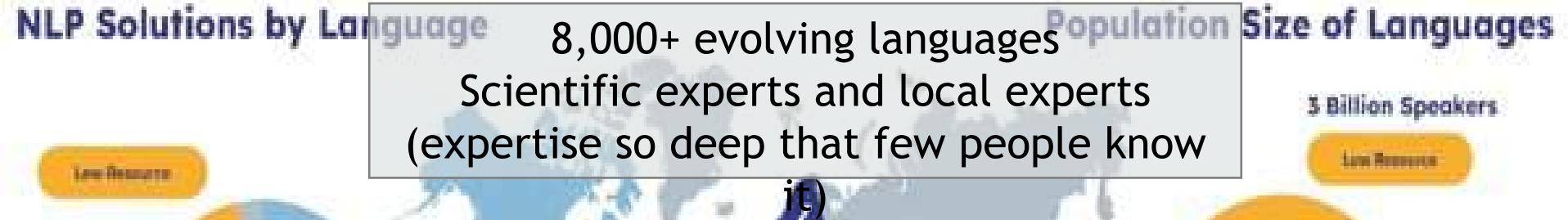
DataScientia

ICTAS – DataScientia Symposium

22.07.2025

Matteo Busso

NLP Solutions by Language 8,000+ evolving languages
Population Size of Languages



The background features a grayscale world map. Overlaid on it are several callout boxes containing text and small pie charts. One large box at the top right contains the text '8,000+ evolving languages' and 'Scientific experts and local experts (expertise so deep that few people know it)'. Another large box in the center contains the question 'How can we scale? How can we engage people in a long-lasting experience?'. To the left, a blue pie chart shows 'Low Resource' (47%) and 'High Resource' (53%). To the right, a yellow pie chart shows 'Low Resource' (33%) and 'High Resource' (67%). Other smaller boxes mention '3 Billion Speakers' (blue), '1 Billion Speakers' (yellow), and '400 Million Speakers' (orange). Buttons for 'Simpler', 'Ingenious', and 'Brilliant' are also visible.

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:

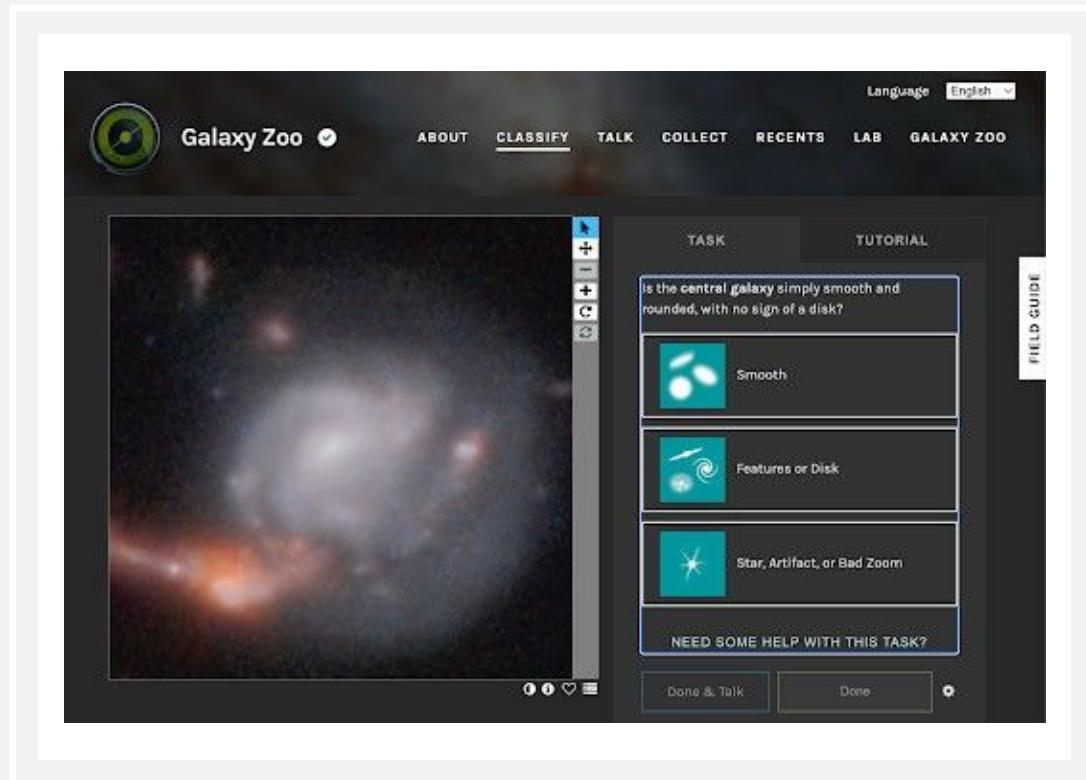
- Results contribute to scientific publications, conservation efforts, or technological advancements.

EXAMPLE: GALAXY ZOO

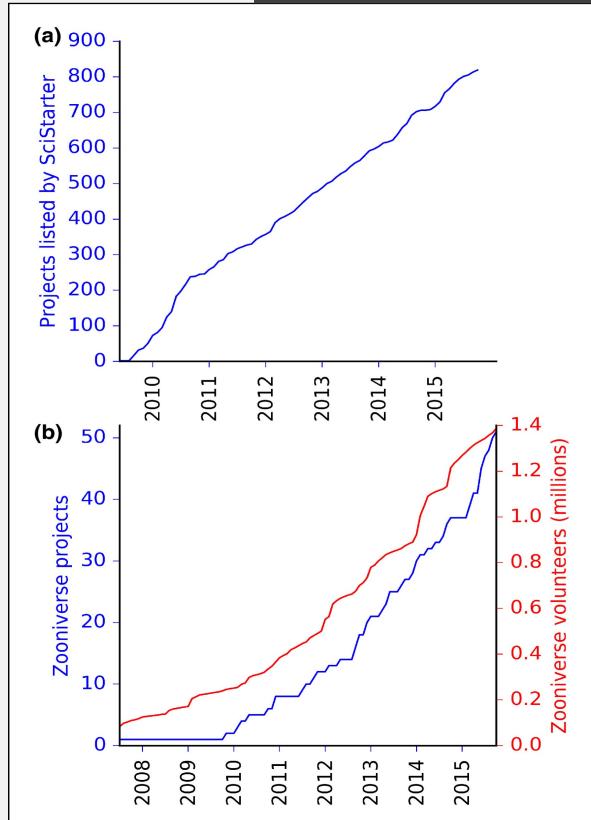
Background: Galaxy Zoo (launched in 2007) asked volunteers to classify images of galaxies from the Sloan Digital Sky Survey (SDSS).

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

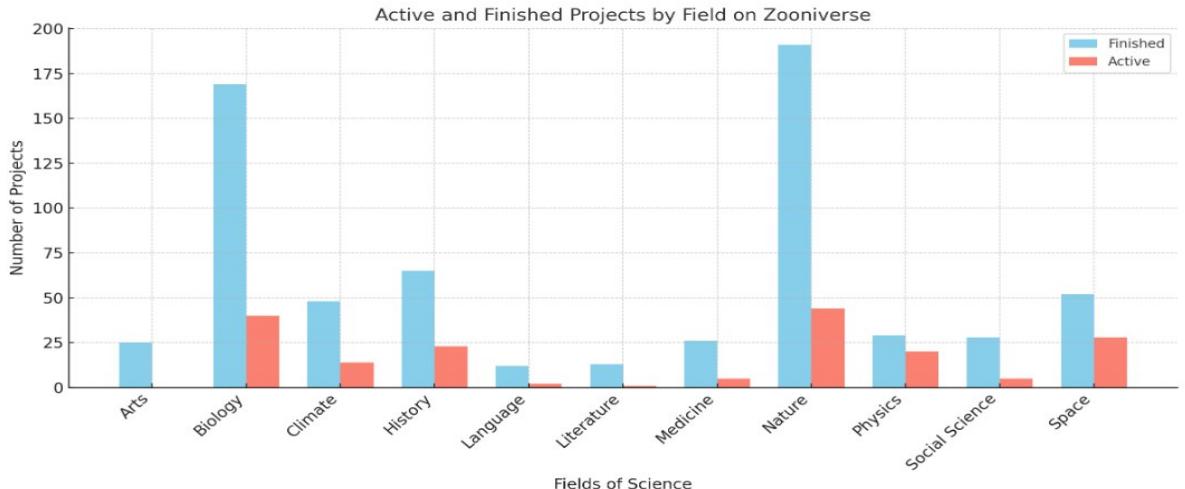
Impact: Millions of classifications were made, leading to numerous scientific papers and discoveries (including new



EXPONENTIAL GROWTH IN POPULARITY

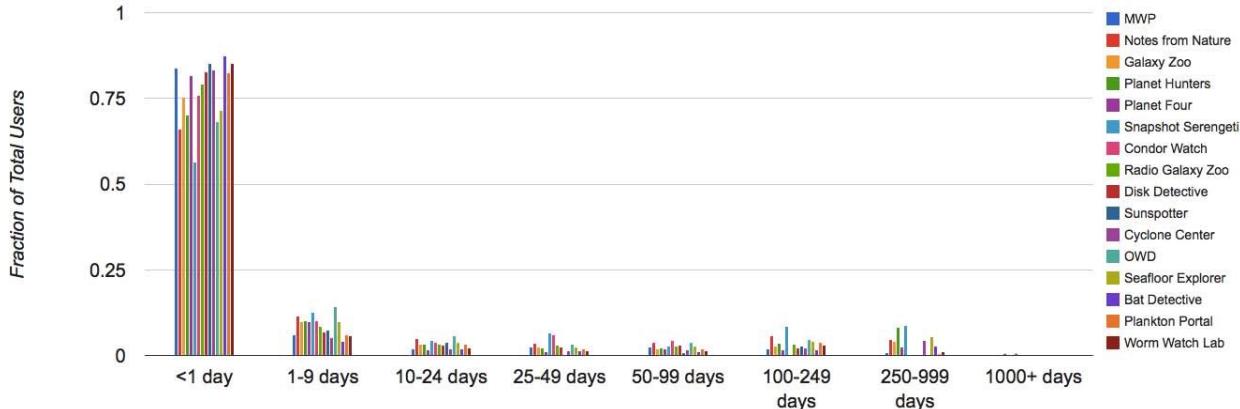


Increase in citizen-science projects and volunteers
 ~ McKinley et al. [2015]



and many more ...

DROP OUT IN 16 CITIZEN SCIENCE PROJECTS



Source: Segal, A., Gal, Y. A., Simpson, R. J., Victoria Homsy, V., Hartswood, M., Page, K. R., & Jirotka, 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
● High dropout rates	● Over half of volunteers leave within a year, 60–83% don't return after the first visit
● Core contributors	● A small group supplies the majority of work: top 10% = ~80% of contributions
● Motivation matters	● Learning, to be part of something, social interaction
● Learning vs doing	● Classifying objects doesn't necessarily mean learning; structured education is needed
● Recognition	● Switching from competitive leaderboards to egalitarian acknowledgment
● Engagement diversity	● Volunteer types vary—understanding this helps tailor retention strategies

Sources:

- Imperiali, H., & Lanzoni, C. (2015). Citizen science and enhanced data quality: Distribution patterns and their implications. *Proceedings of the national academy of sciences*, 112(3), 679-684.
- Shinbrot, X. A., 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.

There is no AI community!

TOWARDS PARTICIPATORY APPROACHES

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



VS.

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



Scientist as
project designer

Participatory
technology
or strategy

Citizens as
data gatherers

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.

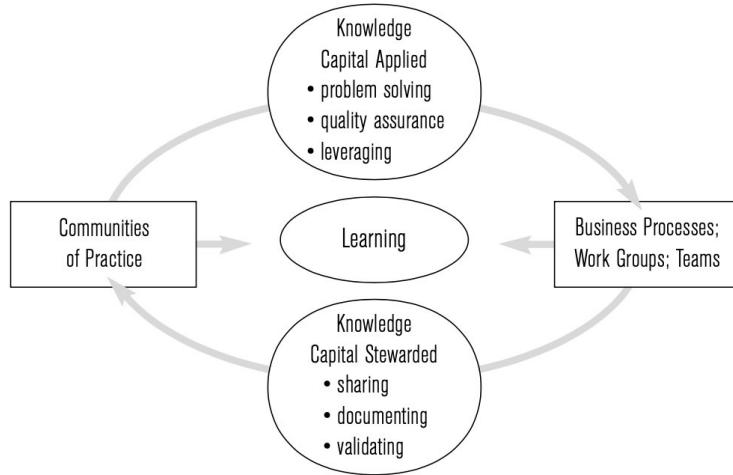
Makes the organization more resilient

- 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)

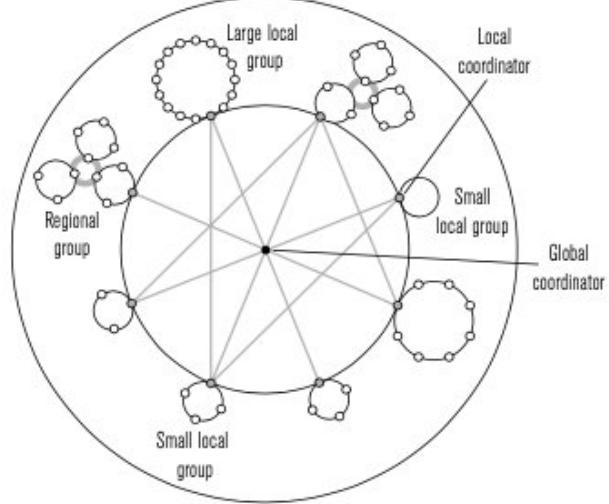
Structured as a Network of Local “Cells”:

- 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

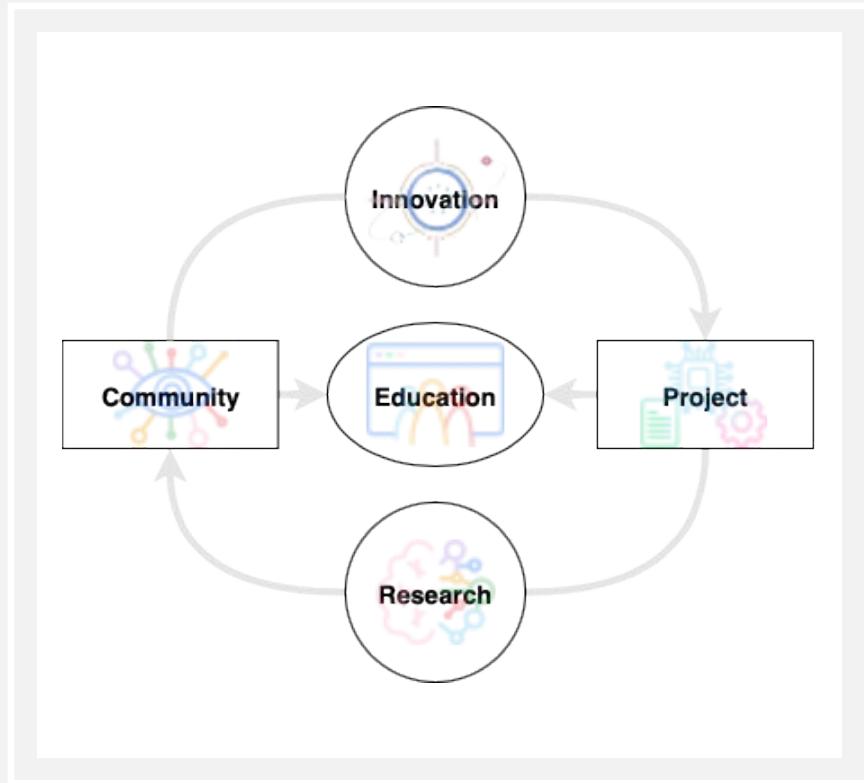
Community: Explore AI & data's real-life relevance with the help of others

Projects: Collaborate or lead ethical, inclusive scientific projects.

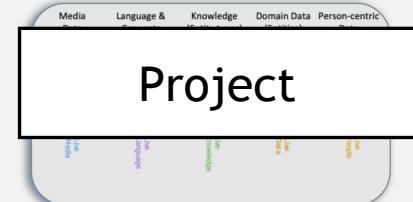
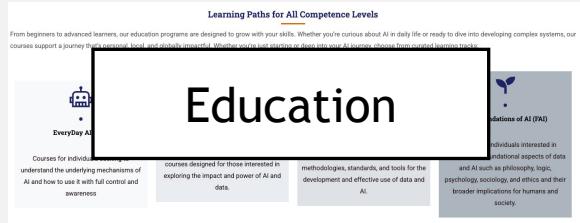
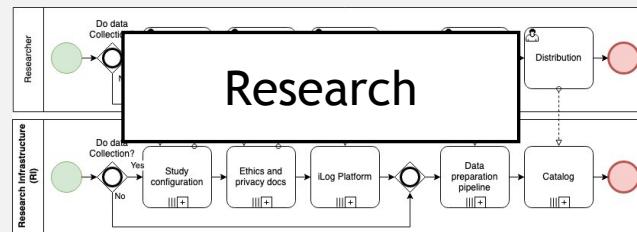
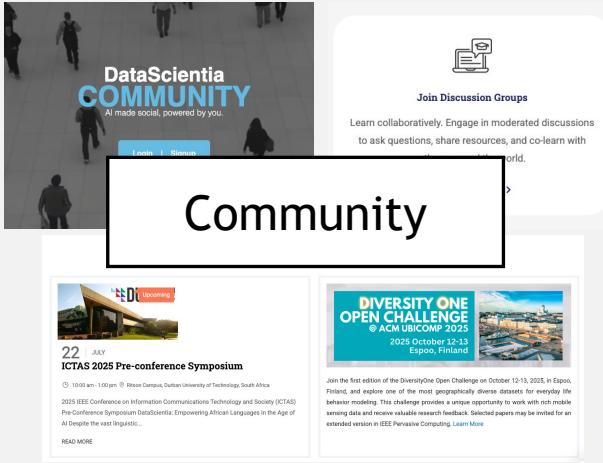
Education: Learn via multilingual courses and local learning paths.

Research: Develop and share new insights and knowledge

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



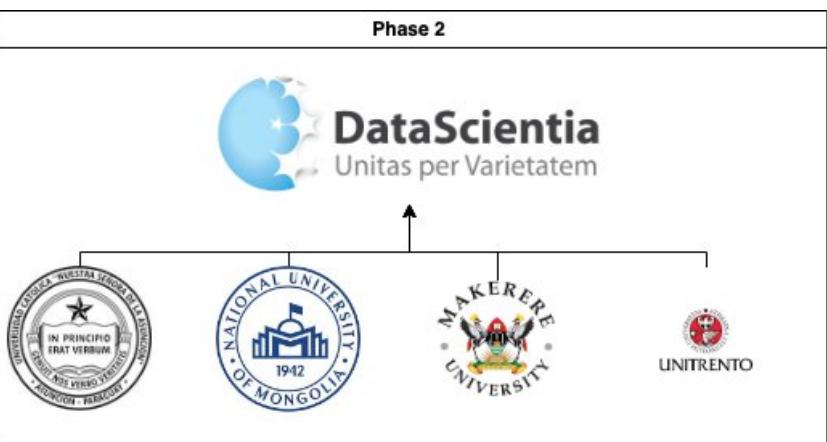
The DataScientia Ecosystem



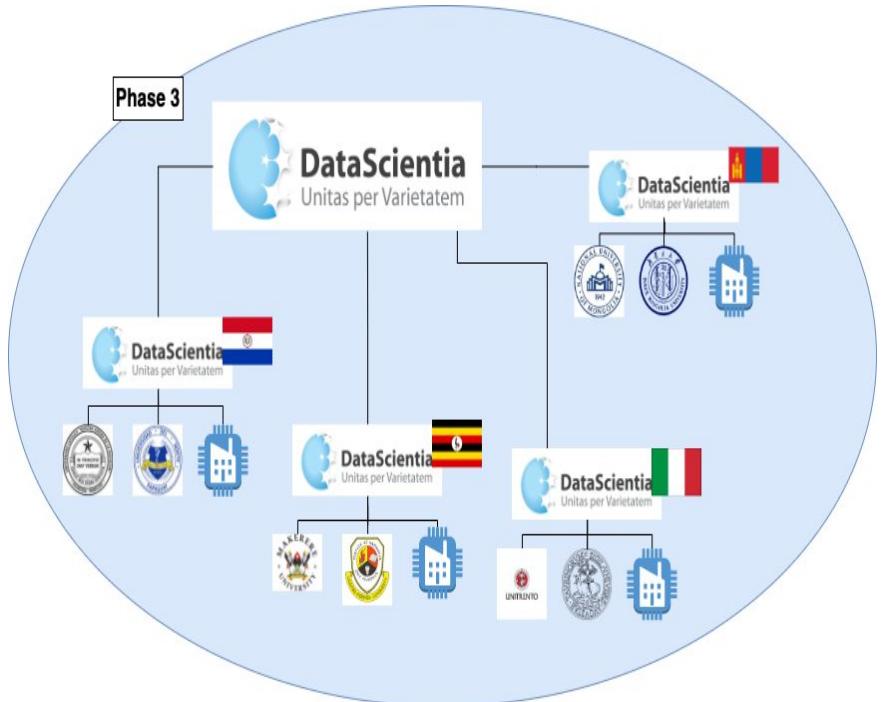
Phase 1



Phase 2



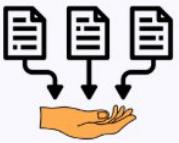
Phase 3





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

EXAMPLE: THE PARTICIPANT EXPERIENCE

WORLDWIDE TOGETHER WITH PARTNER UNIVERSITIES

60+ PROJECTS

A VIBRANT COMMUNITY OF PEOPLE WITH DIVERSE BACKGROUNDS

All focused on human-aware artificial intelligence initiatives, gathering diverse areas 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](#)

No additional metadata available for download



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

Project Name: DiversityOne Open Challenge

Description: The goal of the challenge is to explore people's everyday life behavior with mobile data. Participants will be asked to collect data from their phones and submit it to the challenge. The data will be used to train machine learning models to predict various aspects of people's everyday life.

Start Date: June 1, 2023

The start of the challenge is scheduled for June 1, 2023. Participants will have until July 31, 2023, to submit their data. The challenge will be open to anyone who wants to participate.

Prizes: There will be several prizes for the best submissions, including monetary rewards, badges, and recognition.



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

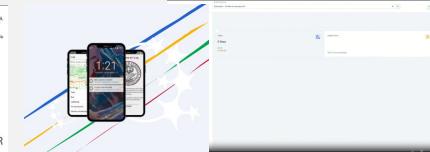
Project Name: Artificial Intelligence in Everyday Life 2.0

Description: This project aims to educate university students from different majors about artificial intelligence in everyday life. It will provide them with the knowledge and skills needed to use AI effectively in their daily lives.

Start Date: October 1, 2023

The start of the project is scheduled for October 1, 2023. Participants will have until December 31, 2023, to complete the course.

Prizes: There will be several prizes for the best submissions, including monetary rewards, badges, and recognition.



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



WORLDWIDE
60+
A VIBRANT COMMU-

all focused on human-aware or
behavior, language, and knowl-

[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

Project Status

inprogress

Participant Enrollment

Open

Start date

Sat Mar 01 2025

End date

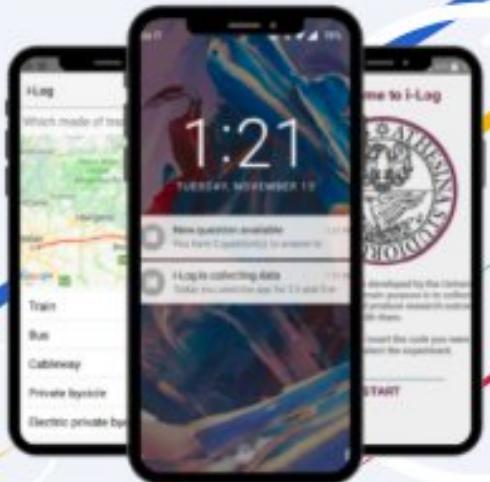
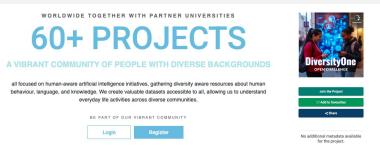
Thu Oct 16 2025

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.





Incentives

- Monetary
- Badges
- Recognition





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

ABSTRACT

With the surge of AI applications in our everyday life, there is a growing concern about the potential risks and benefits of AI. In this report, we present an overview of an introductory course that we offered to students coming from different majors. 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 future plans.

OPEN ONLINE COURSES

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 future plans.

CCS CONCEPTS

- Social and professional topics → Computing education.

KEYWORDS

Artificial Intelligence, AI Education, AI literacy, university students

also seen how serious the consequences of misunderstanding or failing to question AI decisions can be – leading to issues such as viral misinformation [8], biased systems that disproportionately impact marginalized communities [1], and serious concerns about data privacy. This situation highlights the need to bridge the gap between AI's everyday presence and people's lack of knowledge, so we can clear up misconceptions, reduce fears, and embrace a more informed relationship with the AI that is shaping our future [4].

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

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

WE PART OF OUR VIBRANT COMMUNITY

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

No account? [Create one](#)



DataScientia



UNIVERSITÀ
DI TRENTO



DataScientia

WORLDWIDE TEAM
60+ P

A VIBRANT COMMUNITY OF

All focused on human-aware artificial intelligence, machine learning, natural language processing, and knowledge representation.



Everyday AI: Introduzione ai dati dell'Intelligenza Artificiale

TRAILER



Select Experiment

Education - AI data in everyday life

ILOG DAHBOARD

- 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 PAR

All focused on human-aware artificial intelligence initiatives, gather behavior, language, and knowledge. We create valuable shared everyday-life activities across diverse

BE PART OF OUR VIBRANT COMMUNITY

Login Register

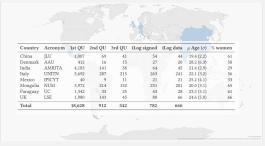
2:41 / 7:06

THE MAKERERE UNIVERSITY EXPERIENCE (UGANDA)

DataScientia Community GLOBAL

DS MAK
LOCAL

Mak2
Oct
2025



MAKERERE UNIVERSITY

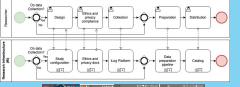


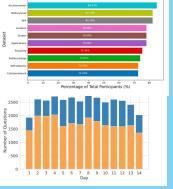
MalDoc: A Multi-Country Descriptive Dataset for Mobile Device and Internet Usage Metrics: An Analysis of College Students in Uganda

Abstract: This paper presents MalDoc, a multi-country descriptive dataset for mobile device and internet usage metrics. The dataset is derived from a survey of college students in Uganda, Kenya, and India. The survey collected information on various aspects of mobile device usage, including ownership, usage frequency, and usage patterns. The dataset also includes information on internet usage, including frequency of use, types of activities, and sources of information. The dataset is available for download and can be used for research purposes. The dataset is intended to provide a comprehensive overview of mobile device and internet usage among college students in three countries.

MalDoc: Behavioral Data of University Students' Smart Devices in Uganda

Abstract: This paper presents MalDoc, a behavioral dataset of university students' smart devices in Uganda. The dataset is derived from a survey of university students in Uganda. The survey collected information on various aspects of smart device usage, including ownership, usage frequency, and usage patterns. The dataset also includes information on internet usage, including frequency of use, types of activities, and sources of information. The dataset is available for download and can be used for research purposes. The dataset is intended to provide a comprehensive overview of smart device usage among university students in Uganda.





DS MAK LOCAL

Mak2 Oct 2025

Country	Acronym	1st QU	2nd QU	3rd QU	iLog signed	iLog data	μ	Age (σ)	% 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
Total		18,628	912	542	782	666			

RESEARCH ARTICLE



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 Bidoglio, George Gaskell,
 Altangerel Chagna, Amarsanaa Ganbold, Tsolmon Zundui, Carlo Caprini, Daniele Miorandi, Alethia Hume,
 Jose Luis Zarza, Luca Cernuzzi, Ivano Bison, Marcelo Rodas Britez, 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

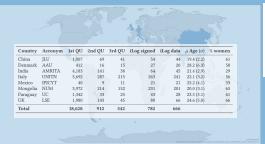
Nicolò Alessandro Girardini , [Simone Centellegher](#), [Andrea Passerini](#), [Ivano Bison](#), [Fausto Giunchiglia](#) & [Bruno Lepri](#)

[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 BUSO¹, ANDREA BONTEMPELLI², and LEONARDO JAVIER MALCOTTI³; University of Trento, Italy

LAKMAL MEDGAJAPOLA, ETH Zurich, Switzerland

PETER KUN, IT University of Copenhagen, Denmark

SHYAM DINGRAK⁴ and CHAITANYA NUTAKKI⁵, Amrita Vishwa Vidyapeetham, India

MARICELA DARIO, RODAS BRITZE⁶, University of Trento & IFPRI, Italy

HAO XU⁷ and DONGLEI SONG⁸, Jilin University, China

SALVADOR RUIZ CORRIDA⁹ and ANDREA RENICA MENDOZA-LARA¹⁰, Instituto Politécnico de Investigación Científica y Tecnológica, Mexico

GEORGE CASARELL, SALLY STABES, and MIRIAM BIDOLQUA, London School of Economics and Political Science, UK

ABARSANAA CANBOLD¹¹ and ALTANGEREEL CHAGNAA¹², National University of Mongolia, Mongolia

LUCA CERNUZZI¹³ and ALETHIA HUML¹⁴, Universidad Católica "Nuestra Señora de la Asunción", Paraguay

RONALD CHIENU ABENTE¹⁵, BOY ALIA ASIKU¹⁶, and IVAN KAYONAO¹⁷, University of Trento, Italy

DANIEL GATICA-PÉREZ¹⁸, Map Research Institute of EPFL, Switzerland

AMALIA DE GOTZEN¹⁹, Aalborg University, Denmark

FEDICO BISONI²⁰, University of Trento, Italy

FAUSTO GRINCCHIOLI²¹, University of Trento, Italy

Understanding everyday life behavior of young adults through personal devices, e.g., smartphones and wearables, is key for various applications, from enhancing the user experience in mobile apps to enabling appropriate interventions in digital health apps. Towards this goal, previous studies have relied on datasets combining passive sensor data with human provided annotations or self-reports. However, many existing datasets are limited in scope, often focusing on specific countries primarily in the Global North, involving a small number of participants, or using a limited range of personal sensors. These limitations restrict the ability to capture cross-country variations of human behavior, including the possibility of studying

“competing” culture.

Authors' Contact Information: Matteo Buso, matteo.bus@unitn.it; Andrea Bontemelli, leonardo.javier.malcotti@unitn.it; Italy; Lakmal Medgajapola, lmk@ethz.ch; Peter Kun, ptkun@itu.dk; Shyam Dingrak, shyam.dingrak@amritavp.ac.in; India; Maricela Dario, maricela.dario@unitn.it; Leonardo Javier Malcotti, leonardo.javier.malcotti@unitn.it; Italy; Hao Xu, hao.xu@jlu.edu.cn; Donglei Song, songdl@jlu.edu.cn; Salvador Ruiz Corrida, salvador.ruiz.corrida@ipicyt.mx; Mexico; Luca Cernuzzi, luca.cernuzzi@polito.it; Alethia Huml, alethia.huml@ucan.edu.py; Paraguay; Ronald Chienu Abente, ronald.chienu.abente@unitn.it; Boy Alia Asiku, boy.alia.asiku@unitn.it; Ivan Kayonao, ivan.kayonao@unitn.it; Philippines; Daniel Gatica-Perez, daniel.gatica-perez@epfl.ch; Amalia De Gotzen, amalia.de.gotzen@au.dk; Denmark; Fausto Grincholi, fausto.grincholi@unitn.it; Italy; Fedico Bisoni, fedico.bisoni@unitn.it; Italy; Faustino Grincholi, faustino.grincholi@unitn.it; Italy; and Andrea Renica Mendoza-Lara, andrea.renica.mendoza-lara@ipicyt.mx; Mexico.



This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International license.

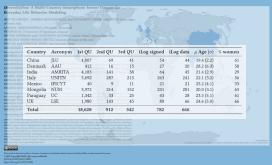
© 2023 Copyright held by the author(s). All rights reserved. Licensee MDPI, Basel, Switzerland.

Received: 10 August 2023; Accepted: 10 September 2023;

Published: 13 September 2023

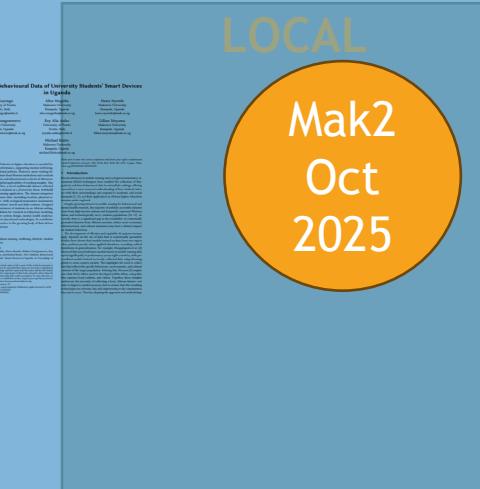
DS MAK
LOCAL

Mak2
Oct
2025



DS MAK LOCAL

Mak2 Oct 2025

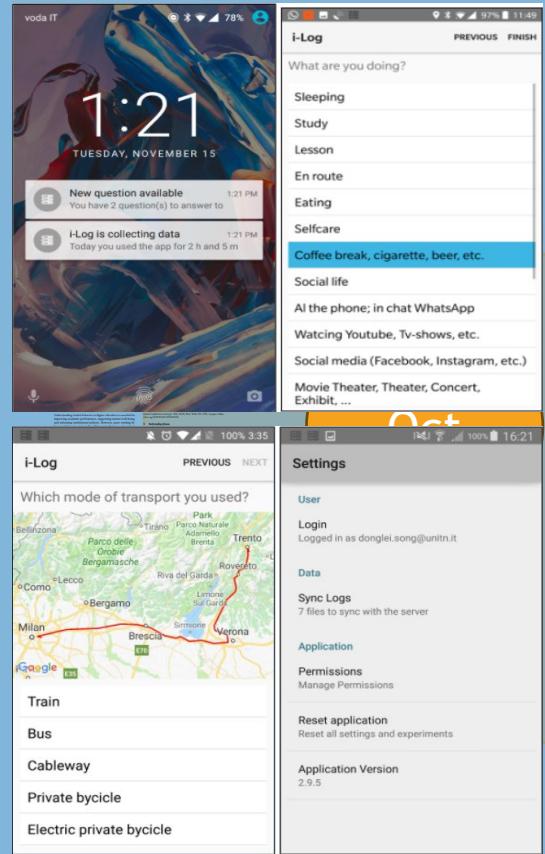
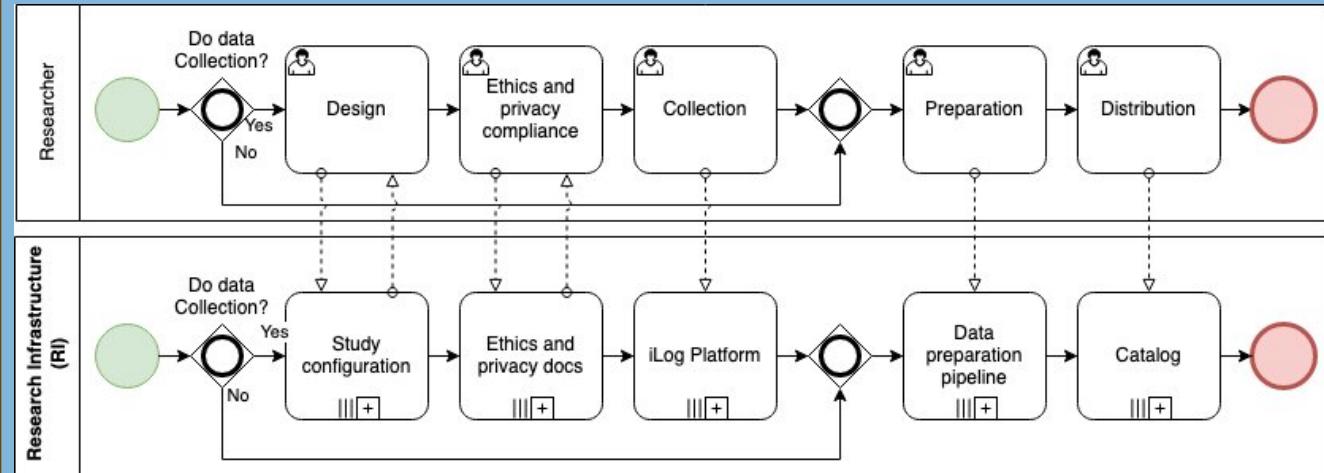


MalDoc: Behavioral Data of University Students' Smart Devices in Uganda

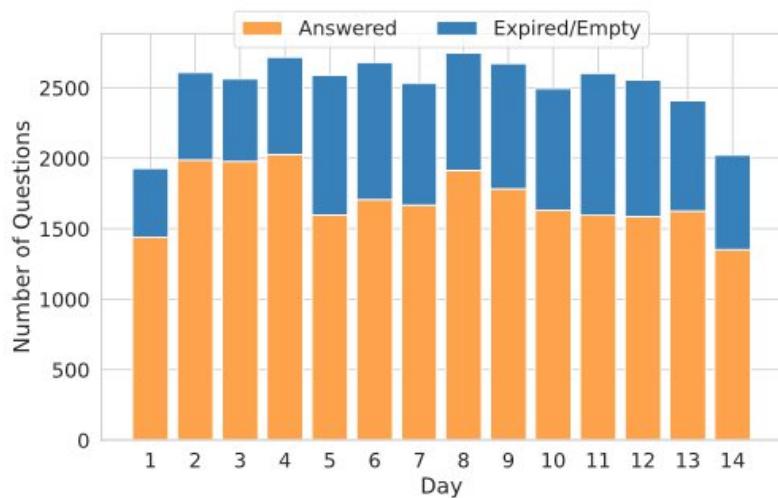
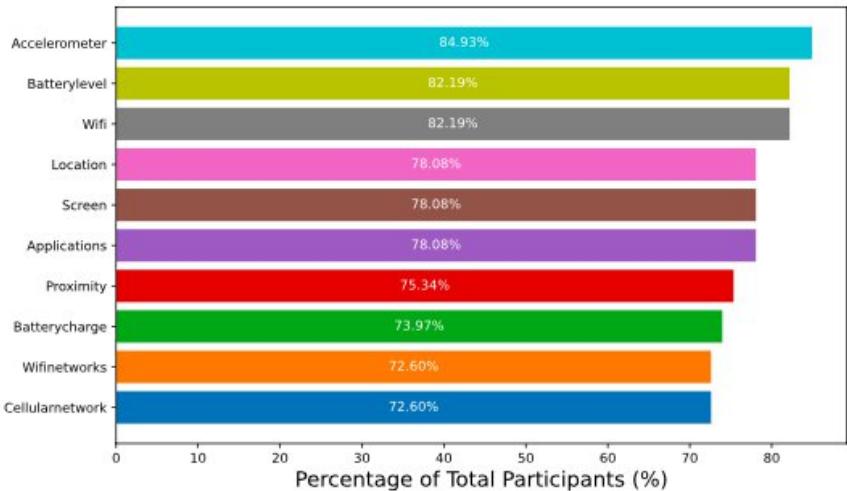
Abstract: This study aims to analyze the behavioral data of university students' smart devices in Uganda. The data was collected from 100 university students using a mobile application called MalDoc. The analysis focuses on the usage patterns of different apps, such as social media, messaging, and productivity apps, and their impact on the students' well-being. The results show that most students use their devices for social media and messaging, while fewer use them for productivity. The study also found that there is a significant correlation between the usage of social media and messaging apps and the students' mental health. The findings suggest that university students in Uganda are heavily dependent on their smart devices and that their usage patterns may have a negative impact on their well-being.

Keywords: University students, Smart devices, Behavioral data, Well-being, Uganda

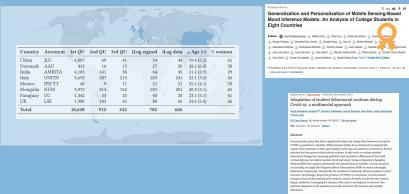
DataScientia Community GLOBAL



DataScientia Community GLOBAL



Scientific paper at ACM Ubicomp Conference



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

Ivan Kayongo
University of Trento
Trento, Italy
ivan.kayongo@unitn.it

Alice Mugisha
Makere University
Kampala, Uganda
alice.mugisha@muk.ac.ug

Hawa Nyende
Makere University
Kampala, Uganda
hawa.nyende@muk.ac.ug

Halima Chongonswero
Makere University
Kampala, Uganda
halima.chongonswero@muk.ac.ug

Roy Alia Asiku
University of Trento
Trento, Italy
royalia.asiku@unitn.it

Lillian Moyanya
Makere University
Kampala, Uganda
lillian.moyanya@muk.ac.ug

Michael Kizito
Makere University
Kampala, Uganda
michael.kizito@muk.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 socioeconomics and institutional contexts of African institutions, leaving the global landscape of modeling sparse. This paper introduces MakOne, a novel longitudinal dataset collected over six weeks from 72 students at a (University Name Withheld) using ILang, a mobile sensing application. The dataset integrates passive smartphone sensor data—including location, physical activity, and screen usage—with ecological momentary assessments (EMA) 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 behaviour, Africa dataset

ACM Reference Format:
Ivan Kayongo, Alice Mugisha, Hawa Nyende, Halima Chongonswero, Roy Alia Asiku, Lillian Moyanya, and Michael Kizito. 2025. MakOne: Behavioural Data of University Students' Smart Devices in Uganda. In *Proceedings of*

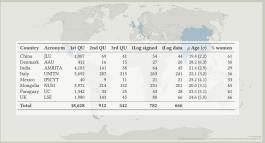
Promises to make digital and hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyright for works that can be reproduced by the author is reserved or registered in the copyright holder's name. Prior permission is granted by the author to reproduce 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 adopting the approach and methodology

Mak2
Oct
2025

DS MAK LOCAL

DataScientia 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)!!!



SCAN ME

COMMUNITY

News & Events

Monthly Newsletter

Data Marketplace

Interest Groups



EDUCATION

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

Data

Personal

Languages

Knowledge

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

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

- data that describes how we are.

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

- 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.



INNOVATION (WITHIN DS INITIATIVE)

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**

...