Human-centric AI challenges and opportunities

WIRTSCHAFTS UNIVERSITÄT WIEN VIENNA UNIVERSITY OF ECONOMICS AND BUSINESS

Sabrina Kirrane, 27.01.2020

AMLD – AI & Policy track

Setting the Scene
About me

Access Policies

Normative Multi-Agent Systems

Consent Policies

Intelligent Agents

Licenses

Cyber Physical Social Systems

Regulatory Constraints

Encryption

Blockchain

Privacy

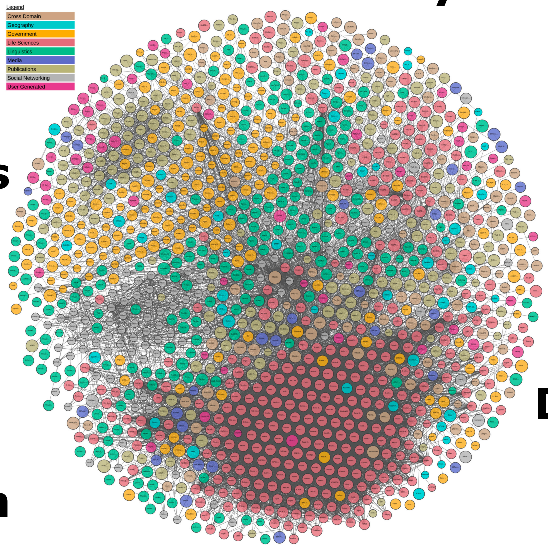
Enforcement

Decentralisation

Data Science

Administration

Artificial Intelligence



Linked Open Data Cloud https://lod-cloud.net/

Transparency

Compliance

Big Data

Web Standards

The World Wide Web

WIRTSCHAFTS UNIVERSITÄT WIEN VIENNA UNIVERSITY OF ECONOMICS AND BUSINESS

Information Management: A Proposal

Tim Berners-Lee, CERN

March 1989, May 1990

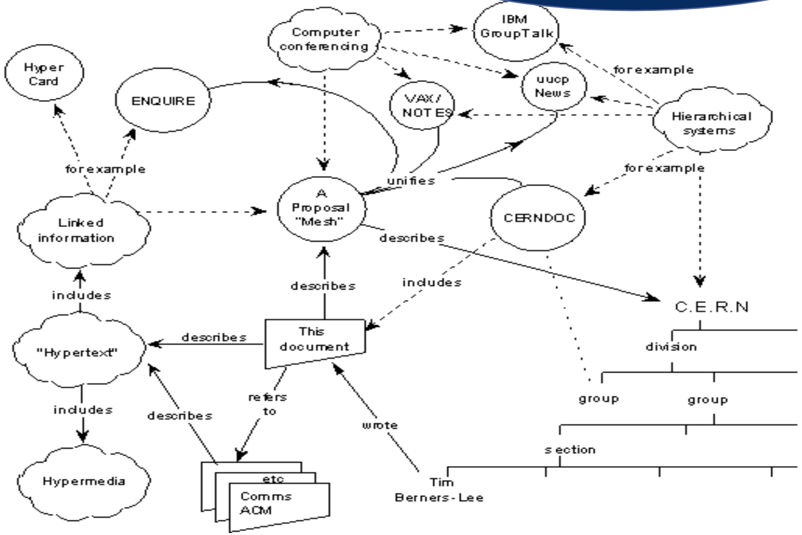
This proposal concerns the management of general information about accelerators and experiments at CERN. It discusses the problems of loss of information about complex evolving systems and derives a solution based on a distributed hypertext system.

Overview

Many of the discussions of the future at CERN and the LHC era end with the question - "Yes, but how will we ever keep track of such a large project?" This proposal provides an answer to such questions. Firstly, it discusses the problem of information access at CERN. Then, it introduces the idea of linked information systems, and compares them with less flexible ways of finding information.

It then summarises my short experience with non-linear text systems known as "hypertext", describes what CERN needs from such a system, and what industry may provide. Finally, it suggests steps we should take to involve ourselves with hypertext now, so that individually and collectively we may understand what we are creating.

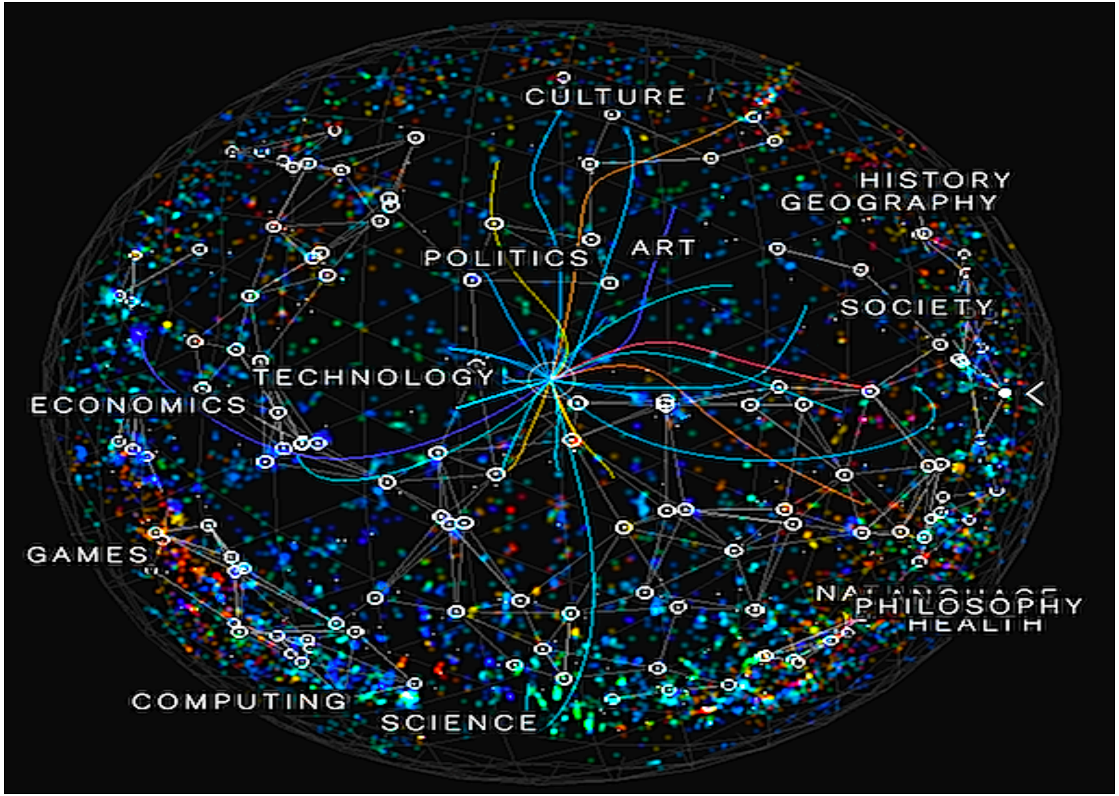
In 1989
Tim Berners Lee
invented the World
Wide Web



1989 The original proposal for the Web
https://www.w3.org/History/1989/proposal.html

EFMD EQUIS ACCREDITED · AACSB ACCREDITED · ASSOCIATION AMBA ACCREDITED

The World Wide Web
As a disturbed data source

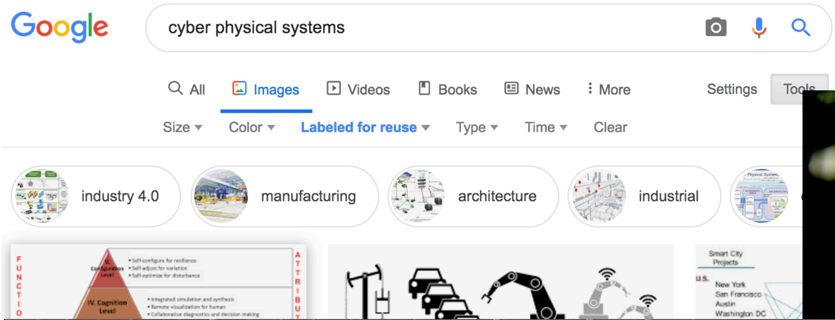


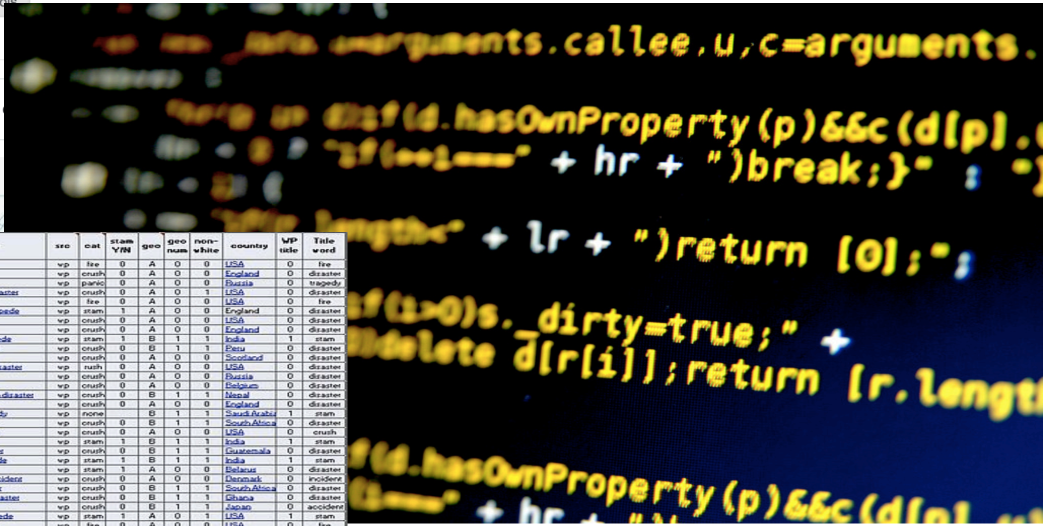
30 years later the Web has become indispensible!

https://commons.wikimedia.org/wiki/File:WikiVerse_at_WM_NYC_WikiWednesday.png

Contracts & Terms of Use
The compliance challenge

- There are many resources without any terms of us
- We need compliance tools



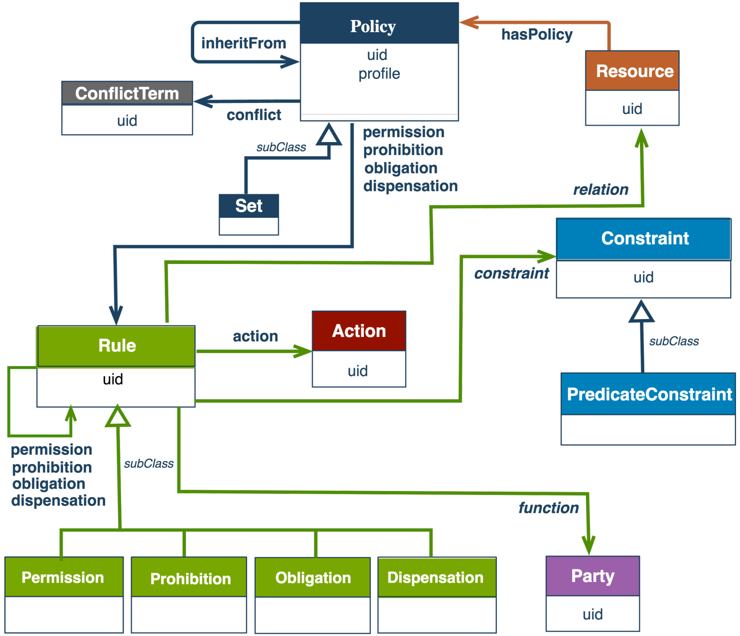


uiwen/3260095534

event	year	dead	injured	crowd size	event type	event type num	mod type num	relig.	indoor	counterflow	bridge	event population	population killed	specific trigger	crowd mgmt failure	escape panic	Event name	src	cat	stam Y/N	geo	geo num	non-white	country	WP title	Title word
5	1876	278		1,000	F	1	1	0	Y	N	N	A	A	Y		Y	Brooklyn Theatre fire	wp	fire	0	A	0	0	USA	0	fire
7	1883	183		1,100	T	8	8	0	Y	N	N	C	C	Y		N	Victoria Hall disaster	wp	crush	0	A	0	0	England	0	disaster
8	1896	1389	1300	500,000	P	3	3	0	N	N	N	G	G	Y		N	Khodynka Tragedy	wp	panic	0	A	0	0	Russia	0	tragedy
9	1902	115		6,000	R	5	5	1	Y	N	N	A	A	Y		Y	Shiloh Baptist Church disaster	wp	crush	0	A	0	1	USA	0	disaster
10	1903	602		2,150	F	1	1	0	Y	N	N	G	G	Y		Y	Iroquois Theatre Fire	wp	fire	0	A	0	0	USA	0	fire
11	1908	16	40		T	8	8	0	Y	N	N	C	C	Y		N	Barnsley Public Hall stampede	wp	stam	1	A	0	0	England	0	disaster
12	1913	73		400	T	8	8	0	Y	N	N	G	C	Y		Y	Italian Hall Disaster	wp	crush	0	A	0	0	USA	0	disaster
14	1946	33	400	85,000	S	7	7	0	N	N	N	G	G	N	Y	N	Burnden Park disaster	wp	crush	0	A	0	0	England	0	disaster
15	1954	800	2000	5,000,000	R	5	5	1	N	N	N	G	G	N	Y	N	1954 Kumbh Mela stampede	wp	stam	1	B	1	1	India	1	stam
17	1964	328		53,000	S	7	7	0	N	N	N	G	G	Y	Y	Y	Estadio Nacional disaster	wp	crush	0	B	1	1	Peru	0	disaster
18	1971	66	200	80,000	S	7	7	0	N	N	N	G	G	N	Y	N	1971 Ibrox disaster	wp	crush	0	A	0	0	Scotland	0	disaster
19	1979	11	26	14,770	M	2	2	0	N	N	N	G	G	Y		N	1979 The Who concert disaster	wp	rush	0	A	0	0	USA	0	disaster
20	1982	66	61	16,643	S	7	7	0	N	N	N	G	C	N	Y	N	Luzhniki disaster	wp	crush	0	A	0	0	Russia	0	disaster
21	1985	39	600	59,000	S	7	7	0	N	N	N	G	G	Y	Y	Y	Heysel Stadium disaster	wp	crush	0	A	0	0	Belgium	0	disaster
22	1988	93	100		S	7	7	0	N	Y	N	G	G	N	Y	Y	1988 Kathmandu stadium disaster	wp	crush	0	B	1	1	Nepal	0	disaster
24	1989	96	766	55,000	S	7	7	0	N	Y	N	G	G	N	Y	N	Hillsborough disaster	wp	crush	0	A	0	0	England	0	disaster
25	1990	1426			R	5	5	1	N	N	N	A	A	N	Y		1990 Mecca tunnel tragedy	wp	none		B	1	1	Saudi Arabia	1	stam
26	1991	40		30,000	S	7	7	0	N	N	N			Y		N	Orkney Stadium Disaster	wp	crush	0	B	1	1	South Africa	0	disaster
32	1993	0	73	77,745	S	7	7	0	N	N	N	G		Y	Y	N	The Camp Randall Crush	wp	crush	0	A	0	0	USA	0	crush
34	1994	113	500	50,000	PO	4	4	0	N	N	N	G	WC	Y	Y	N	1994 Gowari stampede	wp	stam	1	B	1	1	India	1	stam
35	1996	83	147	50,000	S	7	7	0	N	N	N	G	G	N	Y	N	The October 16 disaster	wp	crush	0	B	1	1	Guatemala	0	disaster
37	1999	53			R	5	5	1	N					Y			1999 Sabarimala stampede	wp	stam	1	B	1	1	India	1	stam
38	1999	53	100		M	2	2	0	N					Y		Y	Nyamiha stampede	wp	stam	1	A	0	0	Belarus	0	disaster
39	2000	9			M	2	2	0	N					N		N	Roskilde Festival 2000 incident	wp	crush	0	A	0	0	Denmark	0	incident
41	2001	43		120,000	S	7	7	0	N		N			Y	Y	Y	Ellis Park Stadium disaster	wp	crush	0	B	1	1	South Africa	0	disaster
42	2001	127			S	7	7	0	N	N	N			Y	Y	Y	Accra Sports Stadium disaster	wp	crush	0	B	1	1	Ghana	0	disaster
43	2001	11	247		P	3	3	0	N	N	Y	G	C	N	Y	N	Akashi crowd crush	wp	crush	0	B	1	1	Japan	0	accident
46	2003	21	50	1,500	M	2	2	0	Y	Y	N	A	A	Y	Y	Y	2003 E2 nightclub stampede	wp	stam	1	A	0	0	USA	1	stam
47	2003	100	230	462	F	1	1	0	Y	N	N	A	A	Y	Y	Y	The Station nightclub fire	wp	fire	0	A	0	0	USA	0	fire
52	2005	291		300,000	R	5	5	1	N	N	Y	G	WC	N	Y		Mandher Devi temple stampede	wp	stam	1	B	1	1	India	1	stam
53	2005	953		1,000,000	R	5	5	1	N	N	Y	G	G	Y	Y	N	2005 Al-Aaimmah bridge stampede	wp	stam	1	B	1	1	Iraq	1	stam
55	2006	345	289	2,130,594	R	5	5	1	N	N	N	A	A	N		N	Hajj 2006	wp	ak		B	1	1	Saudi Arabia	1	stam
56	2006	73	400	30,000	S	7	7	0	N	N	N			Y	Y	N	PhilSports Stadium stampede	wp	stam	1	B	1	1	Philippines	1	stam
64	2008	162	47	3,000	R	5	5	1	N		N	G	G	Y	Y	Y	Naina Devi stampede	wp	stam	1	B	1	1	India	1	stam
66	2008	224	425	25,000	R	5	5	1	N		N	M	A	Y	Y	Y	2008 Jodhpur stampede	wp	stam	1	B	1	1	India	1	stam
69	2009	19	135	36,000	S	7	7	0	N	N	N			Y	Y	Y	Houphouet-Boigny Arena stampede	[21]	crush	0	B	1	1	Côte d'Ivoire	1	stam
73	2010	21	500	250,000	M	2	2	0	N	Y	N	G	G	N	Y	N	Love Parade disaster	wp	crush	0	A	0	0	Germany	0	disaster
74	2010	347	755	4,000,000	P	3	3	1	N		Y	G	G	N	Y	N	Phnom Penh stampede	wp	stam	1	B	1	1	Cambodia	1	stam
75	2011	102	100		R	5	5	1	N	N	N				Y	N	2011 Sabarimala stampede	wp	stam	1	B	1	1	India	1	stam
79	2013	60	200	50,000	P	3	3	0	N	Y	N	G	WC				2013 Houphouet-Boigny stampede	wp	stam	1	B	1	1	Côte d'Ivoire	1	stam
81	2013	242	168	1,500	F	1	1	0	Y	N	N	A	A	Y	Y	Y	Kiss nightclub fire	wp	fire	0	B	1	1	Brazil	0	fire
82	2013	36	39		R	5	5	1	Y	N	Y	G	G		Y		2013 Kumbh Mela stampede	wp[25]	stam	1	B	1	1	India	1	stam
83	2013	115	110	500,000	R	5	5	1	N	N	Y	G	WC		Y		2013 Madhya Pradesh stampede	wp[26]	stam	1	B	1	1	India	1	stam
84	2014	32	26		R	5	5	1	N	N	N	G	WC		N		2014 Patna stampede	wp[27]	stam	1	B	1	1	India	1	stam
85	2014	7	40		PO	4	4	0	N	N	N		A	N	Y	N	2014 Multan stampede	wp	stam	1	B	1	1	Pakistan	1	stam
86	2014	11	40	30,000	R	5	5	1	N	N	N	G	G	N	Y	N	Kwekwe stadium stampede	wp	stam	1	B	1	1	Zimbabwe	1	stam
87	2014	36	49	300,000	P	3	3	0	N	Y	N	G	G	N	Y	N	2014 Shanghai stampede	wp[28]	crush	0	B	1	1	China	1	stam
88	2015	28		5,000	S	7	7	0	N	N	N	G	G	Y	Y	N	30 June Stadium Stampede	wp[29]	stam	1	B	1	1	Egypt	1	stam
89	2015	18	78		P	3	3	0	N	N	N			Y		Y	2015 Haiti Carnival stampede	wp	stam	1	B	1	1	Haiti	1	stam
92	2015	1554	934	2,000,000	R	5	5	1	N	Y	N	A	A	N	Y	N	2015 Mina stampede	en	stam	1	B	1	1	Saudi Arabia	1	stam

https://en.wikipedia.org/wiki/File:Data_for_2015_stampede_analysis.jpg

Policies for constraint representation
The interoperability challenge



- Modeling regulatory obligations using an adaption of the Open Digital Rights Language
- Automated compliance checking for business policies

Draft Specification

ODRL Regulatory Compliance Profile

version 0.1

Unofficial Draft 29 May 2019

Editor:

Sabrina Kirrane (Vienna University of Economics and Business)

Authors:

Sabrina Kirrane (Vienna University of Economics and Business)
Marina De Vos (University of Bath)
Julian Padget (University of Bath)

ODRL policy modelling and compliance checking, Marina De Vos, Sabrina Kirrane, Julian Padget and Ken Satoh, Proceedings of the ***3rd International Joint Conference on Rules and Reasoning*** *(RuleML+RR 2019)*

A third 'AI Winter'
The explainabiliy challenge

Explainable AI should help us avoid a third 'AI winter'

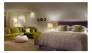
AI researchers are worried that GDPR will limit availability of training data, but there's an upside too, says Gary Richardson

People need to trust companies to do right by their data

MOST READ

5 things women in tech want to see at an event

5 reasons we still need events for women in tech



Why self-care is especially important for women in tech

Huawei CEO Ren Zhengfei admits US sanctions will cut revenues by $30bn

UK Government unveils security standard for surveillance cameras

Gary Richardson – MD of Emerging Tech at 6point6 a technology consultancy with strong expertise in digital transformation, emerging technology and cyber security

- The AI winters of the 1970s and 1990s, which saw research funding slashed and interest in AI wane, were the result of unreaslistic expectations and a failure to scale.
- A third AI winter could be caused by **inadequacies and biases** in the AI algorithms leading to negative impacts on the whole of society.
- **Bias simply does not build value in business**, particularly with regards to important decisions like access to credit and healthcare or increasing diversity through recruitment.

https://www.computing.co.uk/ctg/opinion/3073390/explainable-ai-should-help-us-avoid-a-third-ai-winter

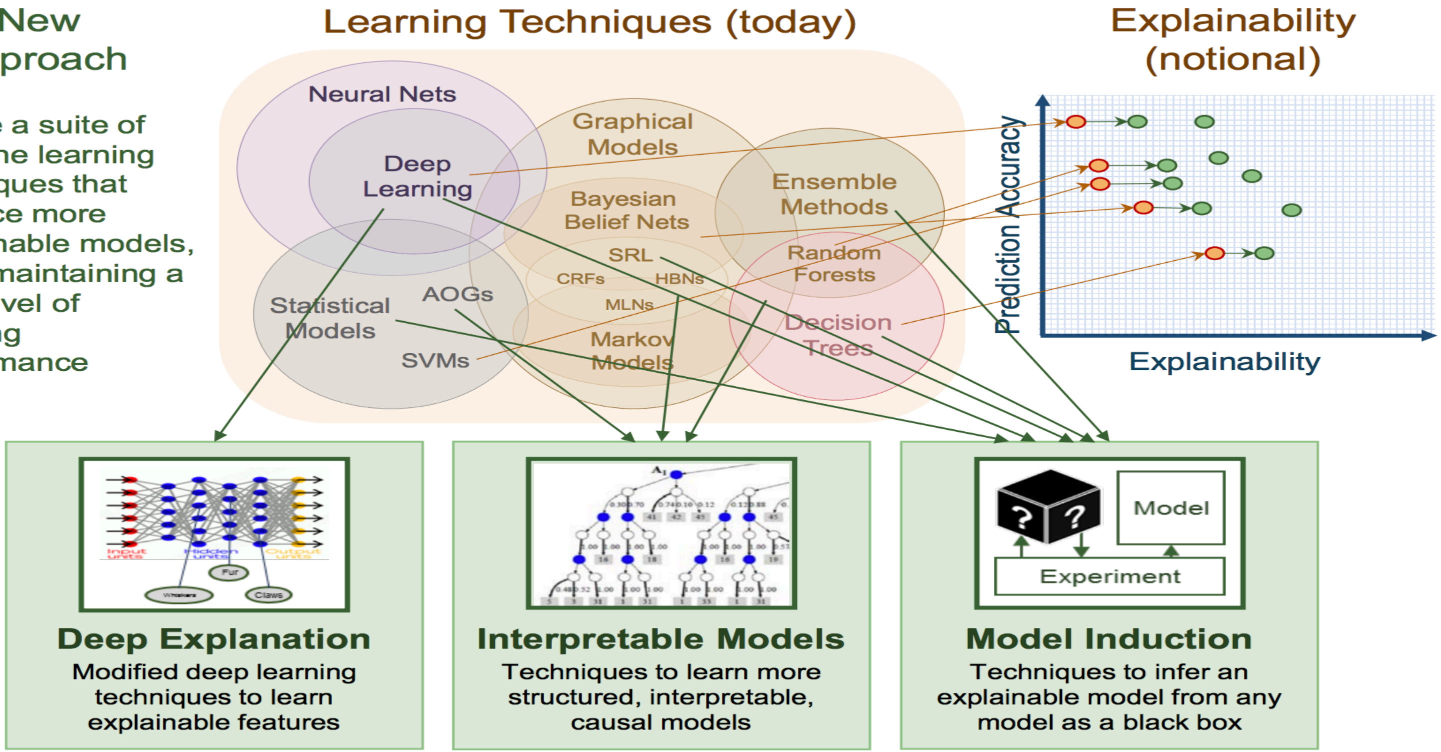
Explainable AI

The human centricity challenge

WIRTSCHAFTS UNIVERSITÄT WIEN VIENNA UNIVERSITY OF ECONOMICS AND BUSINESS

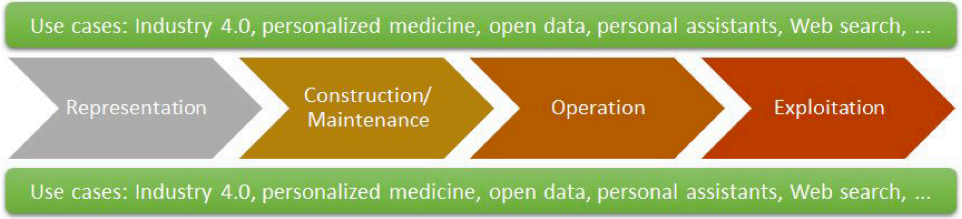
New Approach

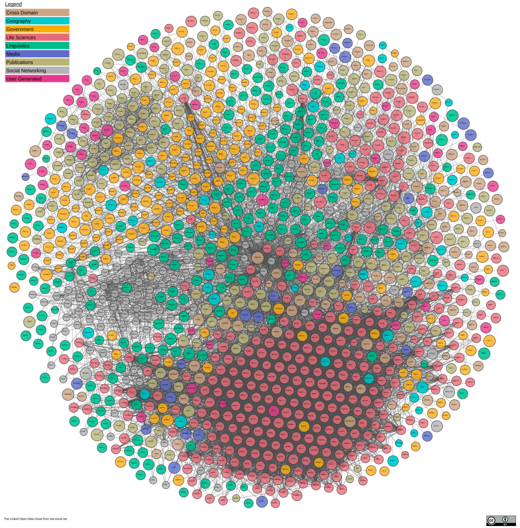
Create a suite of machine learning techniques that produce more explainable models, while maintaining a high level of learning performance



https://www.cc.gatech.edu/~alanwags/DLAI2016/(Gunning)%20IJCAI-16%20DLAI%20WS.pdf

Policies & Knowledge Graphs
Towards Responsible & Explainable AI



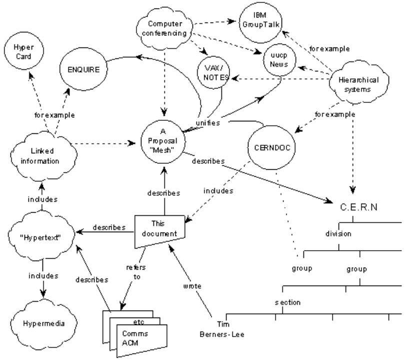


Linked Open Data Cloud https://lod-cloud.net/

- Knowledge aware machine learning
- Constraint aware reasoning and querying
- Using Knowledge graphs for explainability

This work is funded under the ***Marie Skłodowska-Curie Innovative Training Networks*** *Knowledge Graphs at Scale project:*

Policies & Knowledge Graphs
Towards Responsible & Explainable AI

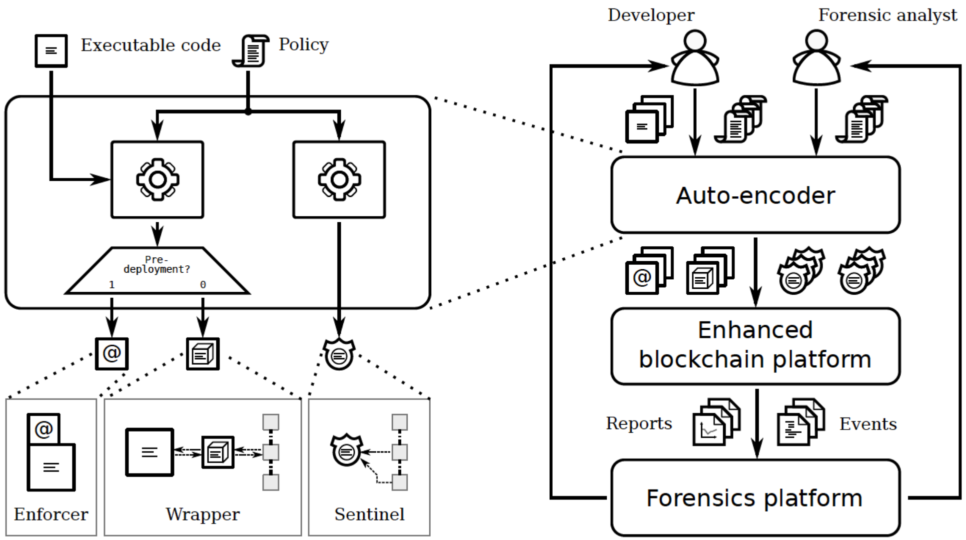


The original proposal for the Web
https://www.w3.org/History/1989/proposal.html

Apple's 1987 Knowledge Navigator
https://commons.wikimedia.org/wiki/File:Knowledge_Navigator.jpg

- Modeling goals and constraints
- Supporting negotiation and explanations
- Algorithm transparency & trust
- Usage control lower down in the technology stack

Policies & Knowledge Graphs
Towards Responsible & Explainable AI



- Constraint representation
- Syntactic and semantic function annotation
- Enforcement and conformance checking
- Forensic architecture and protocols

Collaborators:

➢Claudio Di Ciccio, Sapienza Università di Roma, Italy

➢Ruben Verborgh & Anastasia Dimou UGent-imec, Belgium

Human-centric AI
Challenges & Opportunities

WIRTSCHAFTS UNIVERSITÄT WIEN VIENNA UNIVERSITY OF ECONOMICS AND BUSINESS

- Privacy is only the tip of the iceberg, from a usage control perspective we also need to consider other **regulations, licenses, social norms, cultural differences**
- There are **cognitive limitations** in terms of understanding how data is /will be used
- There is a need for standards, however **standardisation is difficult**
- Ensuring such systems are **comply with usage constraints** is a crucial to success (i.e., all usage policies are adhered to and the system as a whole works as expected)
- We need to embrace **distributed and decentralised systems**, which complicates things further

Thank you / contact details

WIRTSCHAFTS UNIVERSITÄT WIEN VIENNA UNIVERSITY OF ECONOMICS AND BUSINESS

VIENNA UNIVERSITY OF ECONOMICS AND BUSINESS

Department of Information Systems & Operations
Institute for Information Systems & New Media
Welthandelsplatz 1, 1020 Vienna, Austria

Dr. Sabrina Kirrane

T +43-1-313 36-4494
F +43-1-313 36-90 4494
sabrina.kirrane@wu.ac.at
www.wu.ac.at
www.sabrinakirrane.com
@SabrinaKirrane

EFMD EQUIS ACCREDITED AACSB ACCREDITED ASSOCIATION AMBA ACCREDITED