DATA SCIENCE LAW LAB

2024

END OF YEAR MESSAGE

research@datasciencelawlab.africa
www.datasciencelawlab.africa
LinkedIn: Data Science Law Lab
X: @datalawafrica

Dear DSLL community

As 2024 draws to a close, I wanted to share with you some of the work we've done to support innovation and the growth of all aspects of data science research, development and application across Africa.

This year we have focused mainly on driving and contributing to dialogues on innovation in the digital economy that is aligned with the African context and building digital legal tools that help create the Africa we need. We have shared and learnt a lot throughout this process. Here are some of what we have done, shared and learnt:



DIALOGUES, CONFERENCES & WORKSHOPS

This year we organised several key knowledge sharing and dialogue-driven events. In March, it was the first edition of the **University of Pretoria-Meta AI Policy Dialogue**, which presented a Triple Helix dialogue between academia, industry and government, on the current realities and future possibilities for responsible AI development and deployment in South Africa and across the African continent. In May, we brought speakers from academia, civil society and industry and diverse participants in a **Roundtable** to share information on the efforts, contributions, visions, challenges and plans of the triple helix in managing and addressing the relationships between AI, intellectual property rights and the SDGs. July saw us hosting Africa, Law & [Emerging] Technologies Policy Hackathon. We utilised the policy hackathon to brainstorm, develop, prototype and pitch creative policy solutions on two key issues: privacy of children online and privacy of health-related data. We published a **proceedings report** and a **photo-story** on this!

Apart from organising knowledge sharing and dialogue-driven events, we actively contributed to such events organised by some of our collaborators. In February, we hosted a **Fish Bowl dialogue session** at the third edition of the **Hundzula: Natural Language Processing and Linguistics Retreat** which held at Nelson Mandela University in Port Elizabeth, South Africa. Our session focused on legal issues around sharing of language data.



DIALOGUES, CONFERENCES & WORKSHOPS

In May, I joined an expert cast of speakers (including Prof Vukosi Marivate who leads one of our co-affiliates/collaborators, Data Science for Social Impact) to discuss and present on the AI in the Global Majority project of the Carnegie Endowment for International Peace. My presentation and intervention centred on a paper co-written with Prof Vukosi Marivate (University of Pretoria-ABSA Chair of Data Science) entitled "How African NLP Experts Are Navigating the Challenges of Copyright, Innovation, and Access". **The full article presented is available here.**

We were also at the annual conference of the International Association for Teachers and Researchers of Intellectual Property (ATRIP); UCT AI & Law International Interdisciplinary Conference; Deep Learning Indaba𝕏 South Africa;WIPO-WTO Colloquium for Intellectual Property Teachers, Academics and Researchers in Africa; the Intellectual Property Scholars Africa (IPSA) Conference; Deep Learning Indaba; CIPIT workshop on understanding the intersection between Artificial Intelligence and Intellectual Property; workshop on Improving NLP Evaluation for African Languages organised by the AI4D African Language Lab; etc. Look out for the compiled abstracts and presentation slides we shared for these events. Check out publications by our **Sakina Shabani**; **Takudzwa Matekwe** and **Benjamin Potgeiter** on our NOODL License and the Deep Learning Indaba!



NOODL LICENSE – NWULITE OBODO OPEN DATA LICENSE

At a **workshop we organised and held during the Deep Learning Indaba** (Dakar, Senegal), we formally launched the **Nwulite Obodo Open Data License [NOODL License]**! Nwulite Obodo Open Data License is created for those who want to share African datasets and those who want to use/reuse African datasets equitably, especially those from/in/of Africa. Among the unique features of the Nwulite Obodo Open Data License are a single tripartite license agreement involving a licensor/licensor group, licensees from developing nations and licensees from developed nations, an expanded view of pricing and exchanges for open African datasets, and a new mechanism – **African datasets creation split sheet** – that allows data contributors and providers to document their contribution to a given dataset.

Our work received recognition/acknowledgement in the **MIT Technology Review**!

We plan to publish version 1.1 in early 2025. This version has benefitted from feedback various quarters including conversations at the **UP<>Sweden AI Workshop Panel on Low Resource Languages and LLMs**; York University's **IP Osgoode Speaks**!; panellists at the **Data Futures Showcase at MozFest Zambia**; etc. We are truly grateful for the very helpful feedback and engagement with NOODL.

.



Licensing African datasets

FINDINGS & THOUGHTS

There are significant strides made by researchers in Africa who are working on African Natural Langauge Processing to ensure that African languages are not left behind in the AI revolution. Law and policy has a key role to play in determining and establishing the most promising and equitable approaches for developing that African language AI infrastructure.

Regulatory coordination in the AI context is key

The African Union released the **AU Continental Artificial Intelligence Strategy** in July, which identified five focus areas (i.e., harnessing AI's benefits for African people; addressing the risks associated with the use of AI; etc.) and 15 action areas such as the establishment of appropriate AI governance system, adoption of AI in the public and private sectors, etc. One of the key issues that requires thinking relates to regulation, in particular, regulatory actors and spaces. As a technology that finds uses in diverse contexts, many regulatory actors could lay claim to various aspects of AI. Careful thought must be given to regulatory coordination to ensure regulation does not stifle innovation.

FINDINGS & THOUGHTS

Text and data mining for data acquisition

We will continue to see different jurisdictions grapple with the challenges of existing laws (especially copyright laws) and licences in the context of data acquisition for training generative AI systems. This includes whether or not, in the context of African countries, crawling copyright-protected material across the internet to train AI systems is acceptable as fair dealing. With very few judicial precedent available on the subject, whether or not a dealing with copyright-protected material for AI training purposes is fair will depend on context. I have **argued that in the context of research** by university-based researchers and in relation to materials held by a public service broadcaster, dealing for AI training purposes would be fair. We will continue to engage with this issue in 2025.

Still on the subject of limitations and exceptions…When it comes to limitations and exceptions to enable research, the focus has largely been on (computational) research as an activity that engenders innovation. But, it is researchers who make (AI) innovation happen and whose contributions power the innovative outputs that addresses the challenges of our world. And so, ethical innovation must involve the recognition of, and appropriate support for, the position and status of the researcher in the (computational) research processes/pipeline. This leads to the question – How do/should we define 'researcher' in the context of the person or entity to enjoy copyright limitations and exceptions that support research?

FINDINGS & THOUGHTS

Open licenses need to be equitable

We have said that, "...addressing the restrictive and proprietary nature of these rules [Copyright and privacy rules] through openness does not and should not mean that openness is adopted without attending to the nuances of specific concerns, contexts, and people. In adapting openness to the nuances of the contexts of Africa (and the Global South), consideration must be given to the agency and autonomy of specific stakeholders to make decisions about the uses of their data contributions, created and annotated datasets, and the needs that AI tools and development are designed to address in the first place". .

Open licences offer one way of addressing this issue (legislation, competition, consumer advocacy are other ways). In our work on NOODL license, we **shared some of the ways** that the current/traditional data licenses have failed African communities and expressed some of the misconceptions about "open" that are important to dispel in the pursuit of a more equitable data economy. But as I said to some of our collaborators, it's one thing to have a license and yet another thing, to enforce it when there is contravention of the license terms. It's also another thing to have licenses that are compatible and interoperable with existing licenses that dataset (re)users have to adhere to . This bears careful thinking.

LOOKING AHEAD

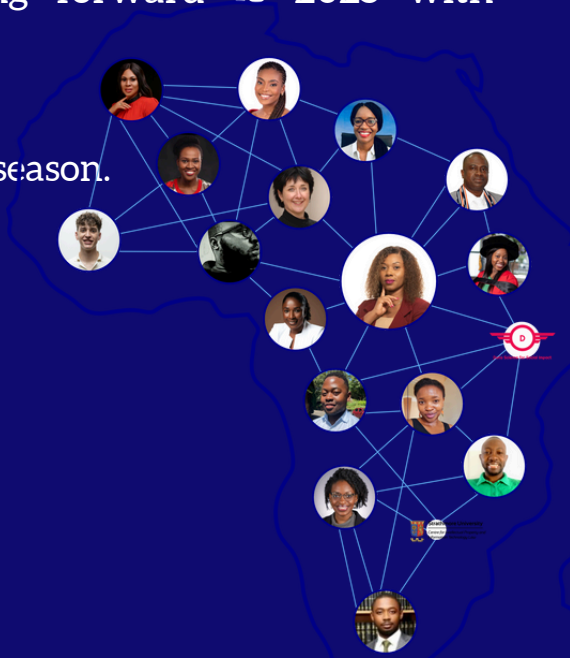
DSLL went from a research group at the University of Pretoria to a research network with an Advisory Committee and researchers from Ghana, Nigeria, Zambia, Senegal, Côte d'Ivoire and the US. Special thanks to our Advisory Committee members – Ms Nontobeko Mtshali; Dr Chinasa T Okolo, Mr Temple Uchegbune and Ms Ronee Robinson for accepting to be part of our work. We are confident that in this animated form, we will deepen the relevance of our contextual approach and significantly increase our research outputs and digital legal tools.

We have utilised various social media channels to reach and engage with our community and the public. On LinkedIn where we have been most active, our community has grown to 683 followers who engage with, support and promote our work! We hope that with dedicated communications support, we can continue to increase the visibility of our research and researchers.

This year, we received funding and in-kind support from Meta, Mozilla Foundation and the University of Pretoria. We are truly grateful to each of these organisations. We deeply appreciate you and every member of the DSLL Community for your support, encouragement, conversations and contributions to us and the work we do. While we will take the time to rest and enjoy interaction with our loved ones, we are looking forward to 2025 with excitement.

We wish you and your loved ones a most joyful holiday season.

Chijioke Okorie
Leader, Data Science Law Lab





Contact

www.datasciencelawlab.africa

research@datasciencelawlab.africa