ifa Institut für Auslandsbeziehungen

ifa Edition Culture and Foreign Policy

Democratising Artificial Intelligence through Culture

Making Generative AI Participatory and Intersectional Through an AI of the Commons

Seoyoung Choi and Miro Leon Bucher

Institut für
Auslandsbeziehungen

Democratising AI through Culture

Making Generative AI Participatory and Intersectional Through an AI of the Commons

Seoyoung Choi and Miro Leon Bucher

Contents

Abbreviations

AI	Artificial Intelligence
AISI(s)	AI Safety Institute(s)
CFG	Classifier-free Guidance
EU	European Union
GPU	Graphics processing unit
ICR	International cultural relations
LLM	Large Language Model
LoRA	Low-rank Adaptation (fine-tuning technique to adapt machine learning models)
MoE	Mixture of Experts
MSIT	(South Korean) Ministry of Science and ICT
SD1.5	Stability Diffuser 1.5
SDXL	Stability Diffuser XL (base model for machine learning from Stability AI)
UI	User interface
UK	United Kingdom
UN	United Nations
US	United States of America

Foreword

The technological race currently taking place in the increasingly competitive global environment raises questions on methods of data acquisition and amplification of certain worldviews through biases, further marginalising certain epistemologies while entrenching existing global power hierarchies.

The authors of this report argue for an "AI of the commons": a publicly-funded, community-led AI infrastructure which encourages participatory processes in shaping and influencing AI systems. To encourage the development of community-owned AI, the authors took an innovative approach focused on capacity building by developing a training toolkit. The toolkit guides lay users through the so-called "critical data injection" process, which aims to locally influence representation in generative AI, and is designed with the purpose of counteracting the biases and omissions which existing data sets collected through crawlers based on algorithms built by only a few.

As the authors state, "AI systems should be viewed and utilized as common cultural assets, designed not just for the public but also by the public." (p. 9) Artists and practitioners in the cultural sector could be key drivers for such a paradigm shift. Furthermore, through evidence collected in a field experiment conducted as part of this study, the authors argue that a co-constructive approach—a "global AI of the commons where AI data is generated through intercultural relations and exchanges" (p. 21)—is feasible. Inclusive AI is made possible through empowering societal participation and the sharing of perspectives.

This study forms part of ifa's Research Programme "Culture and Foreign Policy", in which experts address relevant issues relating to culture and foreign policy with the aim of involving academics, practitioners, policymakers, and civil society. I would like to thank Seoyoung Choi and Miro Leon Bucher for their excellent work and commitment to this research. In addition, I would like to thank my ifa colleagues Sarah Widmaier and Ivana Putri for their work on the conception, coordination and editing of this project.

Inclusive AI which encompasses the plural experiences and conditions of life at the greatest extent possible is important to prevent societal fragmentation and further polarisation. It is a vision for sustainability and more peaceful international relations.

Odila Triebel
Head of Dialogue and Research, ifa – Institut für Auslandsbeziehungen

Abstract

This research proposes a participatory, decolonial and feminist approach to AI through the concept of "AI as commons", offering a practical approach for cultural practitioners to decolonise existing AI systems. Conducted in Seoul, South Korea, the study combines critical theory with practical intervention, introducing "critical data injection" as a method for communities to influence representation in generative AI. At a workshop at Yonsei University, participants created datasets of everyday Seoul images which challenged typical AI-generated urban imagery, then trained Low-rank Adaptation (LoRA) models to generate alternative visual narratives. The accompanying LoRA Training Toolkit makes AI model training accessible to non-expert users. While acknowledging limitations, such as dependence on biased base models and infrastructure constraints, the research demonstrates how local communities can create representational alternatives within existing AI systems. The study argues for publicly-funded, modular AI infrastructure which enables democratic participation in AI development rather than relying solely on corporate or state-controlled systems.

Executive Summary

The Challenge

Artificial intelligence systems increasingly shape how we see and understand the world, yet their development remains concentrated in the hands of a few powerful corporations and nation-states. Current AI models reproduce existing biases around gender, race and cultural representation while extracting data from communities without consent or compensation. This is particularly evident in generative AI, especially in image generation, which tends to reproduce Western-centric, commercialised, including state-approved visual narratives while marginalising diverse lived experiences.

Our Approach

This research builds on the concept of "AI as commons", a participatory model where communities collectively create, govern and benefit from AI systems. Conducted in Seoul, South Korea, the researchers of this study combined critical feminist and decolonial theory with hands-on technical intervention. The researchers introduced "critical data injection" as a method for local communities to influence AI representation by training their own models using locally relevant datasets.

Methodology and Workshop

The research involved a three-day workshop at Yonsei University, bringing together academics, artists and IT professionals. Participants collected personal photographs of Seoul which challenged typical tourist imagery, tagged these images to reflect their lived experiences and used the LoRA (Low-rank Adaptation) Training Toolkit, developed by the researchers specifically for this study, to train AI models. This accessible toolkit, developed as part of and accompanying this report, enables non-experts to create AI image generation models without extensive technical knowledge.

Key Findings

The workshop demonstrated that community-driven AI training can produce significantly different visual narratives than corporate-developed AI systems. Where conventional AI generates a sanitized tourist imagery of Seoul—palaces, skyscrapers, clean modernist spaces—the community-trained models produced images featuring neighbourhood corners, protest banners, rainbow flags, children, urban wildlife, and other markers of lived experience typically absent from official city representations.

Participants described Seoul through their personal lenses which revealed the city's contradictions: spaces of both inclusion and exclusion, sites of memory and forgetting, areas where some groups of peoples' experiences are highlighted while others remain marginalised. These perspectives fundamentally challenged uniform narratives of Seoul as a globalised showcase city.

Limitations and Constraints

Our approach operates within existing AI infrastructure, relying on base models (Stable Diffusion) which contain the same biases we sought to counter. "Critical data injection" through LoRA training can influence but not fundamentally transform these underlying systems. Additionally, the workshop's small scale limits generalisability, though it demonstrates proof of concept for community AI agency.

Significance

This research offers a concrete alternative to both corporate AI monopolies and state-driven AI nationalism by demonstrating that local communities can meaningfully participate in shaping AI systems. While small-scale interventions cannot transform AI wholesale, they create essential spaces for alternative ways of seeing and being seen in increasingly AI-mediated visual cultures. This report provides both the theoretical framework and practical tools for democratising AI development, contributing to broader conversations about technology governance, cultural representation, and digital commons. It also proposes

several examples of measures for actors and institutions working in the field of international cultural relations to support the work towards making AI development more participatory and inclusive.

Policy Recommendations

- **Public AI Infrastructure:** Policymakers and international organisations should invest substantially in open frameworks enabling modular AI systems which integrate community-created models without requiring centralised base models.

- **Community Support:** Public bodies and the private sector should provide funding and infrastructure for local communities, cultural institutions and civic groups to collect, curate and annotate their own datasets.

- **Accessible Education:** Higher education institutions, think tanks and cultural institutions should provide capacity building through programmes and tools to enable non-technical participants to understand, train and audit AI models within their local communities.

- **Accountability Mechanisms:** In consultation with civil society and multilateral frameworks, policymakers should establish legal and technical systems allowing communities to identify and reject harmful AI components while promoting responsibility without censorship.

1. Introduction

Recently, an increasing number of research points to the biases and potential harms of AI for vulnerable populations such as women of colour, non-binary people, immigrants, and refugees. Concerns are also rising around the growing discrepancies between the Global North and the Global South, with the concentration of AI industries in the Global North as the forefront producers of AI innovation and the position of the Global South as mainly a consumer of such systems. With these apprehensions, researchers and AI practitioners have been calling for a feminist and decolonial imagination and ethical (re-)consideration of AI technologies.

This research aligns itself closely with ifa's recent research on AI, in particular with Octavio Kulesz's "Artificial Intelligence and International Cultural Relations" (2024). Where Kulesz gives an extensive overview of the field of AI and its relation to international cultural relations (ICR), this project aims to particularise on the dimensions of intersectionality with decolonial feminist approaches while implementing key aspects of Kulesz's recommendations.

For this, we propose an argument for a form of participatory, decolonial feminist AI. The argument is based on the idea of the commons and that AI systems should be viewed and utilised as common cultural assets, designed not just for the public but also by the public, to incorporate and prioritise often underrepresented perspectives.

Other policy and research initiatives in the field of AI as commons include Open Future and their "AI and the Commons" platform.[1] We aim to expand these efforts with ethnographic approaches to art and the commons, as elaborated by anthropologist Massimiliano Mollona in his book "Art/Commons: Anthropology beyond Capitalism" (2021). Through his participatory observation, Mollona introduces case studies of art-related 'commoning' in which art becomes a situated, collaborative practice where knowledge is co-created through participation, and where creative work sustains the commons rather than serving capital. This translates into our understanding of the conducted workshop, in which we see AI models as an engaged co-creation

[1] https://openfuture.eu/our-work/ai-and-the-commons (accessed on 14 August 2025).

process outside of market interests but as a site for commoning. Similar to Mollona, we use the workshop as a quasi-ethnographic case study from which to derive a broader understanding of AI as a commons.

1.1 Research Methods

As Kulesz (2024, p. 77) and his interviewees point out, there is not only a demand for creating systems which lead to more equitable forms of AI but also a high demand for communicating the workings, issues and potentials of AI to ICR institutions, researchers of other fields, artists, and the public in face-to-face events. Therefore, we conducted a workshop as part of our research project. The workshop was held at Yonsei University, Seoul, South Korea, bringing together members of the ICR community such as artists, academics, activists, and models[2] for image generation to demonstrate that local AI interventions through locally produced data and an AI of the commons which is open source and interactive are feasible (further elaboration can be found in Section 3.1.2).

Similar projects include "Inteligência Artificial, Arte e Indigeneidade" (AI/AI, 2024), where researchers and indigenous communities came together to create AI-generated images from the perspective of indigenous people. However, they used a closed model (MidJourney) to generate the images. While this allowed for critical reflections, it did not allow for adjusting the model. Machine learning design researcher and artist Caroline Sinders is conducting an ongoing art project, "Feminist Data Set"[3], which creates feminist data from a "critical and artistic view on software, particularly machine learning" through public workshops and forums (Sinders, n.d., p. 5). In our workshop, we expand this approach with a decolonial feminist perspective and not only collect data, but also turn this data into locally usable models for diffusion-based image generation[4]. To our knowledge, there has not been any research yet which

[2] The term 'models' used throughout this report refer to 'machine learning models'.

[3] https://carolinesinders.com/feminist-data-set/ (accessed on 31 July 2025).

[4] Diffusion-based image generation is one of the major approaches in the field of generative AI. One of the notable diffusion-based image generation models is Stable Diffusion, incorporating the

connects the idea of an AI of the commons with actually bringing AI models to people without extensive AI experience and enabling them to do 'critical data injections'[5] in meaningful ways on a local and participatory level.

In making critical interventions to AI, it is crucial to incorporate feminist and decolonial approaches to make these systems more inclusive. There are multiple definitions of what makes an approach 'feminist' and 'decolonial'. In this context, we understand a feminist approach to AI as an approach which takes a critical stance toward the gendered inequalities which could be reproduced in our understandings and use of AI. This includes not only the gender inequalities which AI-generated data could reproduce (e.g. AI reproducing a certain image of femininity and masculinity), but also how AI could generate gender binaries (e.g. AI reproducing heteronormative data).

A decolonial approach to AI problematises how AI-generated data can reproduce racial inequalities through their bias. It also problematises how AI data generation and the AI industry operate in a global capitalist order which reproduces the logic of imperialism and colonialism through, for example, the division of labour between the Global North and the Global South. With these perspectives and through ownership of tools and platforms, we aim to take an intersectional approach which considers the multiple layers of discrimination and inequalities which could be (re)produced through and by AI.

Together, feminist and decolonial approaches further allow us to question how AI industry and technology worldwide operate by reproducing different binaries and boundaries—such as binary ideas of the world consisting of developed and developing countries, highlighting the inequalities between them—which marginalise certain countries or groups of people(s) in the global AI market and in access to AI technology.

---

underlying diffusion process in its name. The name is derived from thermodynamics and is introduced and explained in a paper by Sohl-Dickstein et al (2015).

[5] We use the idea of 'critical data injection' to describe the process of adding data to an existing AI model. This way, local actors can introduce their own critical data to work against biases in the already existing AI model. We discuss this in more detail in Section 3, "Critical Data Injection for AI of the Commons".

Concerns on how biased AI can be and to what extent it can reproduce the pre-existing structural inequalities have been raised by decolonial scholars and critical AI scholars concerned with the risks of AI reproducing structural discrimination (see Benjamin, 2019; Chaka, 2022; Drage and Mackereth, 2022; Heaven, 2022; Johnson, 2022; Lohr, 2022; Olson, 2022; Siapera, 2022; Elam, 2023; Hampton, 2023; Iyer et al., 2023; Gengler, 2024, among others). Critical studies of data and technology have moved towards a 'decolonial turn' by problematising and contextualising the history of technology and contemporary data practices in the history of capitalism (Couldry and Mejias, 2019 & 2021; Srnicek, 2017; Zuboff, 2019). These scholars propose the frameworks of data colonialism (Thatcher et al., 2016; Couldry and Mejias, 2019 & 2021) and data extractivism (Mezzadra & Neilson, 2019), studying how, in an era of rampant datafication, lives are constantly translated into data and information is extracted to feed enterprises, as well as how colonial logics and practices of appropriating land, bodies and resources are reproduced by data practices.

1.2 Structure

This report is structured as follows: we first contextualise South Korea, the site of our research, as a complex site from which to contemplate on possibilities for decolonial feminist approaches to AI by situating South Korea's engagement with AI within the larger global contexts. In doing so, we come up with sets of critical questions on AI. We do not aim to find definite answers to these questions; rather, they serve as different entry points for future conversations and critical interventions to AI. Based on that, we introduce our attempt to make critical interventions through the practice of 'critical data injection' and a model of 'AI of the commons' which encourages a participatory use of AI.

We aim to facilitate such participatory engagement with AI by providing a toolkit: the Low-rank Adaptation (LoRA) Training Toolkit. The toolkit aims to make the process of training one's own AI model for image generation accessible to non-expert users. Next, we showcase how the use of the toolkit was put into practice through our workshop at Yonsei University, Seoul, South Korea, with

members of the ICR community, including artists and beyond. We conclude the report by summarising the implications of our interventions, by reflecting on their limits and by providing policy recommendations for a more ethical and sustainable use of AI.

2. South Korea as a Site to Cultivate Feminist and Decolonial Approaches to AI

From their development to their safety, the flaws demonstrated by current AI systems have become a contemporary global concern. While different governments have seen the potential of economic growth and transformations in industries in and through AI, they also have been increasingly concerned with the issue of AI safety and with the need for international cooperation on the matter. As a result, many national governments are seeking a form of global AI governance. This is demonstrated among others through the AI Safety Summit, which are part of the first series of actions that national governments have taken to systematically collaborate on regulating the safety of advanced AI systems.[6]

The First AI Safety Summit was hosted by the United Kingdom in 2023.[7] As a concrete outcome of the summit, twenty-nine countries[8] signed the Bletchley Declaration on AI safety, agreeing to promote "safe, human-centric, trustworthy, and responsible AI" [9]. AI Safety Institutes (AISIs) were created following the First Summit in signatory countries including the United Kingdom and the United States as an attempt to lay the foundations for international AI governance. Though limited in their role of regulating AI, AISIs cooperate with governments and industry on regulatory strategies, AI safety research and model evaluation.

---

[6] The biannual AI Safety Summits build on the work that has already been done at the OECD, the Global Partnership on AI, the Council of Europe, and the Hiroshima AI Process launched by G7, which aim to promote safe, secure, and trustworthy AI through international regulations for advanced AI systems. The AI Safety Summits are hosted by national governments.

[7] https://www.gov.uk/government/topical-events/ai-safety-summit-2023 (accessed on 18 August 2025).

[8] South Korea was among the 29 countries which were represented at the First AI Safety Summit of 2023. The countries represented include Australia, Brazil, Canada, Chile, China, the European Union, France, Germany, India, Indonesia, Ireland, Israel, Italy, Japan, Kenya, the Kingdom of Saudi Arabia, the Netherlands, Nigeria, the Philippines, Republic of Türkiye, Rwanda, Singapore, Spain, Switzerland, Ukraine, the United Arab Emirates, the United Kingdom, and the United States of America.

[9] See the following link for further information on the Bletchley Declaration on AI Safety: https://www.gov.uk/government/publications/ai-safety-summit-2023-the-bletchley-declaration/the-bletchley-declaration-by-countries-attending-the-ai-safety-summit-1-2-november-2023 (accessed on 16 August 2025).

The Second AI Safety Summit, the AI Seoul Summit, was co-hosted by South Korea and the United Kingdom in May 2024.[10] Ten countries and the EU promised to fund AI safety institutes and an international network for strengthened coordination, reinforcing the idea of global AI governance with discussion around safety, innovation and inclusivity. Since the Second Summit, South Korea, Japan, Singapore, France, and Canada have launched AI safety institutes. The EU established its AI Office to serve such a function.

The Third AI Safety Summit, AI Action Summit, was hosted by France in February 2025.[11] The French President Emmanuel Macron had shown his ambition for France to become an AI leader by announcing on the day before the Summit an investment of 109 billion EUR in private AI infrastructure. The European ambition to become a leading power in AI was followed and backed up by the launch of the EU AI Champions Initiative, funded publicly (58 billion EUR), and by 70 European enterprises' investment of 150 billion EUR in AI research and development. According to several policy analysts such as Torchio and Tasin (2025), the Third Summit, which aimed to discuss the safety and risks of AI alongside the opportunities in AI development, was a failure, as the Summit shifted from its initial commitment to safety and global governance to the development of and competition over AI, with the US and the UK withdrawing from the safety commitment they had pledged over the first two Summits.

Though recent efforts around global AI governance exist, such as the AI Summits mentioned above and international efforts for digital global cooperation including UNESCO's 2021 Recommendation on the Ethics of Artificial Intelligence (the first global standard on AI ethics)[12] and the UN

[10] https://aiseoulsummit.kr/aiss/ (accessed on 16 August 2025).
[11] https://www.elysee.fr/en/sommet-pour-l-action-sur-l-ia (accessed on 16 August 2025).
[12] https://www.unesco.org/en/artificial-intelligence/recommendation-ethics (accessed on 16 August 2025).

Summit of the Future 2024[13] which also addresses AI, the growing competition between countries over AI development raise concerns over how AI is going to shape the globe and its ethical implications. In this so-called 'race', certain countries, like the United States, are better positioned to influence how regulations and ethics in using AI are shaped and enforced. Others are left with little or no space and time to consider these in their AI development, as they are pressured to keep up in the global AI 'race' despite having relatively limited infrastructural conditions to meaningfully compete in the global AI market. Some countries are not considered as a place for AI development but as a source of data. For example, the precarious labour required for AI development, such as data labelling, content moderating, warehouse work, and delivery work, have been outsourced to Kenya. Labour exploitation of data workers and content moderators (whose tasks include monitoring violent and sexual content) have been reported over the past few years (Perrigo, 2023). These underpaid workers toil under surveillance software and suffer from trauma and anxiety from the content they are exposed to all day (Williams et al., 2022). Birhane (2020) describes this in the context of the African continent as a new form of "algorithmic colonization", which leads to the call for a "decolonial AI" (see also Hao, 2022).

With this background in mind, the Republic of Korea is a significant site for this research project. As lawyers Ko, Lee and Mok (2024) put it, South Korea has been working on an AI regulation[14] based on the principle of "adopting technology first and regulating later", with much controversy and many social debates on the necessity of AI frameworks which consider safety and transparency. This showcases the perceived need for countries outside the EU, the USA and China[15] to keep up with AI developments by disregarding the need

[13] The outcomes of the UN Summit of the Future 2024 are presented in the Pact for the Future, Global Digital Compact and Declaration on Future Generations: https://www.un.org/sites/un2.un.org/files/sotf-pact_for_the_future_adopted.pdf (accessed on 16 August 2025).

[14] The so-called "basic bills on AI", as will be explained further on pp. 21-22 of this report.

[15] In early 2025, the Chinese AI startup DeepSeek, founded in late 2023 by Chinese hedge fund manager Liang Wenfeng, launched a new open-source AI model, shaking the global AI and stock

to create and implement stronger regulations for more safe and ethical ways to develop AI. AI development has become *the* industry which defines whether a country is technologically and economically 'advanced' or 'developed'. In this global competition, the discrepancies between the Global North and the Global South are laid bare.

The idea of a country being 'backward' or 'developing' as opposed to being 'developed' reproduces (post)colonial conditions which shape how the global order is understood. Here we interpret 'coloniality' in its broader, global context, going beyond merely having colonial ties. Given the influence of colonialism worldwide, coloniality is a condition which continues to be reproduced and lived in present times, not only by former empires but also by postcolonial states which aspire to become developed and, in this process, perceiving the world and people in hierarchical terms (Rutazibwa and Shilliam, 2018). This applies to South Korea in locally-specific ways. As a latecomer in the system of modern nation-states, South Korea has been modernising—including westernising and globalising—rapidly to overcome the idea that it is a 'backward', developing country (Seo and Cho, 2021). South Korea is a postcolonial country which has experienced colonial rule under the Japanese empire and, after liberation, been caught in neocolonial relations with the United States. That South Korea situates itself at present as a 'high-tech country' and brands its culture and its capital city, Seoul, as a 'high-tech city' is not irrelevant to such an attempt to become or to be 'developed' and 'advanced'.

Within such a global context, South Korea is at once in a hurry to surpass other countries in the race for AI technology development and have its own sovereign AI (as will be explained in the following paragraphs) by minimising regulation to incentivise rapid growth in this area. However, South Korea is also in a place where it has to consider the social impact of its brisk implementation of AI

markets as well as the tech world with its significantly low cost of operation. Following its launch and its incredible popularity, the rise of DeepSeek sparked concerns in the United States over the 'invincible' American technology industry being 'threatened' by its rival, China (Goldman, 2025). In addition, under the general assumption that generative AI from a democratic country would be uncensored, there was much curiosity and critique over whether a Chinese generative AI would be 'really' free from censorship (cf. Lu, 2025).

technology, especially concerns over human rights and ethics in relation to the development and usage of AI.

South Korean tech companies are actively participating in the race to create a sovereign AI[16] and achieve data sovereignty. South Korea's largest search engine, Naver, became one of the most prominent players in the market by developing its Large Language Model (LLM) HyperCLOVA X, specialised in the Korean language and trained on Korean datasets. With HyperCLOVA X—which is trained with 6,500 times more Korean data than OpenAI's GPT-4— Naver aims to provide a multilingual AI which is exceptionally sensitive to Korean culture and social norms and could generate "responses that resonate with Koreans", promoting an "AI that can see, hear, and speak like a Korean".[17]

What does it then mean to have an AI which can see, hear and speak like a Korean? Who among Koreans is HyperClova X closest to, in terms of gender, age, class, ethnicity? The Korean online database (i.e. Korean online search engines, news platforms, Korean social media platforms, and so on), which forms the data source for developing Korean sovereign AI, is relevant to discussions on sexist and racist AI, especially considering the relatively highly misogynist online space in Korea. The digital space also plays a prominent role in the Korean feminist movement and its popularisation. It is a site where young women challenge prevalent misogynist language, sometimes confronting it in ways perceived as 'misandry' by traditional Korean society (Seo and Choi, 2020). This extends to the domain of AI, as online communities of university students have been accused of using AI-generated content to promote misogynist ideology (Kim, 2020).

These Korean online (sub)cultures and the way their discourses are reproduced in the media and everyday life make the Korean digital space an important field of research for feminist scholars. Feminist scholars' concerns about AI include questions such as:

[16] "Sovereign AI" refers to a nation's capability to develop, deploy and control independent AI systems without depending on foreign technology, infrastructure, or data.
[17] https://clova.ai/en/hyperclova (accessed on 15 April 2025).

> *"Can AI be feminist? Who makes technology and for whom? When and how should feminists resist new and emerging AI-powered technologies? How do the gendered effects of AI intersect with other forms of power and control, such as ethnicity, race, class, disability, sexuality, and age? How are feminist movements currently using data and for what ends? Can AI infer sexuality? What identities does AI produce? What would a queer technological future look like? How might we begin to look past the liability generated by AI to also consider its much wider structural dynamics?" (Browne et al., 2023, p. xii)*

Gender Studies experts have criticised the lack of ethical concern in AI regulation in South Korea. In recent years, gender bias in AI systems and applications developed by Korean companies elicited public apprehension (Oh, 2023). In 2019, Womenlink, one of the largest feminist activist associations in South Korea, voiced concerns about the gender bias of Giga Genie, an AI speaker developed by Korea Telecom. These concerns arose following a series of experiments in which Giga Genie replied to the question "Do you like cars?" by responding, "I have no interest in cars, probably because I am a woman" (Lim, 2021). In 2021, the open-domain conversational AI chatbot Lee-Luda developed by the Korean start-up Scatter Lab sparked further controversies over the need for appropriate AI regulation, highlighting two critical issues: unethical data collection practices and harmful AI outputs. The chatbot, which assumed the persona of a 20-year-old female college student, was trained using deep learning on real Korean couples' private conversations from the messenger app KakaoTalk. Scatter Lab had collected these conversations through their relationship advice app Science of Love, where users voluntarily uploaded their KakaoTalk conversations to receive dating guidance, but the company used this data to train Luda without users' explicit consent to do so for AI training purposes. The resultant AI then made racist, sexist, misogynistic, and homophobic statements and enabled some users to generate sexual abuse material, demonstrating how inadequate data governance and insufficient content moderation can compound to create significant harm (Choi and Hong, 2021). Outbreaks of deepfake sex crimes in Korean universities, high schools and

middle schools in 2024 raised further concerns over the ethical use of AI (Mackenzie and Choi, 2024).

In 2020, the so-called "basic bills on AI" were first tabled at the South Korean National Assembly to support and promote the research and development of AI. However, the lack of provisions for ethical considerations raised concerns among Korean civil society. In June 2023, amid the national enthusiasm for AI development, the Ministry of Education of South Korea presented its "AI Digital Textbook Promotion Plan", which aims to use AI digital textbooks in Korean schools for personalised learning (Ministry of Education, 2023). This plan sparked controversies and conflicts among the Ministry of Education, teachers' associations and parents, with concerns over the effectiveness of AI textbooks, the current infrastructure in schools and the biases in AI which students might learn. The South Korean progressive organisation of lawyers, Minbyun-Lawyers for a Democratic Society, argued that AI textbooks are against the Constitutional Law, as they degrade students' basic rights to education, possibly disrupting the learning environment due to decreased in-person interactions between teachers and students and the risk of AI-generated curricular misinformation (Hong, 2025). Concerns over national education grew among the public when it was reported that an AI textbook had claimed that Dokdo (an island in the East Sea which South Korea considers to be Korean territory) is "subject to territorial dispute" (with Japan), as framed by the Japanese government (Kim, 2024).

After four years of consideration, the Basic Act on the Development of Artificial Intelligence and the Establishment of the Foundation for Trustworthiness ("AI Basic Act") was passed in South Korea in December 2024. By doing so, South Korea became the second political entity, after the European Union, to establish a comprehensive legislative framework which governs AI. A press release by the South Korean Ministry of Science and ICT (MSIT) states that:

> *"As artificial intelligence (AI) has become so important that it can influence not only technology but also a nation's economy and security, major countries are placing their utmost efforts into fostering AI innovation and establishing international leadership in*

AI, while creating regulations favorable to their own interests. In response, Korea, under the shared understanding that a foundational law is needed to enhance national AI competitiveness and build a trustworthy AI utilisation infrastructure, has decided to pursue the enactment of the AI Basic Act, considering domestic economic and social conditions, making it the second country in the world to do so." (Ministry of Science and ICT, 2024)

The newly passed Act demonstrates South Korea's ambition to excel in the competition for AI, framing AI as a national economy and security issue, similarly like other countries which promote and invest in AI development. While the Act includes measures for AI safety and aims to address AI risks, secure transparency and protect individuals, the law does not regulate online platforms where risky AI-generated content, such as deepfake videos, can be circulated. There is a lack of both national and international consensus on which type of content and data should be regulated as well as how and for whom governments should engage with AI, and thus these questions remain insufficiently discussed amid the global 'AI frenzy' (such consensus may even be impossible).

In this context, what would (or should) a non-Western, sovereign AI, as promoted by companies like Naver, look like? Can a sovereign AI avoid reproducing colonial logics and practices in their data systems? The challenge for sovereign AI projects is not simply avoiding Western data sources, but confronting how colonial processes have shaped local knowledge systems and social hierarchies. Korean online misogyny, while expressed through distinctly Korean cultural forms and language, operates within patriarchal structures which were reconfigured through Japanese colonial rule and subsequent US neocolonialism. As decolonial feminists argue, colonialism and patriarchy function as mutually constitutive systems which shape and sustain each other (e.g. see Lugones, 2010, among others).

Can a sovereign AI which draws from Korean online discourse, including its highly misogynistic spaces, truly decolonise AI representation or does it risk reproducing both local and globally-rooted forms of discrimination? This concern extends to the workings of the global AI market, where even "sovereign"

projects remain embedded in exploitative labour practices, such as the outsourcing of data labelling and content moderation to workers in the Global South. When a country such as South Korea seeks to create a sovereign AI, where does this project stand in the global distribution of labour and power?

The intersection of local patriarchal structures with colonial legacies complicates the question of how "histories of racial violence and colonial control continue to influence who counts as human, and which bodies are considered worthy of protection" (Browne et al., 2023, p. xv). The current situation points towards the need to take an intersectional approach to AI (also see Ulnicane, 2024).

In this research, we problematise the current capitalist system in which big tech firms predominantly based in the United States and China are in control of the global AI market, regardless of the existence of a sovereign AI or lack thereof. While various actors in tech companies, politics, education, and many other sectors often talk about 'AI as the future' or the 'fourth industrial revolution', it seems that not everybody understands what these terms really mean. There is a huge knowledge gap between AI technicians working for big tech companies and most users/consumers of AI, particularly when it comes to understanding how AI works and what it means for AI to generate information or images out of data. Among lay users, there is a lack of understanding of what happens when one interacts with AI, what open-source AI means, and so on.

This research is an attempt to tackle these questions through community-based approaches and practices in understanding and engaging with AI. Though we do not believe that this attempt will change the systems of monopolistic techno-capitalism and of nation-states in which AI operates, we seek to create space for communication of a different engagement with AI across cultures and for feminist and decolonial purposes. Specifically, this means engaging different communities in civil society to open up the tools and platforms for AI-generated data and create other patterns to disrupt the currently heavily biased data sets which regurgitate racist and sexist tendencies.

3. Critical Data Injection for AI of the Commons

3.1 Background and Approach

Through this research project, we propose a model of an AI of the commons which combines technology-centred approaches, such as Open Future's "AI and the Commons"[18], with anthropological and ethnographic approaches to art and the commons, such as those elaborated in anthropologist Massimiliano Mollona's book "Art/Commons: Anthropology beyond Capitalism" (2021). Based on these approaches, we argue that there should be a public infrastructure not only for creating datasets and AI models but also for AI applications which are being made available to the public. [19] These infrastructures and the institutions operating them should be governed by boards representing different communities in society. This allows for societal participation while working against governmental exploitation of AI systems (Feldstein, 2019) and the concentration of power of tech companies (Bremmer and Suleyman, 2023). As security technologist Bruce Schneier and data scientist Nathan Sanders put it, "AI will impact the public like few other technologies, so it should also be developed by the public" (Schneier and Sanders, 2023).

Investing public funds in AI development is an idea which is starting to be explored, as shown by several initiatives in a few countries where government funding is currently being allocated to more open and transparent AI models[20], such as BLOOM[21], which is supported by the French research agencies Centre National de la Recherche Scientifique (CNRS) and Grand Équipment National

[18] https://openfuture.eu/our-work/ai-and-the-commons (accessed on 14 August 2025).

[19] The term 'AI model' refers to the pure model, whereas the term 'AI applications' refers to user interfaces, such as mobile apps or websites which let users interact with the AI model more intuitively. For example, GPT-5 may be the current model trained by OpenAI, but ChatGPT is the web application which lets users use the model in the form of a chatbot. If AI is supposed to become public infrastructure, it needs both components to be usable for society at large: the models based on ethically sourced datasets and applications which allow for easier user interaction with the models.

[20] Closed models typically do not disclose their training process or data, instead allowing users to interact with the model in very narrow ways, e.g. through a specific web service. We use transparency and openness as terms of degree, describing how much of the training data, training process and final AI models is made publicly available.

[21] https://bigscience.huggingface.co/blog/bloom (accessed on 30 July 2025).

de Calcul Intensif (GENCI) (BigScience Workshop, Le Scao et al., 2022). However, there is less effort put into creating an alternative system which provides the public with an ethically sourced form of AI and compensates them for their direct or indirect contributions in the form of public knowledge. As a first step towards achieving this, our proposal promotes participatory modes of data gathering.

Currently, it is common practice for AI-model training to be based on web crawlers, such as Common Crawl[22]. They scrape the internet for information, which is then used to train the AI model. This suggests that the information many prominent AI models use is not actually based on "data" in the Latin sense of "datum," meaning something "given", but rather "capta", meaning something "taken" (Drucker, 2011; Nguyen, 2024; see also Kulesz, 2024). In our approach, we are seeking a system which produces AI models which use 'data' as genuinely 'given' information directly provided by members of society.

In addition, we consider the possibility of a 'global AI of the commons' where the result of the generated AI data is gathered through intercultural relations and exchanges. As there is no consensus on what inclusive AI would look like, considering a 'global AI of the commons' entails questioning what it would mean to aim for such a project. When AI is claimed to be "biased", does this mean the same thing in all regional and cultural contexts? What counts as 'bias'? What does it mean to make AI more "equitable", considering the different interpretations and translations one would get, depending on certain regions and cultural groups? In the globalised and connected world we live in today, these questions point to the continuous need to have open-ended conversations on what is meant by a 'more equitable form of AI'. What is the 'given' information and how is it interpreted differently in different contexts?

Our proposed model aims to facilitate feminist and decolonial interventions in the data gathering and model training processes and leads to what we currently refer to as **'critical data injections'**. The term 'data injection' itself reveals some of the biases of the discourse surrounding AI. In the literature, 'data injection'

[22] See https://commoncrawl.org/ (accessed on 17 August 2025).

implies security breaches, like "false data injection" or "data poisoning," where AI systems are compromised by misleading information (Sarp et al., 2021). Instead, we redefine 'critical data injections' as the addition of diverse, reflective data to AI systems to counteract biases.[23]

The workshop which was conducted as part of this research centred around the idea of demonstrating critical data injections in tangible ways. While it is not possible to train "foundation models"[24] (Kulesz, 2024, p. 18) in the framework of a workshop, it is very feasible to demonstrate the potential of what we refer to as 'critical data injections' in the form of Low-rank Adaptation (LoRA) models.[25]

In the workshop, we presented and discussed our proposal for an AI of the commons and critically questioned the biases and issues of current AI models. We had an open-ended conversation on diverse topics which included, but were not limited to, AI analysed through the lenses of coloniality, questions of AI safety, especially concerning 'deepfake porn' crimes, and AI ethics. Based on these discussion points, the workshop participants took part in critical data injection, a method of de-biasing the AI model by adding specific kinds of image data to it. The participants collected and created their own image data and sorted and tagged this data to prepare the dataset for the model training. We then trained the participants' LoRAs, compared the results with and without the LoRAs and discussed the changes and lessons learned from this process. The

[23] This may have some overlap with model fine-tuning, as discussed by Kulesz (2024, p. 26). However, it is not limited to the fine-tuning process and can also apply to the data collection and selection processes, as demonstrated in the conducted workshop.

[24] "Foundation models" are a type of AI system which are able to perform multiple general tasks like creating or summarising texts, manipulating images or generating audio, among others. They are trained on huge datasets containing texts, audio, video, and other files from the internet. An example of a foundational model is OpenAI's GPT-4, which forms the basis of ChatGPT.
For more information: https://www.adalovelaceinstitute.org/resource/foundation-models-explainer/ and https://hai.stanford.edu/news/what-foundation-model-explainer-non-experts.

[25] Low-rank Adaptation is a technique for adapting machine-learning models for specific uses in or for new contexts (Noble, 2025). LoRAs are used in diffusion-based image generation to extend foundational models (i.e. expand their function, the datasets they contain, among others). They cannot replace them, but have the ability to skew and add image content and styles to the foundational model.

image data created through this workshop is to be understood as an experiment which does not directly influence commercially available AI models but rather encourages participants to understand how AI models work and how they could intervene in data production.

It is important to reflect on the fact that we have to rely on a foundational model, in this case, Stable Diffusion. However, as feminist AI/technoscience researcher and Professor Emeritus at Sheffield Hallam University Alison Adam proposes, "[...] feminist projects can construct systems built with traditional AI technology but where the knowing subjects and the limitations of the knowledge are made explicit [...]" (Adam, 1995, p. 414). This aligns with what Alicja Peszkowska (former Engagement Lead at Open Future) and sociologist Alex Tarkowski (Director of Strategy at Open Future) talk about in their conversation with the co-founder of the French private AI lab PlelAs, Pierre-Carl Langlais, when they conclude that "[f]ine-tuning is an important element of a vision for AI as commons, as it allows general, pre-trained models to be adjusted by communities for their own needs" (Peszkowska and Tarkowski, 2024). As LLM fine-tuning can be less intuitive than LoRA training, we chose the LoRA approach for the workshop.[26]

The workshop served multiple critical functions within this research. For one, it encouraged local actors to take an active approach to AI, something which Kulesz phrases in his recommendations as a "[s]hift from 'it' (AI) to 'we'—adopting a truly participatory, collaborative and co-construction approach" (Kulesz, 2024, p. 87). The workshop participants were presented with a low-threshold but high-impact method to use and shape AI. For another, we treated the workshop as a sort of ethnographic field site. While a true ethnographic project was not feasible in the given time frame of this research project, the method of participatory observation and subsequent qualitative interviews allowed us to obtain a better understanding of how local actors might view and

[26] Since LLMs are based on question-answer text pairs, it can be less intuitive for users to come up with suitable questions and answers to use as training data. On the other hand, image generation models are trained on image-text pairs where the text simply describes the image content. Since many people already take photos in their day-to-day lives, it can be more intuitive for participants to gather images and describe them to train an AI model.

experience their interaction with AI and the critical data injection process. From this, we derived insights into the potentials and limitations of the participatory approach to critical data injections and an AI of the commons.

In order to make this workshop replicable in different local contexts, this report comes with a toolkit which allows anyone to adapt the format for their own work. This can be for the individual use of AI and the creation of 'critical data injections' or for educational purposes in local communities or institutions. The following section provides a summary of the toolkit which was shared with the workshop participants for educational purposes. **The link to the full-length toolkit can be found at the end of this report (see Appendix 2).**

3.2 LoRA Training Toolkit: a Summary

The LoRA Training Toolkit aims to make the rather technically sophisticated process of training one's own AI model for image generation accessible to non-expert users. By following the steps described in the toolkit, anyone with the necessary hardware should be able to create an AI image-generation model which renders images based on their training data. Due to rapid developments in the field of AI, aspects of the toolkit may become outdated over time. Therefore, the instructions aim to convey the general ideas of the process, allowing readers to adapt to any changes in the future.

The focus of this toolkit lies in creating Low-rank Adaptation models for the Stable Diffusion XL (SDXL) base model developed by Stability AI. While newer and more prompt-adherent models have been released since SDXL (e.g. FLUX by Black Forest Labs or Stable Diffusion 3 by Stability AI), it has proven to be broadly used and supported over time. For devices with less capable hardware, the toolkit can be adapted for Stable Diffusion 1.5 (SD15). The LoRA Training Tool which is introduced in the toolkit has presets for all the above-mentioned models so that the process can be adapted to local requirements.

The toolkit was tested on Windows devices. All important steps should be the same on macOS and Linux devices, although LoRA training and image

generation on macOS is only reasonably fast and supported on Apple Silicon chips (e.g. M1, M2, M3, M4). As of this writing, OneTrainer, which is used for training the LoRA model, is not supported on macOS devices. However, there are other macOS applications, such as Draw Things, which support LoRA model training. While the toolkit cannot cover all related tools and applications, the concepts from this toolkit can be applied across a range of different software.

Box 1. Structure of the LoRA Training Toolkit

1. Introduction: This answers the questions of what Stable Diffusion is, what Low-rank Adaptation models are, and how diffusion-based image generation works. It further outlines the necessary hardware, what general steps will ensue, warnings about using the software introduced in the toolkit, and general ethical concerns when using generative AI.

2. Testing the base model: The section introduces the Stable Diffusion XL base model. It gives instructions on how to use Stability Matrix to install ComfyUI, download the Stable Diffusion XL base model and generate the first images inside ComfyUI. This allows the user to reflect on the biases of the base model and gives room to consider what the user wants to do with their LoRA. For example, do they want to change a specific style, image content or composition?

3. Preparing the data set: Based on the understanding from the previous section, the user has set their mind on what their LoRA is going to be about. This section gives further ideas on how to collect and create data based on the user's goals with LoRA training. It suggests a dataset of at least 25 images, but the dataset can also contain hundreds or thousands of images. The user should try to find or create images which roughly resemble what they want to later generate with AI. Some general considerations are:

- when reproducing specific objects, it makes sense to capture them from many different angles to allow for more diverse outputs later;

– when reproducing styles, it makes sense to collect diverse images, all of which are, for example, taken with the same analogue film stock, camera or similar compositions.

Both can be mixed and the process is usually one of trial and error, finding what works best for which particular situation.

4. Image tagging: Once the dataset is gathered, all of the images have to be 'tagged'. The tagging process is one of the most important parts of LoRA training as it connects the image content with semantic descriptions which will later be triggered during the generation process. The toolkit offers some helpful considerations for how to approach tagging.

5. LoRA training with OneTrainer: OneTrainer is an open-source AI model training tool which can run inside Stability Matrix. If the descriptions in section two were followed, OneTrainer can easily be installed alongside ComfyUI. Inside OneTrainer, presets for different base models, such as SDXL and SD15, can be selected. The toolkit provides examples and presets which can be easily adapted for training one's own LoRA model, using OneTrainer.

6. Using the LoRA model to generate images: This section expands the 'Training an AI Model' section by explaining how to add a LoRA model in the generation process and how to do A/B testing[27] with and without LoRA. Respective ComfyUI workflows are provided together with the toolkit. This allows for reflections on the impact of the LoRA model, the dataset and the implications of tagging the images. From here, iterations are possible to make changes to the LoRA and repeat the steps described in sections three to six.

[27] In this context, A/B testing refers to generating two images with the exact same settings, where the only difference between the two images is the use of the LoRA model. By doing so, we can isolate the impact the LoRA model has on the image generation.

3.3 Results of the 'Critical Data Injection' Experiment: Experiencing a Participatory AI of the Commons

As previously mentioned, the workshop which took place as a part of this research project aimed to create space for open-ended conversations on feminist and decolonial approaches to AI between academia, artists and IT professionals. We brought together academics based in East Asia who were working in different fields, including but not limited to International Relations, international migration, political studies, political economy, AI technology and ethics, art history, and comparative literature, as well as Korean artists using different media including AI in their (art)work and Korean IT professionals working on AI development.

This diverse constitution helped the (small-scale) mapping of different levels/ways of engagement with and understanding of AI and enabled the exploration of questions about AI at both the macro and the micro level. Furthermore, the possibilities of feminist and decolonial interventions in the understanding and use of AI were explored in order to seek a form of participatory and intersectional AI. These conversations were followed by a two-day workshop, where the participants could actively engage with AI by creating their own datasets and turning them into models for image generation, a hands-on process we refer to as 'critical data injection'. This activity was followed by reflections and conversations on the possibilities of AI intervention from different positionalities as well as the possibility of an AI of the commons.

Collectively Making Sense of AI

Before the critical data injection process, discussions between participants and different experts took place, specifically on the modernity/coloniality of AI, the possibility of AI being implicitly prejudiced and the blind spots of AI safety. These led to discussions around issues on ethics, materiality, global inequalities, and power which AI could raise.

A major critique, as raised by the political scientist Shang E. Ha (Sogang University) and Critical International Relations scholar Young Chul Cho (Kangwon National University) was the possibility of AI systems being inherently biased, reflecting the social, political and economic conditions in which they are built. In response to this critique, AI technicians within and beyond the workshop questioned the specificities of detecting such biases. Tech companies and policymakers both look for evaluation frameworks to identify what would count as 'bias', to determine where in the dataset or outcome the bias shows and how to remove this bias, to ultimately 'solve' the problem by reducing, mitigating or removing the bias. However, it is naïve to assume that the definition of 'bias' would be the same everywhere, as meanings of what is 'sexist', 'racist' or what counts as 'discrimination' could shift across cultures and contexts. In this globalised world where datasets from different contexts can constitute AI-generated data, how are 'discrimination' and 'bias' defined? According to one of the workshop speakers, political scientist Jungmin Seo (Yonsei University), another question might be whether a universal AI would even be possible, or whether we are facing the emergence of (a) pluriversal AI(s), with different value systems and epistemologies.

At the same time, Young Chul Cho pointed to the problems at the macro level: it is impossible to ignore the deep entanglement of AI technology with capitalism, which both shapes and is shaped by the AI industry. Often forgotten, the material underpinnings of this industry should also be part of discussions around AI. These include the hidden ecological and human costs generated as a consequence of using AI which make this technology possible (i.e. electricity, water and the rare earth materials crucial to AI hardware, as well as exploited labour in the Global South). This raises more questions, such as: for whose benefit is AI being built? Who defines what is or is not 'legitimate' or 'acceptable' data?

The question of 'who' continues to be an issue in AI safety. When governments and enterprises promote and promise AI safety, whose safety matters and on whose terms is it defined? When corporations and policymakers seek an 'ethical AI', in which ways are they seriously tackling the issues of misinformation, privacy violations and dehumanisation? A member of the workshop audience,

Yeonhee Sophie Kim (Sogang University), who works on the politics of ethics and aesthetics, pushed the conversation further by considering AI as a posthuman actor. What would then be the ethical concerns when we consider the relationship between humans and AI? How should we treat AI?

These tensions and questions suggest that AI is not simply about technological development but about humanity and what is humane. In training AI, global conversations about justice, ethics and what kind of world humans imagine or want to have are required.

Creating Other Patterns through Data Injection: a Case Study in Reimagining Seoul

While data injection itself will not change the economic system in which AI technology operates, the workshop envisaged the creation of different patterns of data in order to influence the average data patterns (which are at the moment biased) and to become more aware of the implications of using AI as users and consumers of AI technology. This workshop is one of many attempts to make disruptions in and through everyday practice, by changing individual perceptions about what data could be, what one can do with data by experiencing different possibilities of engaging with AI, and through open-ended conversations on and engagement with the meanings and ethics of and around AI.

The first 'experiment' conducted in the workshop was working on the critical data injection of a 'photo of Seoul' with participants. The aim of choosing such a topic was to reimagine the urban space in which intercultural exchanges were happening, and reenvision it against the typical images of Seoul which AI generates: a global tourist city, a luminous city with palaces and modern buildings, mostly devoid of humans. AI-generated images typically show touristic sites such as Namsan Tower, Kyungbok Palace and skyscrapers. These depictions often align with the images which the state of South Korea curates of itself. According to the South Korean government, the brand image of Seoul is that of a city which is developed, modern, global, clean, and organised.

What kind of imaginary does such an image of Seoul generate, within and outside of South Korea? While these images reinforce the brand image which South Korea is trying to create for itself, they do not reveal what has been systematically erased or concealed from Seoul's representation to serve narratives of development, modernisation, urban cleanliness, and the city's branding as a prosperous global centre. It does not show people's lived experiences of the city nor does it expose the urban planning logic and policies which deliberately enhance Seoul's image as a site of development, modernity and globality.

Research has shown that Seoul's urban space is gendered, classed, racialised, and ableist (e.g. cf. Kim, 2009; Kim, 2015). For instance, even though parts of Seoul reveal the economic disparities within the city, these inequalities are hidden from the images of Seoul which are reproduced by AI. Other research shows that the public space of Seoul is a largely heteronormative space where the queer population is marginalised or made invisible (e.g. cf. Kim, 2015; Lee, 2015; Han, 2018). Efforts to 'queer' Seoul's urban space continue with Seoul Pride, a festival and parade which takes place every June in the city centre. The protests which ensued following Korea's now-impeached president's call for martial law also created a new landscape of the public space in Seoul, occupied with young women, LGBTQ+ and disabled people who emphasise that they had always been part of "the people" and the people's resistance.

Given this context, what does it mean to reimagine Seoul? For one, it is about reflecting on the location where we gathered for intercultural exchanges around the topic of AI. Each participant, either as a visitor or permanent or temporary resident of Seoul, had a different relationship with the city, depending on, for example, which part of the city they lived in, which neighbourhood(s) is/are part of their everyday life and what they had experienced in Seoul. This means that Seoul is not a unitary space but a lived space which can be and needs to be imagined from different perspectives. For another, it is about reflecting on the city from a personal perspective: what would an accurate representation of Seoul look like, if seen from the everyday lived experience of those who live in the city? What kind of images could be generated if we were to focus on *who* constitutes the city, rather than *what* constitutes it? What would be the political implication of reflecting on 'other' possible images of Seoul? Here, the personal is not simply

a private matter, but reveals how much the lives of individuals are imbued by power relations. The personal becomes the political.

By engaging with everyday images of Seoul, we intend to reflect on how a more people-centred approach to imagining a city can be practically applied, which considers the multiple intersectional layers which constitute the city's population and the diversity of lived experiences which may otherwise become lost in the process of AI data generation.

Image Tag and Prompt: "Photo of Seoul": a Feminist and Decolonial Data Set

We asked eight workshop participants to take a minimum of 25 and up to 30 pictures of Seoul or to look into their camera or phone's photo albums and choose pictures they would label as a "Photo of Seoul" for critical data injection purposes. We also conducted interviews or asked participants to write a note to accompany the images they shared, describing what the image of Seoul was for them and/or what Seoul meant for them. Together, these notes and photos make up a little archive of Seoul which is aimed at generating *other* images of Seoul through critical data injection. Parts of the photos and texts were shared among participants to demonstrate how different meanings could be engendered through this collective process of critical data injection. As the following excerpts of reflection from participants demonstrate, the version of Seoul represented in these images and texts are seen subjectively from diverse positions through personal everyday life experiences and memories:

> *[Seoul is a] chaotic, noisy city with a jumbled mixture of various elements. However, just beyond the concrete jungle and bustling streets appear forests, streams and occasional wildlife. Old trees and quiet alleys maintain their places, and the view from the winding hills looks desolate and pitiful, yet somehow familiar and comforting. Contradictory political banners hanging throughout downtown speak of the turbulent political climate, while old houses and shops stand among sleekly polished high-rise apartments and modern glass buildings, creating a strange coexistence. Small neighbourhood*

bookstores or cafés barely offer places to rest, but on the other hand, this is a city with few public spaces which encourages consumption wherever you go, with most spaces inaccessible without spending money. In the alley where tragedy[28] occurred, a very small number of people continue to remember and mourn, and even in a society with extremely low birth rates, children are still growing up. It is a city where leaves sprout between endless deaths and competition, and someone's humour still exists in the form of a snowman.

— Hyunji Choi (writer and translator)

***

Looking at these photos, I feel like living in the outskirts of Seoul gives me a different perspective than what most people think of when they imagine Seoul! It feels closer to nature and animals...? When I walk around, I mostly see trees and mountains, and I even spot water deer and wild boars quite frequently!

The places I visit tend to be residential areas or neighbourhood centres without many people, so my photos capture more of everyday life rather than the bustling crowds.

I also think Seoul is a place where people are passionate and sincere about protests, anniversaries and holidays. When it snows, everyone gathers to build snowmen; in spring, as if coordinated, everyone flocks to cherry blossom spots; and it's the same for Christmas and other holidays! It's a place which is passionate about seasonal fruits and foods... Just like my family making 120 kg of kimchi while living in Seoul.

Among the photos I sent, there are some places which anyone would recognise immediately, but there are also places with birds, cats or no people at all. Those are from my neighbourhood! People are often surprised that just a little distance from central Seoul there are such green areas. There are azaleas, forsythias, black locusts, and so on—it feels close to nature...

[28] The tragedy refers to a crowd crush in Itaewon, one of the most popular nightlife districts of Seoul, where more than 150 people died while celebrating Halloween in October 2022.

And even famous tourist spots like Gyeongbokgung Palace or Namsan Tower become quiet if you step just a little away from them and your perspective of the same buildings or landmarks changes.

— Minyoung Jeong (graduate student)

***

These photos I've shared are ones I took while walking around or meeting friends, so they capture everyday scenes of Seoul.
Each photo has its own story behind it... (meeting places, like the clock tower at Seoul Arts Center, the strikingly large banner with a face at the Kim Dae-jung Library, photos from Uji Coffee which I wanted to post but couldn't before, sunset photos taken from my room, a hanbok shop from Jeonju which replaced a café and moved to Hongdae in Seoul, etc.) I guess you could say I've preserved Seoul's image through these little stories.
The photos of large roads and night views feel somewhat foreign to the Seoul I live in, so I took them more like a tourist would. The busking photo is from Hongdae and the foreign performer was later asked to leave. I think they didn't have a permit, which made me realise that even locals don't fully understand these systems, so it must be even harder for foreigners. Anyway, I hope these various everyday backgrounds are helpful.
When it comes to Seoul as a tourist destination, it seems to be about the difference in perspective. Looking down at Seoul from a bridge or viewing it from a high place feels different and puts me in a tourist's mindset. 'Oh, so this is how Seoul looks. I didn't know it had this side to it.' On the other hand, from my everyday position, commuting to work via public transportation, walking around, or lying in my low-ceilinged room, I see Seoul at my eye level, the Seoul where people work.

— Dongeun Shin (researcher)

***

To me, the image of Seoul seems like this: it monopolises the fruits of growth while wanting to drive away cockroaches (as symbols of pollution or the byproducts of pollution) and then outsources this 'driving away' work to marginalised beings once again... The first photo is of Seoul National University Medical School, which I photographed with a similar feeling.

— Seohwa Park (journalist)

***

This makes me think about whether Seoul is a friendly city. Is it friendly to foreigners, to non-humans, especially to cats, and to people with disabilities? What does it mean for a city to be friendly? [...] Looking at these photos reminds me of Chungking Express [29] and it made me wonder: if Seoul had an expiration date, how many years would it be? I've been thinking about that.

— Hye Min Ryu (graduate student)

***

While collecting these photos, I was reminded once again that queer people have always been present on the streets. Most of the photos I sent you are from rallies and the Seoul Queer Culture Festival[30]. I remember the 2019 Queer Culture Festival particularly, when there was a march route which made a U-turn in front of Gwanghwamun, which I thought was quite symbolic. To be honest, I feel somewhat frustrated that the voices of queer people and young women in the square are being highlighted as if they're new, when we've always

[29] An anthology film directed by the Hong Kong film director Wong Kar-wai (1994). In the film, the expiration dates of pineapple cans symbolise a relationship which is approaching an end.
[30] The Seoul Queer Culture Festival is a yearly festival to celebrate LGBTQ+ communities. Seoul Pride (or the Queer Parade) is part of the festival. Until 2019, the Queer Parade marched from Seoul City Hall through the streets of the city centre to Gwanghwamun, the main gate of Kyungbok Palace, the official palace of the Chosun Dynasty.

been here...?[31] *I think I selected these photos with the thought that my rainbow comrades have been fighting all throughout Seoul.*

— Seongjo Jeong (activist, researcher)

***

The cityscape does not feel so consistent when compared to Europe. It feels comfortable, uncomfortable, familiar, unfamiliar at the same [time]. It's a grey-ish and blue-ish city [sic]. It's simultaneously blurry and vivid [sic]. The city makes me stay up all night.

— Ihntaek Hwang (researcher)

***

I was thinking of a foreigner's perspective. Seoul is for me betweenness. [...] You would find crowd and quiet corner for you book reading at the same time [sic]. [...] Like you would see luxury and electric wires at the same street. Kinda feel like me. You know coming from Turkey between east and west. Electric wires give me vibe like an old town but also, I see great building very next to them see something modern [sic]. [...] More like contradictions is the middle name of Seoul for me. [...] Maybe irony. Beauty and beast is together in this city [sic]. So, I hate but I love. I'm sure many foreigners would say the same thing. Also, I think Seoul is the city where you can find all the colours. I don't know how they did that but just the colours of this city are mesmerising. [...] Seoul makes you want to try its history. Wants you to feel it.

— Nur Cenan Baser (graduate student)

[31] This comment refers to the protests since 3 December 2024, when the now-impeached president of South Korea, Yoon Suk Yeol, declared martial law. Since then, young women and LGBTQ+ people have been protesting actively on the streets and received much media attention.

The texts which workshop participants provided along with the pictures of Seoul which they shared commonly describe the ambiguities and complexities of Seoul as a city. The city of Seoul entails contradictions and in-betweenness, standing on both the presence and absence of specific people and species, giving both an inclusive and exclusionary impression to the people who live in it. It is also a very personal space which brings family and friends to mind, where social movements can be imagined and where deeply personal emotions arise. These understandings of Seoul challenge uniform interpretations of Seoul as the touristic, globalised and hyper-modernised capital city of South Korea. They point to multiple interpretations of Seoul as a space of everyday life and a space of political possibility, to make certain people visible, to mourn, to remember, where some lives are included while others are excluded or even exterminated.

Through LoRA training with the images provided by workshop participants, new imagery was generated which significantly differs from typical AI outputs. Where conventional AI-generated images of Seoul emphasise tourist landmarks and gleaming infrastructure, our community-trained model produced images featuring elements which participants valued: neighbourhood corners, protest banners, rainbow flags, children at play, urban wildlife, and other markers of lived experience which rarely appear in state representations of Seoul. The images also make visible different groups of people: children, LGBTQ+ people (symbolised by rainbow flags) and more, as well as non-human elements which do not appear in conventional AI-generated images of Seoul.

The following images show the difference the LoRA model made when generating images for the prompt "Photo of Seoul". Notably, all of the settings are the same between the image pairs, the only difference being the LoRA model used on the second image.[32] Some of the images are generated with the official Stable Diffusion XL base model[33] and some are generated with the online-community fine-tuned Stable Diffusion model "ZavyChromaXL V6.0"[34].

[32] The notable settings which are the same between images are the base model (i.e. not the LoRA model), seed, prompts, CFG scale, and steps. All of these values are explained in more depth in the LoRA Training Toolkit, see https://miroleon.github.io/lora-training-toolkit/.

[33] https://huggingface.co/stabilityai/stable-diffusion-xl-base-1.0/blob/main/sd_xl_base_1.0_0.9vae.safetensors (accessed on 14 August 2025).

[34] https://civitai.com/models/119229?modelVersionId=416867 (accessed on 14 August 2025).

Image 1: "Photo of Seoul"

AI image generated from the LoRA Training Workshop, December 2024. This image is generated with the community fine-tuned base model "ZavyChromaXL V6.0" ***without*** *the LoRA model.*

Image 2: "Photo of Seoul"

AI image generated from the LoRA Training Workshop, December 2024. This image is generated with the community fine-tuned base model "ZavyChromaXL V6.0 ***with the LoRA model and the same settings as in Image 1****.*

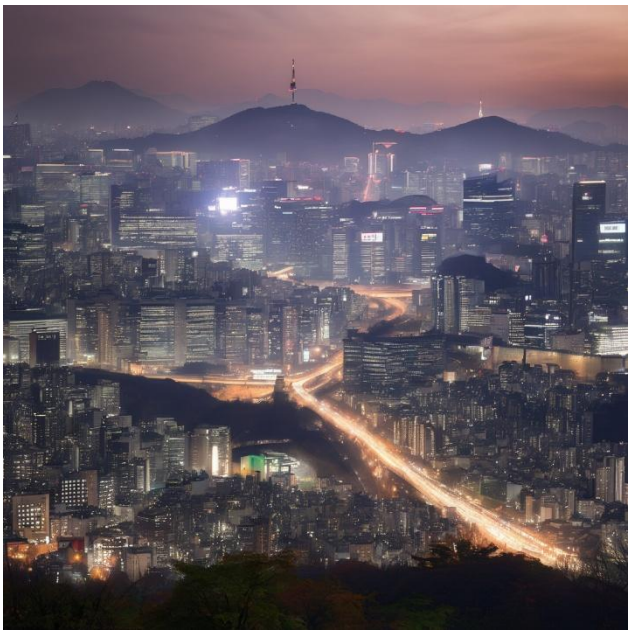


Image 3: "Photo of Seoul"

AI image generated from the LoRA Training Workshop, December 2024. This image is generated with the official Stable Diffusion XL base model ***without*** *the LoRA model.*

Image 4: "Photo of Seoul"

AI image generated from the LoRA Training Workshop, December 2024. This image is generated with the official Stable Diffusion XL base model together ***with the LoRA model and the same settings as in Image 3.***

Image 5: "Photo of Seoul"

AI image generated from the LoRA Training Workshop, December 2024. This image is generated with the official Stable Diffusion XL base model ***without*** *the LoRA model.*

Image 6: "Photo of Seoul"

AI image generated from the LoRA Training Workshop, December 2024. This image is generated with the official Stable Diffusion XL base model together ***with the LoRA model and the same settings as in Image 5****.*

3.4 Analysis

The workshop activity helped explore how to influence the kind of images which can emerge from our use of AI and how these images might be shaped through reflections on daily life. It helped reimagine worldviews; in this case, providing alternative ways of viewing Seoul. While this approach has significant limitations, given our dependence on base models with existing biases embedded in their training data, this project demonstrates possibilities for community agency in AI representation.

The work could be expanded to rethink how spaces are represented, through processes of visualising, hyper-visualising and invisibilising certain groups of people and things. It demonstrates how personal experiences can reshape our understanding of what spaces can be and do, even within the constraints of existing AI infrastructure. The key insight from this experiment is not that small-scale interventions can transform AI wholesale, but that they can create spaces for alternative ways of seeing and being seen: spaces which might otherwise remain invisible in increasingly AI-mediated visual cultures.

As previously mentioned, we follow an idea of 'the commons' based on anthropologist Massimiliano Mollona's interpretation. In reference to De Angelis (2017), Mollona defines his idea of the commons as a conjunction of three conditions. Therein, the commons is "(1) a pool of natural and/or human resources, (2) a community of people with reciprocal and sharing relations and (3) acts of working together towards the reproduction of the community" (Mollona, 2021, p. 9ff).

If we apply this to the concept of an 'AI of the commons', we can see how these conditions may apply to the communal creation of AI models. Keeping in mind that AI models are trained with an accumulation of different data points, whether these are question-pairs for language models, image-text pairs for image generation models or any other type of data, the first condition can be met when these data points are put together in a mutual context or 'pool'. This condition

is practically always met when AI models are trained, even if not in the context of a commons.

The second condition adds a social element to the commons, where the human actors are standing in reciprocal relations and sharing them with one another. This is where commercial AI models deviate from an AI of the commons. Commercial AI models commonly take data points without explicit consent from the creator of that data point, train their AI model and then offer a paid service which the potential creators of the data points can use. If we assume that the initial data acquisition is non-consensual and if we further assume that a reciprocal or sharing relation always has to include some form of consensus, most (if not all) commercial AI models fail this definition. This may also extend to open-source models which have not been trained on appropriately licensed training data. If we were to have an AI of the commons, it would require all of the participating members to voluntarily share their data in a mutual interest for the AI model to be trained with such data. This is what we can see in the context of our workshop, where participants offered their photos of Seoul in a mutual interest to change the representation of such. This can be adapted to other local communities and theoretically scaled up, as long as there is a general consensus process wherein data points are provided voluntarily in a mutual interest to train the AI model based on such.

The third condition expands on this social element in the sense that this act is not singular but an ongoing process of communal engagement. It aims not just at one result but is rather a system of actions which are continuous and repeating. This also applies to the (re-)generative qualities of AI. As new data points are provided within the commons and as the data set evolves, new AI models can and should be trained. Just as values and norms in communities change, their representative models should adapt and change as well.

3.4.1 Expanding from Local Models to AI systems

When we combine the idea of an 'AI of the commons' with our practical experiments of local LoRA model training, we can derive a larger system from the two which can go beyond what commercial AI models currently offer.

To recall, the Low-rank Adaptation model allows for compatibility with a base model and multiple other LoRA models. One of the disadvantages of these smaller models is their limited data. Although our trained LoRA for the 'photo of Seoul' is useful for reproducing and representing this pre-defined context, it will not be very good at doing so for other contexts. However, if we create a technical system which can combine multiple small models without the need for a base model and if this is accompanied by a social system which supports the creation of data sets and AI models in the context of local commons, this can lead to a more powerful AI tool which can combine a multitude of representations in the future. This will have to be the task for AI development, particularly in the context of publicly funded research institutions.

Another critical dimension in evaluating the representational nature of current artificial intelligence systems is their operational paradigm. Most contemporary AI models, especially large language models (LLMs), are fundamentally probabilistic systems which rely on statistical pattern recognition. These models do not 'understand' language in the human sense but rather generate outputs by sampling from distributions of token probabilities shaped by massive datasets. This design means that phrases, facts and associations which occur more frequently in training data are more likely to reappear in model responses (Coronado-Blázquez, 2025). While this mechanism supports the accurate reproduction of widely accepted knowledge, such as scientific constants or measurement conversions, it also risks replicating and amplifying cultural biases or oversimplified generalisations embedded in the data. Such limitations pose significant concerns in contexts like policymaking, education and social discourse, where the reproduction of dominant narratives may reinforce existing power imbalances or marginalise minority perspectives.

To address these issues, recent developments in AI architecture have explored modular frameworks designed to improve both model efficiency and content sensitivity. A promising direction involves Mixture of Experts (MoE) AI models, which break away from monolithic architectures by routing inputs to specialised subnetworks (or 'experts') tailored for different types of information.[35] One recent paper proposes the DeepSeekMoE architecture (Dai et al., 2024), which introduces two key strategies to enhance expert specialisation: fine-grained expert segmentation and shared expert isolation. The former allows knowledge to be distributed more precisely by decomposing tasks into smaller, more focused components, while the latter centralises general or common-sense knowledge to reduce duplication across experts. This modular design not only improves computational efficiency but also opens possibilities for more targeted and interpretable outputs, potentially mitigating the uncritical repetition of statistical biases and enabling more nuanced knowledge representation.

This concept is not limited to image generation models; it can also extend to language models and other common AI technologies, such as object detection. To our knowledge, no such system currently exists and private companies are unlikely to pursue the development of an open infrastructure due to the lack of commercial viability. Given these circumstances, it appears that creating such an open AI infrastructure is most feasible through public funding and research efforts.

3.4.2 AI and its Limits of Representation

The discussion, to this point, may suggest that the proposed AI of the commons is an ideal state of representation and ethical AI practice. However, this is not the case in at least one central understanding of 'the commons'. Just because a local group of actors comes together and agrees on providing data to a common AI

[35] The general idea behind the Mixture of Experts approach for AI models is that instead of a prompt triggering the entire model to generate an answer, only the relevant parts of the AI model are triggered. These relevant parts are what are referred to as 'experts'. This can improve the efficiency of the model and potentially make outputs more specific to the prompt.

model, this does not make the data 'good' or 'right'. Theoretically, a local group or 'commons' can gather racist, misogynistic or otherwise ethically problematic data and train an AI model which reproduces such views. Thus, one may say that the proposed system is even worse than a commercial AI model which may use some form of guardrails when gathering data.

However, the AI of the commons maintains a few central advantages compared to the privatised model. For one, assuming that this group acting unethically is part of a country which enforces its laws, any unlawful conduct can be prosecuted. In the current state of AI development, this is much more complicated with privately funded tech companies, as their data sets are often not publicly accessible (see Widder et al., 2024). For another, once a system exists in which any number of local models can be combined, there is the option to avoid using such unethical models, whereas one could not avoid them in a singular commercial AI model.

To ensure that this infrastructure is maintained, we suggest that there should be councils of representatives from a broad and inclusive range of social groups overseeing local models and allocating resources for creating data sets, models and AI applications. This could be structured similarly to, for example, the German public broadcasting agencies, which are responsible for broadcasting local and regional as well as national events, independent of the government and overseen by councils of representatives. Similarly, these AI councils could ensure representation and local allocation of the necessary means to keep the infrastructure intact.

4. Conclusion and Policy Recommendations

4.1 What is there to learn from this experiment?

Values and norms are constantly changing. Even if it were possible to create a singular AI model which seems perfectly balanced today, that model could quickly become outdated. Ideas which feel reasonable or neutral today may later be recognised as biased or exclusionary. Turning toward an AI of the commons does not eliminate this risk; rather, it embraces it as an inherent part of democratic knowledge production. By centring an ongoing, participatory process of adaptation, the 'commons' approach enables a more responsive and accountable model of AI development, one which reflects the evolving needs, values and identities of its contributors.

The modular nature of Mixture-of-Experts architectures offers a compelling technical blueprint for such a system. Just as MoE models route different types of inputs to specialised expert modules, a decentralised, publicly funded AI infrastructure could allow communities to create their own expert models, trained on consensually provided, locally relevant datasets. These models could then be recombined in flexible ways to serve diverse tasks, such as image generation, language modelling or civic applications (e.g. to make policies, regulations, and administrative processes more accessible to citizens), without being beholden to centralised platforms or commercial gatekeepers. Crucially, this does not require a flawless or morally superior dataset. It simply requires transparency, consent and the ability to challenge or avoid problematic sub-models, possibilities largely foreclosed by proprietary systems trained on opaque, inaccessible datasets.

In this vision, the 'experts' are not just technical modules; they are communities. By investing in an infrastructure which enables communities to build, share and govern their own representational models, we shift AI development from an extractive industry into a participatory and deliberative process. Rather than being provided with outputs from deterministic systems trained behind closed doors, users can become co-creators, shaping the outputs and ethics of the tools they use.

The following section presents policy recommendations for governments, multilateral institutions, various actors in the cultural sector—especially those operating in international cultural relations (ICR)—to make AI development a more inclusive and participative process.

4.2 Policy Recommendations

Technical Infrastructure Development

Invest substantial public funding into the research and development of open AI frameworks. National governments (especially science and technology ministries), EU institutions and programmes (e.g. Horizon Europe), UN Agencies (e.g. UNESCO), and international funding bodies can support the development of technical infrastructure by providing substantial public funding for research and development of one or multiple open frameworks which allow modular AI systems, comparable to the Mixture of Experts, to integrate small, local models without requiring centralised base models.

What can actors in international cultural relations do?

- Cultural institutions can advocate for these investments through existing policy networks,
- demonstrate use cases through pilot projects with artists and specialists from the tech sector, and
- provide evidence of community demand for alternative AI systems.

Community Capacity Building

Provide **funding and infrastructure support** for local communities, cultural institutions and other civil society organisations to **collect, curate and annotate their own datasets in transparent and democratic ways**. Cultural

institutions, arts councils, civil society organisations, and local communities could do this through the following suggested measures.

Examples of measures actors and stakeholders in ICR can potentially take to support capacity building

- Bilateral cultural cooperation agreements which include provisions for collaborative community AI projects.
- Cultural and/or academic exchange programmes (e.g. British Council, Fulbright, DAAD) can fund artist/researcher residencies focused on community AI development.
- Museum networks and international arts festivals could pilot collaborative dataset creation across cultures.
- Cultural foundations could establish grant programmes, specifically for "AI as commons" projects.
- Practical mechanisms: facilitate mobile digital storytelling labs, equip community media centres with AI training capacity and facilitate travelling workshops using toolkits like the LoRA Training Toolkit developed in this research.

Democratic AI Literacy

Create and/or provide **educational programmes and accessible tools which enable non-technical participants to understand, train and audit AI models created within their communities**. Educational institutions, cultural organisations and other civil society actors (including artists and others working on digital rights) could cooperate and facilitate the creation and provision of these programmes and tools. Furthermore, various actors in international cultural relations can implement certain programmes and initiatives to bring this effort further (see next page).

What can actors in international cultural relations do?

- Cultural institutes (e.g. Goethe-Institut, Institut français, etc.) could integrate community AI training exchanges into their programmes.
- International university partnerships could develop exchange programmes focused on participatory AI development.
- As current tensions and questions on AI are not simply about technological development but about humanity and what is humane, higher education institutions and think tanks should host public discussions between policymakers, scholars, artists, and actors from the tech sector on justice, ethics and what kind of world humans imagine or want to have.
- Arts education networks could incorporate community AI training into their curricula.
- Digital rights organisations could provide technical support and advocacy training for non-experts/lay users.

Accountability and Governance

Establish **legal and technical mechanisms** which allow communities to **flag or reject unethical or harmful sub-models within larger AI systems, promoting accountability without censorship**. Policymakers in national governments, multilateral organisations and civil society watchdogs could work together to establish and enforce these mechanisms.

What can actors in international cultural relations do?

- Cultural institutions can serve as trusted intermediaries between civil society, government officials and the tech sector,
- host community AI auditing workshops, and
- provide safe spaces for discussing AI ethics across cultural contexts.

References

Adam, A. (1995). "Artificial Intelligence and Women's Knowledge", *Women's Studies International Forum*, 18(4), pp. 407–15. https://doi.org/10.1016/0277-5395(95)80032-K.

AI/AI (n.d.). *AI/AI: Artificial Intelligence, Art and Indigeneity*. Available at: https://aei.art.br/aiai/en/ [Accessed on August 27, 2024].

Ali, S. M. (2016). "A Brief Introduction to Decolonial Computing", *XRDS: Crossroads, The ACM Magazine for Students,* 22(4), pp. 16–21. https://doi.org/10.1145/2930886.

Baack, S. (2024). *Training Data for the Price of a Sandwich: Common Crawl's Impact on Generative AI. Mozilla Insights*. Available at: https://foundation.mozilla.org/en/research/library/generative-ai-training-data/common-crawl/ [Accessed on August 26, 2024].

Benjamin, R. (2019). Race After Technology: Abolitionist Tools for the New Jim Code. Cambridge: Polity Press.

Birhane, A. (2020). "Algorithmic Colonization of Africa", *SCRIPT-Ed*, 17 (2), pp. 389–409. https://doi.org/10.2966/scrip.170220.389.

Bremmer, I., and Suleyman, M. (2023) "The AI Power Paradox", *Foreign Affairs*, 16 August. Available at: https://www.foreignaffairs.com/world/artificial-intelligence-power-paradox [Accessed on November 11, 2024].

Browne, J., Cave, S., Drage, E., and McInerney, K. (eds). (2023). *Feminist AI: Critical Perspectives on Data, Algorithms and Intelligent Machines.* Oxford: Oxford University Press.

Chaka, C. (2022). "Digital Marginalization, Data Marginalization, and Algorithmic Exclusions: A Critical Southern Decolonial Approach to Datafication, Algorithms, and Digital Citizenship from the Souths", *Journal of E-Learning and Knowledge Society,* 18(3), pp. 83-95. https://doi.org/10.20368/1971-8829/1135678.

Choi, M., and Hong, J. (2021). "The Bare Face of Misogyny and Discrimination Revealed by Chatbot 'Iruda' (챗봇 '이루다'가 보여준 여성 혐오와 차별의 민낯)". *Kyunghyang*, January 11. Available at: https://www.khan.co.kr/article/202101111031011 [Accessed on November 11, 2024].

Chowdhury, R., and Lakshmi, D. (2023). "Your Opinion Doesn't Matter, Anyway". Exposing Technology-Facilitated Gender-Based Violence in an Era of Generative AI. World Trends in Freedom of Expression and Media Development Series. Available at: https://unesdoc.unesco.org/ark:/48223/pf0000387483 [Accessed on November 11, 2024].

Coronado-Blázquez, J. (2025). "Deterministic or probabilistic? The psychology of LLMs as random number generators" (Version 1), *arXiv preprint*. https://doi.org/10.48550/ARXIV.2502.19965

Costanza-Chock, S. (2018). "Design Justice, A.I., and Escape from the Matrix of Domination." *Journal of Design and Science*, https://doi.org/10.21428/96c8d426.

Couldry, N., and Mejias, U. A. (2019). The Costs of Connection: How Data Is Colonizing Human Life and Appropriating It for Capitalism. Stanford, California: Stanford University Press.

Couldry, N. and Mejias, U. A. (2021). "The Decolonial Turn in Data and Technology Research: What Is at Stake and Where Is It Heading?", *Information, Communication & Society,* 26(4), pp. 786-802. https://doi.org/10.1080/1369118X.2021.1986102.

De Angelis, M. (2017). *Omnia Sunt Communia: On the Commons and the Transformation to Postcapitalism*. London: Zed Books. https://doi.org/10.5040/9781350221611.

Dai, D., Deng, C., Zhao, C., Xu, R. X., Gao, H., Chen, D., Li, J., Zeng, W., Yu, X., Wu, Y., Xie, Z., Li, Y. K., Huang, P., Luo, F., Ruan, C., Sui, Z., and Liang, W. (2024). *DeepSeekMoE: Towards Ultimate Expert Specialization in Mixture-of-Experts Language Models* (Version 1). *arXiv*. https://doi.org/10.48550/ARXIV.2401.06066.

Drage, E. and K. Mackereth. (2022). "Does AI Debias Recruitment? Race, Gender, and AI's 'Eradication of Difference'", *Philosophy & Technology,* 35(89). https://doi.org/10.1007/s13347-022-00543-1.

Drucker, J. (2011). "Humanities Approaches to Graphical Display", *Digital Humanities Quarterly* 5(1). https://dhq.digitalhumanities.org/vol/5/1/000091/000091.html.

Elam, M. (2023). “Signs Taken for Wonders: AI, Art, and the Matter of Race”, in Browne, J., Cave, S., Drage, E., and McInerney, K. (eds.) *Feminist AI: Critical Perspectives on Data, Algorithms and Intelligent Machines.* Oxford: Oxford University Press, pp. 237-257.

Feldstein, S. (2019). “How Artificial Intelligence Systems Could Threaten Democracy.” *The Conversation*, 22 April. Available at: http://theconversation.com/how-artificial-intelligence-systems-could-threaten-democracy-109698 [Accessed on February 13, 2025].

Gengler, E. J. (2024). “Sexism, Racism, and Classism: Social Biases in Text-to-Image Generative AI in the Context of Power, Success, and Beauty”, *Wirtschaftinformatik 2024 Proceedings*. 48. https://aisel.aisnet.org/wi2024/48.

Goldman, D. (2025). “What is DeepSeek, the Chinese AI Startup that Shook the Tech World?” *CNN*, 27 January. Available at: https://edition.cnn.com/2025/01/27/tech/deepseek-ai-explainer/index.html [Accessed on March 2, 2025].

Hampton, L. M. (2023). “Techno-Racial Capitalism: A Decolonial Black Feminist Marxist Perspective”, in Browne, J., Cave, S., Drage, E., and McInerney, K. (eds.) *Feminist AI: Critical Perspectives on Data, Algorithms and Intelligent Machines.* Oxford: Oxford University Press, pp. 119-136.

Han, W. (2018). "Proud of Myself as LGBTQ: The Seoul Pride Parade, Homonationalism, and Queer Develpmental Citizenship." *Korea Journal* 58(2): 27-57.

Hao, K. (2022). “Artificial Intelligence Is Creating a New Colonial World Order.” *MIT Technology Review*, 19 April. Available at: https://www.technologyreview.com/2022/04/19/1049592/artificial-intelligence-colonialism/ [Accessed on 26 August 2024].

Heaven, W. D. (2022). “How to Make a Chatbot That Isn’t Racist or Sexist”, in Martin, K. (ed.) *Ethics of Data and Analytics: Concepts and Cases*. Boca Raton: CRC Press, pp. 60-62.

Hong, S. (2025). “AI Textbook Introduction: Controversies and the Reasons behind the Failure of the Ministry of Education’s Education Policies” (in Korean), *Gyobasa*, 22 January. Available at: http://21erick.org/column/14519/ [Accessed on April 2, 2025].

Iyer, N., Chair, C., and Achieng, G. (2023). “Afrofeminist Data Futures”, in Browne, J., Cave, S., Drage, E., and McInerney, K. (eds.) *Feminist AI: Critical Perspectives on Data, Algorithms and Intelligent Machines.* Oxford: Oxford University Press, pp. 347-369.

Johnson, C. Y. (2022). “Racial Bias in a Medical Algorithm Favors White Patients over Sicker Black Patients”, in Martin, K. (ed.) *Ethics of Data and Analytics: Concepts and Cases*. Boca Raton: CRC Press, pp. 10-12.

Johnson, P., Menedez-Gonzalez, N., Duran, L. D. D., Panai, E., Kalpokiene, J. and Bertulfo, D. J. (2022). “The Ghost in the Machine Has an American Accent: Value Conflict in GPT-3.” *arXiv.* https://doi.org/10.48550/arXiv.2203.07785.

Kaplan, J., McCandlish, S., Henighan, T., Brown, T. B., Chess, B., Child, R., Gray, S., Radford, A., Wu, J., and Amodei, D. (2020). “Scaling Laws for Neural Language Models”, *arXiv*. https://doi.org/10.48550/ARXIV.2001.08361.

Kim, E. (2009). “Minority Politics in Korea: Disability, Interraciality, and Gender” in Grabham, E., Cooper, D., Krishnadas, J., and Herman, D. (eds.) *Intersectionality and Beyond: Law, Power and the Politics of Location*. New York, NY: Routledge.

Kim, H. (2015). “Sexual Dissidents and the Publicity of Urban Space: A Case Study on ‘2014 Sinchon Queer Parade’ in Seoul, South Korea”, *공간과 사회 (Space and Environment),* 51, pp. 12-62.

Kim, M. (2020). “The Politics of the Abolition of Women’s Student Councils and the Depoliticized Campus in Korea”, *Journal of Asian Sociology,* 49(4), pp. 449-476.

Kim, W. (2024). ""Dokdo Is a Disputed Territory" AI Textbook Response - Government: "Not a Serious Error" ("독도는 분쟁지역" AI 교과서 답변에 정부 "심각한 오류 아냐")", Kyunghyang Shinmun, 28 November. Available at: https://www.khan.co.kr/article/202411271734001 [Accessed on February 22, 2025].

Ko, H. K., Lee, I. S., and Mok, M. Y. (2024). “Analysis of AI Regulatory Frameworks in South Korea.” *Law.Asia*, 15 April. Available at: https://law.asia/ai-regulatory-frameworks-south-korea/ [Accessed on November 27, 2024].

Kulesz, O. (2024). *Artificial Intelligence and International Cultural Relations: Challenges and Opportunities for Cross-Sectoral Collaboration*. Stuttgart: ifa. https://doi.org/10.17901/1203.

Lee, H. (2015). "Reconstruction of the Struggle for Rigth to the City of Sexual Minority through the Concept of Post-Traditional Solidarity - Focused on Honneths Theory of Recognition and Butlers Critique." *Journal of Pan-Korean Philosophical Society* 78: 355-379. (In Korean)

Lim, J. (2021). "AI 'Lee Luda' Has Been Stopped... But Gender Discrimination and Hate Return to Humans (AI '이루다' 멈췄지만... 성차별.혐오는 인간에게 돌아온다)." *Hankyoreh*, 11 January. Available at: https://www.hani.co.kr/arti/society/women/978313.html [Accessed on December 2, 2024].

Lohr, S. (2022). "Facial Recognition Is Accurate, If You're a White Guy", in Martin, K. (ed.) *Ethics of Data and Analytics: Concepts and Cases*. Boca Raton: CRC Press, pp. 143-147.

Lu, D. (2025). "We Tried Out DeepSeek. It Worked Well until We Asked it about Tiananmen Square and Taiwan.", *The Guardian*, 28 January. Available at: https://www.theguardian.com/technology/2025/jan/28/we-tried-out-deepseek-it-works-well-until-we-asked-it-about-tiananmen-square-and-taiwan [Accessed on February 15, 2025].

Luccioni, A. S., Jernite, Y., and Strubell, E. (2023). "Power Hungry Processing: Watts Driving the Cost of AI Deployment?", *arXiv*. https://doi.org/10.48550/ARXIV.2311.16863.

Lugones, M. (2010). "Toward a Decolonial Feminism", *Hypathia* , 25(4), 742-759. https://www.jstor.org/stable/40928654.

Mackenzie, J., and Choi, L. (2024). "Inside the Deepfake Porn Crisis Engulfing Korean Schools", *BBC*, 3 September. Available at: https://www.bbc.com/news/articles/cpdlpj9zn9go [Accessed on November 10, 2024].

Mezzadra, S., and Neilson, B. (2019). *The Politics of Operations: Excavating Contemporary Capitalism*. Durham, NC: Duke University Press.

Mollona, M. (2021). Art/Commons: Anthropology beyond Capitalism. London: ZED.

Ministry of Education. (2023). *Briefing on the Plan for AI Digital Textbooks*. Available at: https://english.moe.go.kr/boardCnts/viewRenewal.do?m=0202&s=english&page=2&boardID=254&boardSeq=95291&lev=0&opType=N [Accessed on November 11, 2024].

Ministry of Science and ICT. (2024). "A New Chapter in the Age of AI: Basic Act on AI Passed at the National Assembly's Plenary Session", Press Release, 26 December. Available at: https://www.msit.go.kr/eng/bbs/view.do?sCode=eng&mId=4&mPid=2&pageIndex=&bbsSeqNo=42&nttSeqNo=1071&searchOpt=ALL&searchTxt= [Accessed on March 2, 2025].

Nguyen, C. T. (2024). "The Limits of Data", *Issues in Science and Technology* 40(2), 1 January, pp. 94–101. https://doi.org/10.58875/LUXD6515.

Nicholas, G., and Bhatia, A. (2023). "Lost in Translation: Large Language Models in Non-English Content Analysis", *arXiv*. https://doi.org/10.48550/ARXIV.2306.07377.

Noble, J. (2025). "What is LoRA (low-rank adaption)?", *IBM*, 28 January. Available at: https://www.ibm.com/think/topics/lora#:~:text=LoRA%20adds%20low%2Drank%20matrices,tuning%20or%20training%20large%20models [Accessed on July 31, 2025].

Oh, S. (2023). "Have They Forgotten About Iruda's Gender-Biased Learning? AI Bills Lack Gender Equality Considerations. ('이루다' 성차별 학습 잊었나... AI 법안에 성평등 관점 없다)", *Hankyoreh*, 6 August. Available at: https://www.hani.co.kr/arti/society/society_general/1103232.html [Accessed on November 11, 2024].

Olson, P. (2022). "The Quiet Growth of Race-Detection Software Sparks Concerns Over Bias", in Martin, K. (ed.) *Ethics of Data and Analytics: Concepts and Cases*. Boca Raton, London and New York: CRC Press, pp. 201-205.

Perrigo, B. (2023). "Exclusive: Open AI Used Kenyan Workers on Less Than $2 Per Hour to Make ChatGPT Less Toxic", TIME, 18 January. Available at: https://time.com/6247678/openai-chatgpt-kenya-workers/ [Accessed on May 18, 2025].

Peszkowska, A., and Tarkowski, A. (2024). "Common Corpus: Building AI as Commons." *Open Future*, 2 May. Available at: https://openfuture.eu/blog/common-corpus-building-ai-as-commons [Accessed on December 1, 2024].

Rahman, M., Yan, J., and Fapi, E. T. (2024). "Adversarial Artificial Intelligence in Blind False Data Injection in Smart Grid AC State Estimation", *IEEE Transactions on Industrial Informatics* 20, no. 6 (June): pp. 8873–83. https://doi.org/10.1109/TII.2024.3374374.

Rutazibwa, O. U. and Shilliam, R. (2018). *Routledge Handbook of Postcolonial Politics*. New York and London: Routledge.

Schneier, B., and Sanders, N. E. (2023). "Build AI by the People, for the People", *Foreign Policy*, August 26. Available at: https://foreignpolicy.com/2023/06/12/ai-regulation-technology-us-china-eu-governance/ [Accessed on August 13, 2025].

Sarp, S., Kuzlu, M., Cali, U., Elma, O., and Guler, O. (2021). "Analysis of False Data Injection Impact on AI Based Solar Photovoltaic Power Generation Forecasting", *arXiv*, https://doi.org/10.48550/ARXIV.2110.09948.

Seo, J., and Cho, Y. C. (2021). "The Emergence and Evolution of International Relations Studies in Postcolonial South Korea", *Review of International Studies,* 47(5), pp. 619–636. https://doi.org/10.1017/S0260210521000504.

Seo, J. and Choi, S. (2020). "Introduction", *Journal of Asian Sociology,* 49(4), pp. 501-525.

Siapera, E. (2022). "AI Content Moderation, Racism and (de)Coloniality", *International Journal of Bullying Prevention,* 4, pp. 55-65. https://doi.org/10.1007/s42380-021-00105-7.

Shumailov, I., Shumaylov, Z., Zhao, Y., Gal, Y., Papernot, N., and Anderson, R. (2023). "The Curse of Recursion: Training on Generated Data Makes Models Forget", *arXiv*. https://doi.org/10.48550/ARXIV.2305.17493.

Sinders, C. (n.d.). *Feminist Data Set.* Available at: https://carolinesinders.com/wp-content/uploads/2020/05/Feminist-Data-Set-Final-Draft-2020-0526.pdf [Accessed on July 28, 2025].

Sohl-Dickstein, J., Weiss, E. A., Maheswaranathan, N., & Ganguli, S. (2015). *Deep Unsupervised Learning using Nonequilibrium Thermodynamics* (No. arXiv:1503.03585). arXiv. https://doi.org/10.48550/arXiv.1503.03585.

Srnicek, N. (2017). *Platform Capitalism*. Cambridge: Polity Press.

Taulli, T. (2024). "3 Most Common Problems with Small Language Models", *AI Business.* Available at: https://aibusiness.com/nlp/3-most-common-problems-with-small-language-models [Accessed August 26, 2024].

Thatcher, J., O'Sullivan, D., and Mahmoudi, D. (2016). "Data Colonialism Through Accumulation by Dispossession: New Metaphors for Daily Data", *Environment and Planning D: Society and Space*, 34(6), pp. 990-1006. https://doi.org/10.1177/0263775816633195.

Torchio, G., and Tasin, F. (2025). "The Paris Summit: Au Revoir, Global AI Safety?", *European Policy Centre*, 14 February. Available at: https://www.epc.eu/publication/The-Paris-Summit-Au-Revoir-global-AI-Safety-61ea68/ [Accessed on June 20, 2025].

Toupin, S. (2024). "Shaping Feminist Artificial Intelligence", *New Media & Society* 26(1) (January), pp. 580–595. https://doi.org/10.1177/14614448221150776.

Ulnicane, I. (2024). "Intersectionality in Artificial Intelligence: Framing Concerns and Recommendations for Action", *Social Inclusion* 12. https://doi.org/10.17645/si.7543.

UNESCO and International Research Centre on AI (IRCAI). (2024). *Challenging Systematic Prejudices: An Investigation into Bias against Women and Girls in Large Language Models*. Available at: https://unesdoc.unesco.org/ark:/48223/pf0000388971 [Accessed on June 20, 2025].

Villalobos, P., Ho, A., Sevilla, J., Besiroglu, T., Heim, L., and Hobbhahn, M. (2022). "Will We Run out of Data? Limits of LLM Scaling Based on

Human-Generated Data", *arXiv*. https://doi.org/10.48550/ARXIV.2211.04325.

Widder, D. G., Whittaker, M., and West, S. M. (2024). "Why 'open' AI systems are actually closed, and why this matters", *Nature*, 635(8040), pp. 827-833. https://doi.org/10.1038/s41586-024-08141-1.

Williams, A., Miceli M., and Gebru, T. (2022). "The Exploited Labor Behind Artificial Intelligence: Supporting Transnational Worker Organizing Should Be at the Center of the Fight for 'Ethical AI.'", *NOEMA Magazine*, 13 October. Available at: https://www.noemamag.com/the-exploited-labor-behind-artificial-intelligence/ [Accessed on November 11, 2024].

BigScience Workshop, Le Scao, T., Fan, A., Akiki, C., Pavlick, E., Ilić, S., et al. (2022). "BLOOM: A 176B-Parameter Open-Access Multilingual Language Model", *arXiv*. https://doi.org/10.48550/ARXIV.2211.05100.

Wudel, A., and Ehrenberg, A. (2025). *What is Feminist AI?*. Friedrich-Erbert-Stiftung Competence Centre on the Future of Work. Available at: https://library.fes.de/pdf-files/bueros/bruessel/21888-20250304.pdf [Accessed on November 11, 2024].

Zuboff, S. (2019). The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power. New York, NY: Public Affairs.

Appendix 1. Workshop Programme

"Feminist and Decolonial Approaches to AI: From Global Politics to Knowledge Production/Cultivation" Workshop

- Dates and times:
  - Day 1: December 20, 2024 | 14:00 - 17:00
  - Day 2: December 23, 2024 | 14:00 - 17:00
  - Day 3: December 30, 2024 | 10:00 - 12:30, 14:00 - 17:00
- Location: Yonsei University, Yonhee Hall, Room 201 [TBC]
- Hosts: ifa Research Programme "Culture and Foreign Policy", Yonsei University | Critical Approaches to Political Science Lab, Department of Political Science, Yonsei University
- Number of Participants: 15
- Target group: Social scientists, IT/AI specialists, students, artists.

Background

Globally, there are not only demands for creating systems that lead to more equitable forms of AI, but also a high demand for communicating the workings, issues, and potentials of AI in various sectors. Recently, an increasing number of research points to the biases and potential harms of AI for vulnerable populations such as women of colour, non-binary people, immigrants, and refugees. Concerns are also rising around the growing discrepancies between the Global North and the Global South, with the concentration of AI industries in the Global North as the forefront producers of AI innovation and the position of the Global South as mainly a consumer of such systems. With these apprehensions, researchers and AI practitioners have been calling for a feminist and decolonial imagination and ethical (re-)consideration of AI technologies.

Objectives and Workshop Structure

The workshop aims to create space for conversations on feminist and decolonial approaches to AI. On the first day of the workshop, with the participation of political scientists and AI specialists, we will critically question AI and its impacts on knowledge production and politics, as well as in the global landscape of AI development. Further, we will discuss possibilities of feminist and decolonial interventions in the understanding and use of AI, seeking a form of participatory, intersectional, decolonial and feminist AI. These conversations

will be followed by a two-day workshop, where the participants will be able to actively engage with AI by creating their own datasets and turning them into models for image generation. This participation in "critical data injection" will be followed by conversations on possibilities of AI interventions from different positionalities and of an AI of the commons.

In continuing the conversations from Day 1 of the workshop, participants will engage with critical data injection by taking a participative, collaborative and co-construction approach to AI. Participants will collect and create—or curate—their own data to train AI and reflect on the possibilities of making interventions to AI by de-biasing its models from a feminist and decolonial perspective.

We will first discuss how particular data are generated and provided in relation to particular themes by AI and what sort of biases and potential harms such data generation reproduces. Then, participants will create and add to AI their own images (e.g., drawings, photos) on these themes in a way that reflects on the positionalities of participants and in a way that goes against the biases that are currently reproduced by the AI model.

In between the dates of the workshop, we hold office hours for participants so that we can assist them with their data collection and AI model training and expand conversations on critical data injection.

Day 1 (20 December 2024, 14:00 - 17:00)

Feminist and Decolonial Approaches to AI: From Global Politics to Knowledge Production/Cultivation

Introduction and Critical Conversations on AI
The first day of the workshop will start with a critical conversation on AI with specialists from the fields of political science, international relations, and AI. Through these conversations, we will introduce the conceptual framework of our project and set the grounds for feminist and decolonial interventions to AI in the following two days of the workshop.

Chair: Seoyoung Choi
Discussant: Prof. Jungmin Seo (Yonsei University) and Seoyoung Choi

Presenters and Programme:

1. Miro Leon Bucher and Seoyoung Choi: Introduction to the research project
2. Prof. Young Chul Cho (Kangwon National University) – Presentation: *The Modernity/Coloniality of AI: A Personal Speculation*. (Keywords: Modernity, coloniality, decolonisation, species, AI)
3. Prof. Shang E. Ha (Sogang University) – Presentation: *Can AI be implicitly prejudiced?* | It is well-documented that artificial intelligence (AI) has the potential to reinforce pre-existing biases such as racism and sexism because it relies on machine learning techniques that consume a large amount of text data tainted by individual and institutional prejudice. Much ink has been spilled to find out how to reduce or even eliminate explicit or blatant prejudice that AI reproduces, but little is known whether AI can be implicitly prejudiced. This essay aims at critically evaluating two mutually exclusive arguments. One is that AI can be implicitly prejudiced, even though implicit bias, by definition, cannot be expressed in the texts. The other argument is that AI is free from implicit prejudice because it simply replicates what it learned without consciousness.
4. Chaewon Yun (Max Planck Institute for Human Development) – Presentation: *AI Safety's Blind Spot: Gendered Technology Facilitated Abuse (TFA)* | Myriad efforts to make AI 'better' are being made under the umbrella of fairness, safety, ethics and alignment, among other 'essentially contested (Gallie 1956)' terms. A more critical question is: On whose term of AI safety (Lazar and Nelson 2023), fairness, and alignment are defined, operationalised, and measured? In my presentation, I draw on discourses from computer science, science and technology studies (STS), and criminology to identify the blind spot of the current paradigm of AI safety. Specifically, I examine the case of 'deepfake porn' crimes through the lens of AI safety. What constitutes safety, for whom, and from what? Is it a failure of AI safety or rather a 'malign neglect'? By situating the discourse within the context of real-

world events, we examine how AI perpetuate or exacerbate existing power structures. *Chaewon Yun is a predoctoral fellow at the Center for Humans and Machines at the Max Planck Institute for Human Development in Berlin, Germany. Her PhD thesis critically challenges 'human-centered' AI both conceptually and empirically. She studied Computational Social Systems at RWTH Aachen University (MSc) and Political Science and International Studies (BA) at Yonsei University. In between academic studies, she worked as a software engineer at Samsung Electronics' AI Development Team.*

5. Discussion with all workshop participants

Day 2 & Day 3: Critical Data Injection as Feminist and Decolonial Interventions to AI

Participants will be introduced to concepts and practices of critical data injection and share their ideas on how to make their own interventions through critical data injection.

Day 2 (23 December 2024, 14:00-17:00)

Introduction to Critical Data Injection

1. Understanding the basics of diffusion-based image generation, answering the following questions:
    a. What is diffusion-based image generation, and what is Stable Diffusion in particular?
    b. How does Stable Diffusion "learn" what different terms "mean" visually?
    c. How can Stable Diffusion generate a theoretically infinite amount of images based on a finite dataset?

2. Live demonstration of LoRA training:
    a. Showing an example dataset and how to use TagGUI to create image-text pairs

b. Showing how to use Stability Matrix (free and open-source tool) to install AI tools
c. Using OneTrainer (free and open-source tool in Stability Matrix) to train a LoRA model based on the example dataset
d. Using ComfyUI (free and open-source tool in Stability Matrix) to generate images for A/B tests with and without the LoRA to visualise the impact of the model training

3. Discussing approaches for data collection based on the prior discussion and impressions
    a. What biases do participants expect in the original AI model?
    b. How do they want to impact them?
    c. What kind of data do they need to collect to cause the desired impact?

Participation Between Workshop Days

From December 24th to December 29th, the participants are expected to collect data on their own, i.e., collect and create images on their own and tag the images with TagGUI (or alternative tools). The researchers will also do their own critical injection during this period.

Office hours:
Times
- 26 December, 12:00 – 15:00
- 27 December, 12:00 – 15:00
- 28 December, 12:00 – 15:00
- 29 December, 12:00 – 15:00

Purpose:
- Discussing data collection
- Offering technical help
- Training LoRA models together
- Doing test generations

The office hours play a crucial role in the workshop schedule. Doing the AI model training fully during the workshop sessions is not possible since the

training is very time-consuming. By offloading this step to individual appointments, all participants receive enough attention and technical assistance to realise their desired outcomes of the workshop. The time that is not taken up by office hours serves us as researchers the critical function to do the same process as the participants by creating and collecting our own data and training our own model to have better participant observation and insight into the experience and to reflect on this better in our research report.

——

Day 3 (December 30, 10:00-12:30 & 14:00-17:00)

Testing and Reflections on Societal Implications

Morning Session (10:00-12:30):
- (Training LoRAs in the background for those who missed the office hours)
- Group Presentation: A/B testing of the base AI model versus the model enhanced with critical data injections using the LoRA (Low-Rank Adaptation) technique.
- Analysis: Groups analyse the outcomes, focusing on changes in model behaviour and personal experience.
- Reflection: Reflecting on the experience of critical data injection on a personal level.

Afternoon Session (14:00-17:00):
- Roundtable Discussion: Explore the potentials and limitations of critical data injection.
- Concluding Panel: Discuss how the findings from the workshop are relevant on a societal level, considering global and local contexts. Consider the next steps following this experience and its implications for an AI of the commons.

***

Appendix 2. LoRA Training Toolkit

This toolkit is designed to provide an accessible introduction to local AI image generation and training using open-source tools. It is aimed at cultural practitioners, researchers, artists, and other practitioners new to machine learning who wish to explore how generative AI models—specifically **Stable Diffusion** and **Low-Rank Adaptation (LoRA)** models—can be used in creative and academic contexts. This toolkit can also be used within the scope of in international cultural relations, particularly in artistic and/or cultural projects that aim to make AI more inclusive through (international) cross-sectoral collaboration with partners from different backgrounds.

The aim of this toolkit is to provide practical knowledge backed by a basic understanding of the underlying workings of AI. The toolkit discusses:

- Running image generation locally using Stable Diffusion
- Navigating open-source platforms such as **Stability Matrix** and **ComfyUI**
- Preparing datasets for training
- Training your own LoRA model using **OneTrainer**
- Generating images with your own AI model

The toolkit aims to equip you with both the practical tools and conceptual understanding necessary to engage meaningfully with local, open-source AI image generation and LoRA training. From installing Stability Matrix and ComfyUI, to building your own dataset, training a LoRA model using OneTrainer, and finally applying it through hands-on workflows, each step has been designed to demystify the process and empower you as a critical, creative user of AI technologies.

The toolkit places particular emphasis on transparency, local control, and ethical reflection. All software used is **free, open-source, and designed to run locally**, ensuring users maintain control over their data and workflows at all times. While many AI tools today rely on commercial application programming interfaces (APIs) or cloud services, this guide intentionally avoids those, favouring self-contained tools that respect user privacy and provide insight into how generative AI models actually function.

By focusing on locally run, transparent, and open-source systems, this guide encourages a move away from opaque, centralised AI platforms and toward user autonomy and deeper engagement. Whether you are a cultural practitioner, researcher, or simply curious, the ability to train and test your own models not only expands creative possibilities but also allows for a more nuanced critique of the biases, potentials, and limits embedded in generative models.

Furthermore, by following the basic technical explanations and step-by-step instructions to gain practical experience using local AI models, we want to help readers have a more deeply founded critical view of AI. Through this process, we hope to show that users with diverse backgrounds can become engaged actors in forming AI, questioning the top-down approach used by many commercial AI tools.

As AI becomes increasingly embedded in visual culture, design, and research, we believe it is essential that these tools are not only used—but understood, challenged, and reshaped by a broad community of practitioners. This toolkit is one small contribution toward that goal.

For better legibility, this toolkit is **only available in a digital version**. To access the toolkit, including all additional assets, please follow this link:

https://miroleon.github.io/lora-training-toolkit/

Disclaimer: Please note that by using this toolkit, you accept full responsibility for any consequences that may arise. The authors and the publisher of this toolkit (Institut für Auslandsbeziehungen – ifa) cannot be held liable for any issues related to the use or misuse of the tools described herein. Use this toolkit at your own risk.

About the Authors

Seoyoung Choi (she/her) is a political scientist based in Seoul, South Korea, whose research centres on decolonial feminist international relations theory and methodology. She holds a PhD in Political Science from Yonsei University. She is currently a Visiting Researcher at the Social Science Research Institute at Yonsei University and an Adjunct Lecturer at Kangwon National University and University of Seoul. Contact: syseochoi@gmail.com.

Miro Leon Bucher (they/he) is a media artist and researcher exploring the ethical and aesthetic dimensions of emerging technologies, with a particular focus on open-source AI in art and architecture. They hold an M.A. in Philosophy and Social and Cultural Anthropology from the University of Cologne and currently work as an Artistic Scientific Associate in the Academic and Research Department for Artistic Methods at RWTH Aachen University. Their transdisciplinary practice spans artistic research, curation, and production, drawing on utopian thinking to explore visions for the future. Bucher has lived and worked between Germany and the Republic of Korea, contributing to international projects at the intersection of art, science, and technology. Contact: contact@miroleon.de.

Imprint

The study is created within the framework of ifa's Research Programme "Culture and Foreign Policy" and is published in the ifa Edition Culture and Foreign Policy.
The Research Programme is funded by the Federal Foreign Office.
The views expressed are those of the author and do not necessarily reflect those of ifa.

Publisher:
ifa (Institut für Auslandsbeziehungen e. V.),
Charlottenplatz 17,
70173 Stuttgart,
P.O. Box 10 24 63, D-70020 Stuttgart,
info@ifa.de,
www.ifa.de

Authors:
Seoyoung Choi, Miro Leon Bucher

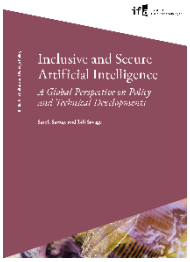
Editing:
ifa's Research Programme "Culture and Foreign Policy" and Vera Draack

Cover Photo Credits:
Pexels/Google DeepMind & XK Studio | An artist's illustration of artificial intelligence (AI), depicting AGI (artificial general intelligence)'s potential to enrich lives. It was created by XK Studio as part of the Visualising AI Programme, an initiative of Google DeepMind.

Design:
Andreas Mayer, Stuttgart

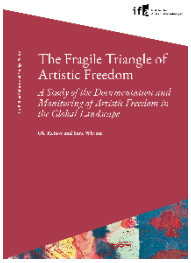
DOI: https://doi.org/10.17901/1571

Other ifa Edition Culture and Foreign Policy publications



Saiph Savage, Lili Savage: Inclusive and Secure Artificial Intelligence. A Global Perspective of Policy and Technical Developments, Stuttgart: ifa, 2025 (ifa Edition Culture and Foreign Policy)
https://doi.org/10.17901/1570

Octavio Kulesz: Artificial Intelligence and International Cultural Relations. Challenges and Opportunities for Cross-Sector Collaboration, Stuttgart: ifa, 2024 (ifa Edition Culture and Foreign Policy)
https://doi.org/10.17901/1203



Ole Reitov, Sara Whyatt: The Fragile Triangle of Artistic Freedom. A Study of the Documentation and Monitoring of Artistic Freedom in the Global Landscape, Stuttgart: ifa, 2024 (ifa Edition Culture and Foreign Policy)
https://doi.org/10.17901/1201

Manouchehr Shamsrizi: Metaverse and Gaming, Potentials for Foreign Cultural and Educational Policy, Stuttgart: ifa, 2023 (ifa Edition Culture and Foreign Policy)
https://doi.org/10.17901/957

More publications are available at: culturalrelations.ifa.de/en/research

Democratising Artificial Intelligence through Culture

Making Generative AI Participatory and Intersectional Through an AI of the Commons

This research proposes a participatory, decolonial and feminist approach to AI through the concept of "AI as commons", offering a practical approach for cultural practitioners to decolonise existing AI systems. Conducted in Seoul, South Korea, the study combines critical theory with practical intervention, introducing "critical data injection" as a method for communities to influence representation in generative AI. At a workshop at Yonsei University, participants created datasets of everyday Seoul images which challenged typical AI-generated urban imagery, then trained Low-rank Adaptation (LoRA) models to generate alternative visual narratives. The accompanying LoRA Training Toolkit makes AI model training accessible to non-expert users. While acknowledging limitations, such as dependence on biased base models and infrastructure constraints, the research demonstrates how local communities can create representational alternatives within existing AI systems. The study argues for publicly-funded, modular AI infrastructure which enables democratic participation in AI development rather than relying solely on corporate or state-controlled systems.

ifa Institut für Auslandsbeziehungen