GLOSSARY

This glossary consists of key terms from the digital world and digitized workplaces. They are in alphabetical order.

Algorithm: In its simplest form, an algorithm is a set of rules in computer programming code to solve a problem or perform a task. An algorithm is fed data. It is a person, such as a manager or programmer, who sets the goal of the task or problem and writes algorithms to tell the computer system what to do and how to do it. These are the so-called instructions. The computer system is then able to independently perform tasks by following the instructions outlined by the algorithm. (For more information, see: What is an algorithm? See also the section below "*Machine Learning"*

Algorithmic system: An algorithmic system is a system that uses one or more algorithms, usually as part of computer software, to produce output that can be used to make decisions.

Algorithmic management: The concept of algorithmic management can be broadly defined as the delegation of management functions to algorithmic and automated systems. According to the organization Data & Society, algorithmic management is a diverse set of technological tools and techniques that structure working conditions and remotely manage the workforce.

Artificial intelligence (AI): Artificial intelligence refers to machines (e.g. computer systems) that are able to mimic human abilities to perceive their environment, learn, reason and act (perform tasks). In the broadest sense, AI refers to machines that can learn, reason and act by themselves. They can make their own decisions when faced with new situations, in the same way that humans and animals can.

The vast majority of systems that are called artificial intelligence today are actually *machine learning systems* (see below). In addition to machine learning and *automated decision-making systems*, there are some other types of artificial intelligence that have specific applications in the workplace:

Computer Vision (CV): analysis of visual information (static images or video streams) to recognize and classify images, objects, activities or events, individual faces, intentions. Some CV systems are designed to monitor human-to-human interaction.

Natural Language Processing (NLP): is the analysis of written and spoken language to recognize and classify words and to understand and generate written and spoken language. Other types of NLP applications are machine translation, chatbots, social media analysis, voice assistants, text summarization, information retrieval and emotional analysis.

Speech recognition: Analyzing audio (e.g. phone calls, conversations, voice commands) to recognize and process spoken language into text. Speech recognition can also be used to process text into spoken language.

> **Robotics:** Hardware systems that can perform physical tasks such as movement and interact with and adapt to changes in the physical world. Robots run on software systems that have varying degrees of complexity; the most advanced robots rely on learning algorithms, computer vision and natural language processing.

For more information see this short video: Artificial Intelligence Explained in 2 min and this article: What is AI?)

Automated decision-making systems (ADS): Semi-automated decision-making systems can be used to support human decision-making by providing recommendations to humans, while fully-automated ADS systems execute the decision without human involvement. Examples of ADS in use are fraud detection, social welfare eligibility determination, scheduling optimization, driving route optimization, planning and task allocation. Thus, there is often an overlap between ADS and "*algorithmic management*."

Big data: Extremely large and complex data sets that are analyzed using very high computing power and speed. Data is continuously collected from a variety of sources, including business transactions, *IoT devices, sensors*, RFID tags, industrial equipment, videos, social media, etc. Big data is used by artificial intelligence or machine learning to reveal patterns, trends, and associations, especially in the context of human behavior and interactions.

Data: Data can be seen as the smallest unit of information that can be used as a basis for calculation, reasoning or discussion. Data must be processed to be meaningful. See more under *Data analytics*, *Employee data* and *Personal data* respectively.

Data analytics: Data analytics is a broad term that encompasses many different types of analysis designed to extract insights, identify trends, optimize processes or solve problems. Two common types of advanced data analysis techniques used in the workplace are predictive and prescriptive analytics. **Predictive analytics** uses techniques such as forecasting, statistical modeling or machine learning to make predictions about what outcome is likely to happen in the future. **Prescriptive analytics** uses techniques such as machine learning to recommend a course of action that will deliver the desired results. These forms of advanced analytics are increasingly embedded in automated digital systems that combine data collection and analysis to make predictions and decisions.

Controller and processor: Concepts used in data protection regulations. In short, the controller determines why (for what purpose) and how (by what means) personal data is processed. A data processor, on the other hand, is someone who processes personal data on behalf of the controller - i.e. on the instructions of the controller.

Data protection: In the US, data protection regulations vary across states. The General Data Protection Regulation (GDPR) in Europe is said to be the gold standard setting a number of requirements towards employers, including: *transparency, data minimization* and *impact assessments*. See the descriptions for these in this document.

Data minimization: The principle of "data minimization" is especially known from the European Data Protection Regulation, the GDPR. It implies that a data controller should limit the collection of personal data to what is directly relevant and necessary to achieve a specific purpose. They should also only retain the data for as long as necessary to fulfill that purpose. In other words, controllers should only collect the personal data they really need and should only keep it for as long as they need it.

This means that employers are not allowed to collect a lot of data because they might need it one day. They are also not allowed to permanently store data and personal information about their employees because it goes against the protection of individual rights.

Data Protection Impact assessments (DPIA): Required by law in 2023 in California, Colorado and Virginia. The requirements are similar in many ways to the existing requirements for DPIA's under the European Union's General Data Protection Regulation (the "GDPR"). Although the terms differ among jurisdictions, the basic concepts are substantially similar.

See the UK DPA's detailed guidance on DPIAs here. The European Commission's here

Digital systems: Digital systems include hardware and software designed to collect, store, process and communicate information (*data*) in digital (binary numbers) form. Key components of digital systems include 1) input and output devices (e.g. keyboard, camera, microphone, monitor, speakers), 2) memory and 3) central processing unit (CPU). Computers and smartphones are examples of digital systems. Algorithms are also a key component of digital computing systems. Digital systems can be connected to form a network (see also *Internet of Things*).

Electronic monitoring: Electronic monitoring is a particularly invasive form of data collection that involves systematic and continuous monitoring and recording of employee behavior and actions. Although not new, electronic monitoring has become more common with the development of internet-connected devices with built-in sensors that can capture a wide range of data about employees' location, activities and interactions with coworkers (see *Internet of Things* definition). Electronic monitoring is often embedded in the measurement of the work process rather than specifically focusing on tracking the employee.

Employee data - collecting data from workers: Employers can collect a wide range of data about employees. Some of this data is collected in the workplace, such as computer activity, location in the building, customer reviews, bathroom usage, coworker interactions and smartphone app interactions. Other types of data are purchased from third parties, such as social media activity, credit reports, driving history and consumer activity. Some of this data, such as criminal background checks, has been collected by employers abroad for decades. In some countries, employers have partnered with wellness programs and technology providers to collect biometric and health and wellness data as new wearable *sensors* have become available. Data collected from employees is *personal data* (see below)

Generative AI: Typically takes on 2 forms:

- **Large language models** (LLMs), such as the one that underpins ChatGPT, which generate plausible-sounding text in response to a human prompt in the form of a request (e.g. 'write a sonnet about the risks of AI in the style of Shakespeare')

- **Multi-modal models,** such as Stable Diffusion, Midjourney, or OpenAI's DALL-E 2, typically take text prompts (e.g. 'a purple penguin wearing sunglasses') and generate images as an output. Some models, such as GPT-4, can also take images as input (e.g. a photo of your fridge's contents) to produce text as the output (e.g. a recipe for the ingredients you have). Multi-modal models that can generate audio and video outputs are also in development.

Generative AI systems are trained on huge datasets of text, images, and other media in order to produce similar but synthetic content. These systems make predictions about the text likely to follow a given prompt: they generate content as their output - hence the term 'generative AI.' Such systems can imitate the work of writers or artists included in their training data – but they will also replicate any biases from the content they are trained on, such as racist language or sexist imagery.

Read more about Generative AI by Access Now here

Human oversight/human in command: These two concepts are used in the context of artificial intelligence systems that are designed to support human autonomy and decision-making, as prescribed by the principle of respect for human autonomy. This requires that AI systems must both enable a democratic, thriving and just society by supporting user agency and promoting fundamental rights, and allow for human oversight (stewardship of the systems). Read more in the document The Ethics Guidelines for Trustworthy Artificial Intelligence (AI).

Internet of Things (IoT): IoT refers to a system of devices ("things") connected to the internet to transmit and receive data, such as physical objects, industrial machines, WiFi-connected cameras, workplace and handheld devices, wearables (e.g. wristbands), smartwatches and fitness trackers, etc. IoT devices use embedded sensors (see sensor definition) to collect data and then share the data through a wireless network to other internet-connected devices (e.g. smartphones) for remote monitoring and interaction (control) or computers for processing, storage and in some cases real-time analysis and use. (For more information, see this video: What is the internet of things).

Machine learning, deep learning and neural networks: These are subcategories of AI and cover the more advanced algorithms and algorithmic systems. A common term is that they are "learning algorithms". They enable computers to perform a specific task without a human explicitly writing the rules (instructions) for how the computer should perform the task. What happens is that the algorithm is given data, a goal and feedback when it's on the right path. It then learns on its own how to continue. We humans can control what data goes in and we can see the result that comes out, but these forms of artificial intelligence are often so complicated that we can't always understand how the algorithm arrived at a certain result (for more information, see article: What is machine learning?).

Personal data: GDPR defines personal data as "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 ID 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. Online identifiers such as IP addresses are now considered personal data unless they are anonymized. Pseudonymized personal data is also subject to GDPR if it is possible to identify whose data it is by reverse engineering.

"Privacy by design and default": Privacy-by-design and/or default is about how data protection is built into a company's products, services and business processes from the outset.

- Privacy-by-design is an approach that ensures that a company incorporates data protection as an integral part of its business processes, value chain and product lifecycle. From the production phase to the product reaching the end user.
- Privacy-by-default means that products are set up from the start to ensure the highest level of personal data protection.

Profiling: Profiling, in short, is the classification of a person's personality, behavior, interests and habits. It's important to remember that profiling is also done on everything we DON'T do, have interests or habits. Profiling is typically used to make predictions and is based on analysis of collected data (for more information, see this webpage from the UK Data Protection Board). Profiling can have *immediate* effects – someone is hired, promoted, disciplined or fire. And importantly, profiling can have *future* effects as profiles created can open or close opportunities for future workers.

Sensors: Sensors detect, measure and transmit information about the environmental context around the sensor and/or physical and behavioral characteristics of a human wearing the sensor. They can capture precise measurements of the physical environment and can distinguish human characteristics, activities and interactions with machines and devices. Sensors can be embedded in a variety of objects (see *Internet of Things* definition), wearables, personal devices (e.g. smartphones), etc.

Transparency: Transparency requirements in the private sector vary by state: **In the CCPA-CPRA** transparency requirements cover 1. What types of information are collected; 2. For what purpose they are being collected; 3. Specifics of what is being collected; 4. Disclosure of where data is being shared. In Virginia, the **VCDPA requirements include:** 1. Stating what categories of personal data are collected; 2. Obtaining affirmative consent for sensitive data before collecting it; 3. Providing an option for access and correct personal information; 4. Providing opt-out mechanisms; 5. Providing data protection assessments; 6. Honor deletion requests; 7. Provide data breach notifications.