Dr. Jochen Friedrich – IBM Technical Relations Europe
jochen@de.ibm.com

Artificial Intelligence

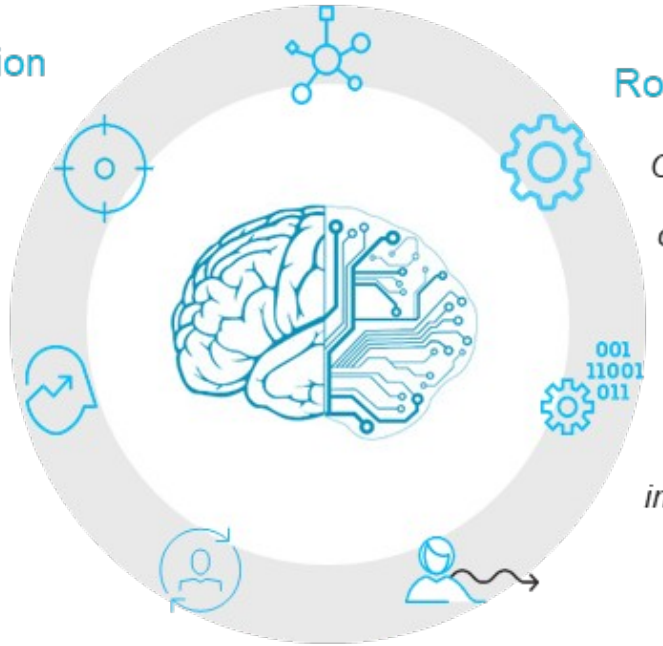
Responsible computing
Technology topics
Open technologies

18 November 2020

Cognitive Technologies

Artificial intelligence

Simulation of human intelligence processes



Recommendation engines

Analyze data and suggest something as per user's interest

Robotics

Conception, design, manufacture, and operation of robots

Predictive analytics

Predicting outcomes using statistical algorithms and machine learning

Machine learning systems

Ability to learn and improve without explicit instructions

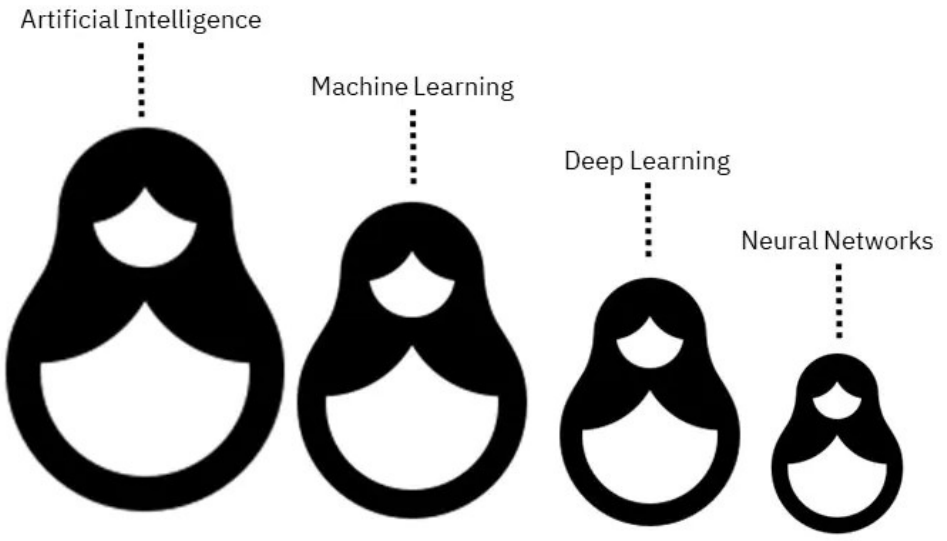
Deep learning

Machine learning with artificial neural network algorithms

Natural language processing

Ability to understand human speech as it is spoken

Some Terminology and Concepts



Compare this to Russian nesting dolls: Machine learning is a subfield of artificial intelligence. Deep learning is a subfield of machine learning, and neural networks make up the backbone of deep learning algorithms. In fact, it is the number of node layers, or depth, of neural networks that distinguishes a single neural network from a deep learning algorithm, which must have more than three.

AI mimics human intelligence. It is used to predict, automate, and optimize tasks of humans, e.g. speech and facial recognition, decision making, translation.

Deep learning is a subset of machine learning. The way in which they differ is in how each algorithm learns. A machine learning model can cluster and classify inputs.

The “deep” is referring to the depth of layers in a neural network. More than three layers can be considered a deep learning algorithm.

Neural networks mimic the human brain through a set of algorithms (4 components: inputs, weights, bias/threshold, output).

Source: https://www.ibm.com/cloud/blog/ai-vs-machine-learning-vs-deep-learning-vs-neural-networks#:~:text=Each%20is%20essentially%20a%20component,backbone%20of%20deep%20learning%20algorithms.

Good AI

Bad AI

Some Trust and Transparency Principles on AI

The purpose of AI is to augment human intelligence

- Augment – not replace – human intelligence
- Enhance and extend human capability and potential

Data and insights belong to their creator

- Clients' data is their data, and their insights are their insights

New technology, including AI systems, must be transparent and explainable

- AI must be transparent. It must be clear about who trains AI systems, what data was used in that training
- Data governance policies to ensure people understand how an AI system came to a conclusion or recommendation
- Address the issue of data bias pro-actively and try to minimise it and continuously look for better data sets

Imperatives for Artificial Intelligence

Five imperatives for companies, based on whether they are a provider or owner (or both) of an AI system

Designate a lead AI ethics official

Responsible for trustworthy AI

Accountable for internal guidance and compliance mechanisms

Oversee risk assessments and harm mitigation strategies

Different rules for different risks

Initial high-level assessment of the technology's potential for harm

In-depth and detailed assessment for higher-risk applications

Document assessment processes in detail to be auditable, and retain them for a minimum period of time

Don't hide your AI

Promote transparency is through disclosure

Make purpose of an AI system clear to consumers and businesses

Explain your AI

Maintain audit trails surrounding input and training data

Make available documentation informing about confidence measures, levels of procedural regularity, error analysis

Test your AI for bias.

Responsibility that AI systems are fair and secure

Take remedial actions as needed, both before sale or deployment and after an AI system is operationalised

EU High Level Expert Group on AI

Ethics Guidelines on Artificial Intelligence:

The Guidelines put forward a human-centric approach on AI and list 7 key requirements that AI systems should meet in order to be trustworthy.

- Human agency and oversight
- Technical robustness and safety
- Privacy and Data governance
- Transparency
- Diversity, non-discrimination and fairness
- Societal and environmental well-being
- Accountability

Standards versus Open Source

Open Technologies

Standards

A building plan (methods, metrics, processes, protocols, …)

Developed collaboratively in standards developing organisations in open and transparent processes

Not controlled by one single vendor or group of vendors

Available for free, for small administrative fee, or for sale

May include patented technologies

Open Source

Source code (Software)

Developed collaboratively in open source projects / communities / foundations

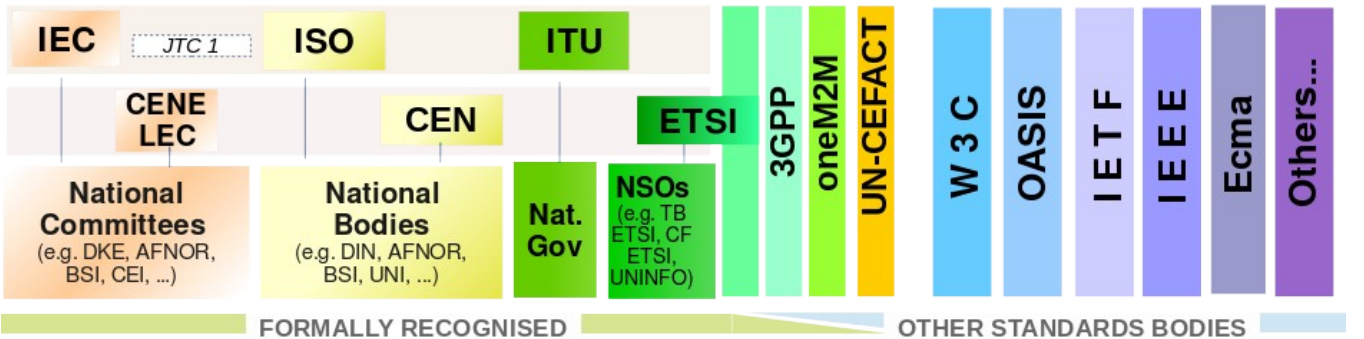
Open source governance for decision making

Openly available

Licensed under open source license (typically OSI approved license)

Standardisation and Open Source Ecosystem

standardisation



open source

- Linux Foundation
- Eclipse Foundation
- Linux Foundation AI
- Iot eclipse
- node
- kubernetes
- Many others

Major Standardisation Committees Active in AI

International / global	ISO/IEC JTC 1 SC 42	IEEE	(ITU)
Regional	CEN	CENELEC	ETSI
National	DIN	BSI	NEN
	afnor	UNINFO	NIST
			others

Standards Under Development … Some Concrete Examples

ISO/IEC JTC 1 SC42	Concepts and Terminology; Definition; ISO/IEC 42001 – Management Systems Standard on AI (Process standard) ISO/IEC 24368 – Artificial Intelligence (AI) - Overview of ethical and societal concerns (TR) ISO/IEC 24027 – Bias in AI systems and AI aided decision making (TR) ISO/IEC 24028 – Overview of trustworthiness in AI (TR) ISO/IEC 38507 – Governance implication of the use of AI by organizations
IEC SEG 10	Ethics in Autonomous and Artificial Intelligence Applications
IEEE	Global Initiative on Ethics of Autonomous and Intelligent Systems (P7000) Governance of AI systems Several activities addressing Machine Learning
CEN/CENELEC	Roadmap on European AI Standardisation Under preparation: Setting up CEN/CENELEC JTC on AI

In a Nutshell …

AI is a broad field and differentiation on the exact technology and its use is important.

Imperatives and guidelines are available regarding a transparent and responsible use of AI.

Open Technologies – open standards and open source – play a major role in addressing societal values and ethical topics around AI.

Very concrete and relevant standards are already under way and consider European requirements, e.g. those coming from the HLEG.

Thanks very much for your attention

Dr. Jochen Friedrich
jochen@de.ibm.com
https://www.linkedin.com/in/jochenfriedrich/
https://twitter.com/jfopen