This lexicon was first drafted as material for the first Digital Rights for All workshop called “Talking Digital”. We struggled back then to find one document with different levels of definitions for tech jargon. Most notions in this lexicon were part of our first workshop activity called “Days of our lives”. During this exercise, the participants were to match notions such as Algorithm, interoperability, big tech etc. to a situation described in the text, as we followed one person going through their day.

**This revised lexicon has a similar ambition:** having one document where definitions of often used tech vocabulary can easily be found and highlighted through their social justice lens. Therefore, the lexicon offers multiple interpretations for these definitions. It is meant to propose different approaches to the concepts and invite the reader to use the one best suited to their unique context.

**In this lexicon, you will find technical, legal, historical, sociological, and philosophical definitions.** You will also see short descriptions and longer analyses. There are varied sources, from well-known free-to-use online encyclopaedias, non-profit reports, and legal texts to news articles, research articles, and essays.

**Many of the definitions in the Lexicon come from the work of black and brown scholars, activists, and collectives.** They have been, and continue to be, pioneers in the study and function of how racial and social injustices are also embedded in digital realities. They are part of traditions that have constantly affirmed that defining, far from being a neutral act, is indeed a political one.

We hope this lexicon will support organisers, individuals, and activists in finding suitable definitions. Hopefully, it will inspire new ones.

Introduction by
Laurence Meyer,
DFF Social & racial justice lead.

This lexicon was compiled by:
Adélaïde Hirwe
César Manso-Sayao
Laurence Meyer

“**As designers, we wanted to help visualize the various topics touched upon in this lexicon, from systemic oppressions to capitalism drifts, by appealing to our own subjectivity.**

**We imagined an ironical universe mixing the narratives of digital self-defense or non-global internet, and borrowing from science fiction as well as both web and anti-racism, environmental activism culture and imagery.**

**We are pleased to highlight the political messages conveyed by the Talking Digital Lexicon. We hope that our part will serve anyone who wants to overcome the current digital world’s challenges, and build a fairer one.”**

Designers of the Talking Digital Lexicon
Claire Zaniolo
Estelle Pom
“A series of steps that allow you to perform a particular task. Every algorithm takes in defined inputs (the things being acted upon) and has the goal of producing defined outputs (the results you want). Different algorithms can exist for accomplishing the same task. Algorithms are often judged by their efficiency, and efficiency can be science, an algorithm usually indicates a mathematical procedure for solving a recurrent problem.”

“Part of the challenge of understanding algorithmic oppression is to understand that mathematical formulations to drive automated decisions are made by human beings. While we often think of terms such as ‘big data’ and ‘algorithms’ as being benign, neutral, or objective, they are anything but. The people who make these decisions hold all types of values, many of which openly promote racism, sexism, and false notions of meritocracy, which is well documented in studies of Silicon Valley and other tech corridors.”

In addition to classical sequential algorithms, in use from antiquity, we have now parallel, interactive, distributed, real-time analog, hybrid, quantum, etc. algorithms, evaluated in different ways.”
“They are, after all, programmed using algorithms that are constantly updated on the basis of human behavior and are learning and replicating the technology of race, expressed in the many different associations that the users make. This issue came to light in 2016, when some users searched the phrase ‘three Black teenagers’ and were presented with criminal mug shots.

Then when they changed the phrase to ‘three White teenagers’, users were presented with photos of smiling, go-lucky youths; and a search for ‘three Asian teenagers’ presented images of scantily clad girls and women.

Taken together, these images reflect and reinforce popular stereotypes of Black criminality, White innocence, or Asian women’s sexualization that underpin much more lethal and systemic forms of punishment, privilege, and fetishism respectively.”

“The definition of ‘algorithmic bias’ is a hotly contested topic in the ML (machine learning) literature, so here I will define it broadly as when ML algorithms systematically perform less well for or penalize certain subgroups.”

“Taken together, these images [...] reinforce popular stereotypes of Black criminality, White innocence, or Asian women’s sexualization that underpin much more lethal and systemic forms of punishment, privilege, and fetishism respectively.”
“Automated decision-making is when a computer makes a decision based on information it has gathered about you; a person uses information gathered by a computer to make a decision.”

“Automated decision-making is when:
- a computer makes a decision based on information it has gathered about you
- a person uses information gathered by a computer to make a decision. Automated decision-making is also called ADM for short.”

“Algorithmic decision-making” refers to decisions made using an ML or other statistical model trained on data. These decisions can be completely automated or used to inform human decision-makers.”

“Automated decision-making is a type of automated processing facilitated by technological means without human involve-
Artificial Intelligence (AI)

“The use of computer technology to make computers and other machines think and do things in the way that people can.”

SABELO MHLAMBI
“From Rationality to Relationality: Ubuntu as an Ethical and Human Rights Framework for Artificial Intelligence Governance”
CARRCENTER.HKS. HARVARD.EDU
CARR CENTER FOR HUMAN RIGHTS POLICY, HARVARD KENNEDY SCHOOL, NO. ISSUE 2020-009, PP. 1–27.
JULY 2020

“In a sense, AI represents a quest to imitate the human brain. It’s about making a machine brain that can mimic the kinds of tasks that we think are unique to humans. In other words, AI is about developing machines that can do what humans can. Considering how complicated the human brain is, it’s no wonder that we have only recently been able to get machines to do this. After all, the human brain is so complex that we as humans have not yet tapped into its full potential. What we do understand about the brain is that there are multiple intelligences within humans and that our social environment influences how these intelligences are regarded. If AI is about imitating human intelligence, which forms of intelligence are we actually trying to mimic?”

DIANA NUCERA
AND MIMI ONUOHA
“The People’s Guide to AI”
ALLIEDMEDIA.ORG
2020

“Artificial: Intelligence”
MACMILLAN-DICTIONARY.COM
2022

“The use of computer technology to make computers and other machines think and do things in the way that people can.”

Today’s salvation, deeply motivated by rationality as personhood, and enabled by capitalism and modern colonialism, is the use of artificial intelligence to automate decision making about the lives of humans. It is the same weaponization of rationality that has dominated Euro-American conquests.

The belief in the neutrality of automated decision-making systems is deeply misguided and shares the same flaws and contradictions of its predecessors. The negative effects of ADMS on groups historically marginalized by Euro-American modernity affirms the dehumanizing effects of basing the essence of personhood on rationality.”

DIANA NUCERA
AND MIMI ONUOHA
“The People’s Guide to AI”
ALLIEDMEDIA.ORG
2020

“Artificial Intelligence Regulation. Article 3/Annex I”
EUROPEAN-COMMISSION.ORG
2021

| Machine learning approaches, including supervised, unsupervised and reinforcement learning, using a wide variety of methods including deep learning; |

| Logic- and knowledge-based approaches, including knowledge representation, inductive (logic) programming, knowledge bases, inference and deductive engines, (symbolic) reasoning and expert systems; |

| Statistical approaches, Bayesian estimation, search and optimization methods.” |
“Big Tech is a term that refers to the most dominant and largest technology companies in their respective sectors.”

“Big Tech refers to the major technology companies such as Apple, Google, Amazon, Facebook, and Microsoft, which have inordinate influence.”

“First, the Big Tech ‘visionaries’ invariably started with technologies based on research at institutions that rely on significant public funding. The venture capitalists did not take most of the risk, nor did the technologists create the basis for the market valuation of their developments on their own. Instead, as economist Mariana Mazzucato (2013) and others have shown, far from getting out of the way of private innovation the State paves the way. It is the State, not private capital, that funds the long-term, high-risk research and development (R&D) that underpins Big Tech.”

“Defund Big Tech, Refund Community, Anti-Trust is Not Enough, Another Tech is Possible”
TECHOTHERWISE, PUBPUB.ORG
2021
"A computer or similar piece of equipment that performs a complicated job, although the person using it does not understand how it works."

“AI models often are ‘black boxes’ which we are not able to understand. Relying on these models has a multifaceted impact and raises significant concerns about their transparency. Applications in sensitive and critical domains are a strong motivational factor in trying to understand the behavior of black boxes […] Relying on black box systems is becoming increasingly risky both for their lack of transparency and the systematic bias they have shown in real-world scenarios.”
Content moderation is when an online platform screen and monitor user-generated content based on platform-specific rules and guidelines to determine if the content should be published on the online platform, or not.

In other words, when content is submitted by a user to a website, that piece of content will go through a screening process (the moderation process) to make sure that the content upholds the regulations of the website, is not illegal, inappropriate, or harassing, etc.

Content moderation as a practice is common across online platforms that heavily rely on user-generated content, such as social media platforms, online marketplaces, sharing economy, dating sites, communities and forums, etc.

There are a number of different forms of content moderation, pre-moderation, post-moderation, reactive moderation, distributed moderation, and automated moderation.”
“The screening, evaluation, categorisation, approval or removal/hiding of online content according to relevant communications and publishing policies. It seeks to support and enforce positive communications behaviour online, and to minimise aggression and anti-social behaviour.”

Notably, Perspective decided a tweet from the virulent white nationalist Stefan Molyneux claiming ‘the three major races have different brain volumes and different average IQs,’ a blatantly hateful, ignorant, and inflammatory statement, was 21.7% toxic. Meanwhile, it decided that a tweet by drag queen Mayhem Miller that simply states, ‘I am black. I am gay. I am a man. I am a drag queen. If those are not enough for you… kindly, fuck off!!!’ was 95.98% toxic.”

As platforms grew and began to turn a profit, sex workers, who were some of the first to use these platforms (Patreon and Tumblr, for example), were then deemed high-risk and deplatformed. A shadowban ensures that content that platforms deem inappropriate, high-risk, or low value speech is invisible to other users, but still allows an account to remain on the platform in order to make ad revenue off of the shadowbanned individual. The shadowban is an integral part of surveillance capitalism, where the user is still on the metadata collection and surveillance matrix. The platform can still surveil, but the user loses their voice.”
“Data can be many different things. It can be digital, personal, interpersonal (person to person). It can be collected by us or from us, provided to us by companies, or taken from us by companies. Data can be used to make decisions about us, to craft or tell our stories, or even connect us to something, someone, a service or to be criminalized. Data and how we use or understand it is vast, as
we can see from the list we generated. ‘Data is: facts, details, statistics, or any information collected together for reference or analysis.’ Data is not all digital. Not all data is equal (created equally)."

“Data is kind of capital, on par with financial capital, for creating new digital products and services.”

“[...] On one hand, data is cast as a digital raw material – constant capital – necessary in the production of commodities. It is hard to read media articles and business reports about data without seeing it called ‘the new oil.’

“[...] On the other hand, data is cast as a commodity produced by the digital labour of people posting on Facebook, clicking on Google, exercising with Fitbits, and all the other things we do that create data and that data is created about.

“[...] I suggest a better framing of data is as a form of capital that is distinct from, but has its roots in, economic capital. Data capital is more than knowledge about the world; it is discrete bits of information that are digitally recorded, machine processable, easily agglomerated, and highly mobile. Like social and cultural capital, data capital is convertible, in certain conditions, to economic capital.

“[...] Data capital is institutionalised in the information infrastructure of collecting, storing, and processing data; that is, the smart devices, online platforms, data analytics, network cables, and server farms.

“[...] When data is treated as a form of capital, the imperative to collect as much data, from as many sources, by any means possible intensifies existing practices of accumulation and leads to the creation of new ones.”

“Data’s abstract nature makes it hard to conceptualise and talk about. And so we use metaphors. We speak of data being like oil, water or carbon dioxide; of...
The data generated by online activity, the quantities of which are so large that new tools and methods are needed to analyse it.”

“Data is kind of capital, on par with financial capital, for creating new digital products and services. Unlike the metaphors we use about data – that it’s the new oil, gold or new electricity – what we’re saying with data capital is that it fulfils the literal textbook definition of capital. It is a produced good, as opposed to a natural resource, and it’s a necessary input into other goods or services.”

“Big Data is a phrase used to mean a massive volume of both structured and unstructured data that is so large it is difficult to process using traditional database and software techniques. In most enterprise scenarios the volume of data is too big, or it moves too fast, or it exceeds current processing capacity.”
“[…] biometrics, in its simplest form, is a means of body measurement that is put to use to allow the body, or parts and pieces and performances of the human body, to function as identification. […] What I am suggesting here is that branding in the transatlantic slave trade was a biometric technology, as it was a measure of slavery’s making, marking, and marketing of the black subject as commodity.”

“Biometric data is anything that relates to the measurement of people’s physical features and characteristics. In digital identity terms, this data is used to prove a person’s uniqueness and verify that someone is who they say they are. […] However, biometric data also presents privacy concerns. Data breaches could have severe consequences, as users are not able to simply reset a password when their biometric data is stolen. Biometrics also present issues around privacy,…”

“Branding in the trans-atlantic slave trade was a biometric technology, as it was a measure of slavery’s making, marking, and marketing of the black subject as commodity.”
surveillance and state control, and biometric ID programmes have been met with mistrust from users because of the nature of the data being gathered by authorities."

“Biometric data is widely used in systems that attempt to identify a specific user or other human through unique characteristics. Computer image processing is one form of biometric analysis that uses biometric data. Digital fingerprint analysis also relies on the use of biometric data for identification purposes.

In most biometric analysis systems, there is demand for a large amount of biometric data. This data must be stored and somehow secured from unauthorized access. These systems rely on complex algorithms that sort data in ways that will achieve an identifying result in a given application. Developers use key features that are unique from one person to another in order to make biometric identification effective.”

“Discrete parts of our whole selves that are collected, stored in databases, the cloud, and other spaces of digitally networked flows, and used to make decisions or determinations about us. They are a manifestation of our relationships with our communities and institutions, including institutions of privilege, oppression, and domination.”

“Personal data is any information that relates to an identified or identifiable living individual. Different pieces of information, which collected together can lead to the identification of a particular person, also constitute personal data. Personal data that has been de-identified, encrypted or pseudonymised but can be used to re-identify a person remains personal data and falls within the scope of the GDPR.

Personal data that has been rendered anonymous in such a way that the individual is not or no longer identifiable is no longer considered personal data. For data to be truly anonymised, the anonymisation must be irreversible.”

“Article 4 (I): ‘personal data’ means any information relating to an identified or identifiable natural person (‘data subject’): an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person.”
“Digital describes systems that generate and process binary data. Computers are fundamentally digital machines because they process information that has been encoded as binary values either one that’s positive (represented as 1) or one that’s non-positive (represented as 0). These values, called bits, are combined to form bytes that serve as the foundation for all computer systems. […]

Similar to the Industrial Revolution that was characterized by new manufacturing processes in the 18th and 19th centuries, the Digital Revolution marked a global shift in culture and economics that was influenced by the rise of digital electronics. The Digital Revolution also signified the start of the Information Age.”

“1 | Composed of data in the form of especially binary digits.
2 | Characterized by electronic and especially computerized technology.”
1. Is there a face in the image?

Face or facial detection is the process of detecting the presence of faces and locating those faces in an image or video. Detecting the presence of a face and locating it in the image is not the same as assigning a unique identity to a detected face or trying to determine attributes like gender or age.

In particular, the process of face detection does not report anything about who someone is or what kind of person someone might be. It is merely the process of attempting to find and locate faces in an image. Subsequent analysis performed on a face often depends on the successful completion of face detection.

2. What kind of face is shown in the image?

Face attribute classification and face attribute estimation. Software can be developed to assess the attributes of a person from their face. When these attributes have been separated into distinct categories, such as gender, race, or ethnicity, this may be called face attribute classification.

When the attribute is a number, like an age, the term face attribute estimation is more appropriate. Finally, software to detect and locate accessories like glasses and scarves or face attributes like beards or moustaches may be referred to as face attribute detection.

Emotion, affect, and facial expression classification. Facial recognition technologies can be used to classify facial expressions, such as ‘smile’, ‘frown’, or ‘scowl’. They can also be used for the closely related problem of inferring the emotional state or effect of a person, such as ‘happy’, ‘sad’, or ‘angry’.

“Face identification software can only match the image of a person for whom it already has some appearance information.”

JOY BUOLAMWINI, VICENTE ORDÓÑEZ, JAMIE MORGENSTERN, AND ERIK LEARNED-MILLER
“Facial Recognition Technologies: A Primer”
GLOBAL-UPLOADS.WEBFLOW.COM
2020
The final category of applications is related to establishing the identity of a person.

**Face recognition or facial recognition** - the process of using digital representations of faces to try to identify or verify the identity of a unique individual. The image of a particular individual we wish to recognize is often referred to as the query image or a query. There are two subtly different types of recognition, referred to as face verification and face identification.

**Face verification is one type of face recognition.** It attempts to determine whether an image shows a particular person. For example, software on a cell phone may try to answer the question, ‘Can it be verified that the camera shows the phone’s owner?’ A query image is deemed to be either a match, if it appears to show the owner, or a mismatch otherwise.

There are two common ways to perform face verification. In the first, one asks a question such as ‘Does this image show Janelle Smith?’, in which the person of interest is named. A common use for this type of face verification is access control, such as software that allows the owner of a device or a service to access it. In the second common version of face verification, one is given two pictures and asks, ‘Is the first person the same as the second person?’ In this case, it is not necessary to know the identity of either person to answer the question. Face verification is also referred to as 1-to-1 matching or 1-to-1 comparison.

**Face identification attempts to answer the question, ‘Whose face is this?’**. Face identification software can only match the image of a face to a person for whom it already has some appearance information. The set of people for whom an application has stored appearance information is called the gallery.

Simply put, this is the set of people that a face identification system could possibly identify. Face identification can be used for surveillance, to find a person of interest, or for the identification of subjects that are either unable or unwilling to respond. It may be referred to as 1-to-many comparison, 1-to-many matching, 1-to-many identification, or 1-to-N identification.”
Interoperability is the property that allows for the unrestricted sharing of resources between different systems. This can refer to the ability to share data between different components or machines, both via software and hardware, or it can be defined as the exchange of information and resources between different computers through local area networks (LANs) or wide area networks (WANs).

Broadly speaking, interoperability is the ability of two or more components or systems to exchange information and to use the information that has been exchanged.

“Think of interoperability as a necessary, but not sufficient,
“At the click of a button, they are able to check whether he has applied for asylum before; [...] has any previous criminal records; whether police are actively searching for him and even how many times he has entered the EU before”

CRISTINA BLASI CASAGRAN
“Fundamental Rights Implications of Interconnecting Migration and Policing Databases in the EU”
HUMAN RIGHTS LAW REVIEW VOL. 21, ISSUE 2, PP. 433–457, 2021

“M.A. is a thirteen-year-old Syrian boy who has managed to get to the Greek border looking
“Essentially, ML is the capability of software or a machine to improve the performance of tasks through exposure to data and ex-

“Field of study that gives computers the ability to learn without being explicitly programmed.”
[This definition is often attributed to Samuel, who coined the term ‘machine learning’, but this is not found verbatim in this publication, and may be a paraphrase that appeared later.]”

“Essentially, ML is the capability of software or a machine to improve the performance of tasks through exposure to data and ex-

DAVID D. LUXTON
“Artificial Intelligence in Behavioral and Mental Health Care”
ELSEVIER/ACADEMIC PRESS
SCIENCEDIRECT.COM
2016

ARTHUR L. SAMUEL
“Some Studies in Machine Learning Using the Game of Checkers”
IBM JOURNAL OF RESEARCH AND DEVELOPMENT
“Essentially, Machine Learning is the capability of software or a machine to improve the performance of tasks through exposure to data and experience.”

Machine Learning (ML) is a sub-branch of Artificial Intelligence (AI) that enables computers to learn, adapt, and perform the desired functions on their own. ML algorithms can learn patterns from the previous input and results and adjust tasks accordingly. Machine learning can be categorized in one of three major ways.

| Supervised Learning: |
Uses labeled data that includes inputs and rectified outputs to train models.

| Unsupervised Learning: |
Uses unlabeled data to train models in which the output variable is unknown. Therefore, the models need to learn from the data, discover patterns, and provide the desired output.

| Reinforcement Learning: |
Algorithms need to learn from their environment, like human beings. It gets favorable or unfavorable rewards based on the environment, favorable or unfavorable rewards based on the environment.
“A range of policing practices that claim to use demographic, environmental and historic crime data to predict future patterns of crime”

“Predictive Policing refers to a range of policing practices that claim to use demographic, environmental and historic crime data to predict future patterns of crime as well as presuming "where will the crime occur", and "who" will commit the crime. Predictive policing can be better understood within the broader creep of data-surveillance on the part of law-enforcement.”

“At a recent workshop sponsored by a grassroots organization called Stop LAPD Spying, the facilitator explained that community members with whom she works might not know what algorithms are, but they know what it feels like to be watched.

Feelings and stories of being surveilled are a form of 'evidence', she insisted, and community testimony is data. As part of producing those data, the organizers interviewed people about their experiences with surveillance and their views on predictive policing. They are asked, for example: ‘What do you think the predictions are based on?’

One person, referring to the neighborhood I grew up in, responded: ‘Because they over-patrol certain areas – if you're only looking on Crenshaw and you only pulling Black people over then it's only gonna make it look like, you know, whoever you pulled over or whoever you searched or whoever you criminalized that's gonna be where you found something.’”
“Technology is the result of accumulated knowledge and application of skills, methods, and processes used in industrial production and scientific research. [...] Technology means ‘science of craft’, from Greek τέχνη, techne, ‘art, skill, cunning of hand’; and -λογία, -logia.”

“Race itself is a kind of technology— one designed to separate, stratify, and sanctify the many forms of injustice experienced by members of racialized groups”
“The belief that every problem has a solution based in technology.”

“Recasting all complex social situations either as neatly defined problems with definite, computable solutions or as transparent and self-evident processes that can be easily optimized—if only the right algorithms are in place!—this quest is likely to have unexpected consequences that could eventual-

I call the ideology that legitimates and sanctions such aspirations ‘solutionism.’ I borrow this unabashedly pejorative term from the world of architecture and urban planning, where it has come to refer to an unhealthy preoccupation with sexy, monumental, and narrow-minded solutions—the kind of stuff that wows audiences at TED Conferences—to problems that are extremely complex, fluid, and contentious...

Design theorist Michael Dobbins has it right: solutionism presumes rather than investigates the problems that it is trying to solve, reaching ‘for the answer before the questions have been fully asked.’ How problems are composed matters every bit as much as how problems are resolved.”
<table>
<thead>
<tr>
<th>INDEX</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
</tr>
<tr>
<td>ALGORITHM</td>
</tr>
<tr>
<td>ALGORITHMIC BIAS</td>
</tr>
<tr>
<td>AUTOMATED DECISION-MAKING</td>
</tr>
<tr>
<td>ARTIFICIAL INTELLIGENCE</td>
</tr>
<tr>
<td>B</td>
</tr>
<tr>
<td>BIG TECH</td>
</tr>
<tr>
<td>BLACK BOX</td>
</tr>
<tr>
<td>C</td>
</tr>
<tr>
<td>CONTENT MODERATION</td>
</tr>
<tr>
<td>D</td>
</tr>
<tr>
<td>DATA</td>
</tr>
<tr>
<td>BIG DATA</td>
</tr>
<tr>
<td>BIOMETRIC DATA</td>
</tr>
<tr>
<td>DATA BODY</td>
</tr>
<tr>
<td>PERSONAL DATA</td>
</tr>
<tr>
<td>DIGITAL</td>
</tr>
<tr>
<td>F</td>
</tr>
<tr>
<td>FACIAL RECOGNITION TECHNOLOGIES</td>
</tr>
<tr>
<td>I</td>
</tr>
<tr>
<td>INTEROPERABILITY</td>
</tr>
<tr>
<td>M</td>
</tr>
<tr>
<td>MACHINE LEARNING</td>
</tr>
<tr>
<td>P</td>
</tr>
<tr>
<td>PREDICTIVE POLICING</td>
</tr>
<tr>
<td>T</td>
</tr>
<tr>
<td>TECHNOLOGY</td>
</tr>
<tr>
<td>TECHNOLOGICAL SOLUTIONISM</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>BARENDREGT Wolmet, BECKER Christoph, CHEON EunJeong, CLEMENT Andrew, REYNOLDS-CUÉLLAR Pedro, SCHULER Douglas, and SUCHMAN Lucy. 2021. Defund Big Tech. Refund Community Anti-Trust is Not Enough, Another Tech is Possible [online]. Available at: <a href="https://techotherwise.pubpub.org">https://techotherwise.pubpub.org</a>.</td>
</tr>
</tbody>
</table>