July 2019

A visual approach to explainable machine learning. A design *framework*.

Thematic Scholarship *Information design, data visualisation and social media*

Beatrice Gobbo

Supervisor - Paolo Ciuccarelli
Co-Supervisor - Michele Mauri

Coordinator - Paola Bertola

POLITECNICO MILANO 1863

DOTTORATO DI RICERCA IN DESIGN

DEN-SITY GN+

First Midterm Assessment / XXXIV Cycle

* Topic Tomographies

— Collaboration between ISI and DensityDesign

#machine learning

#algorithms

#computerscience

#topicmodeling

#visual analytics

#datavisualisation

#infovisualisation

#metaphors

#interaction

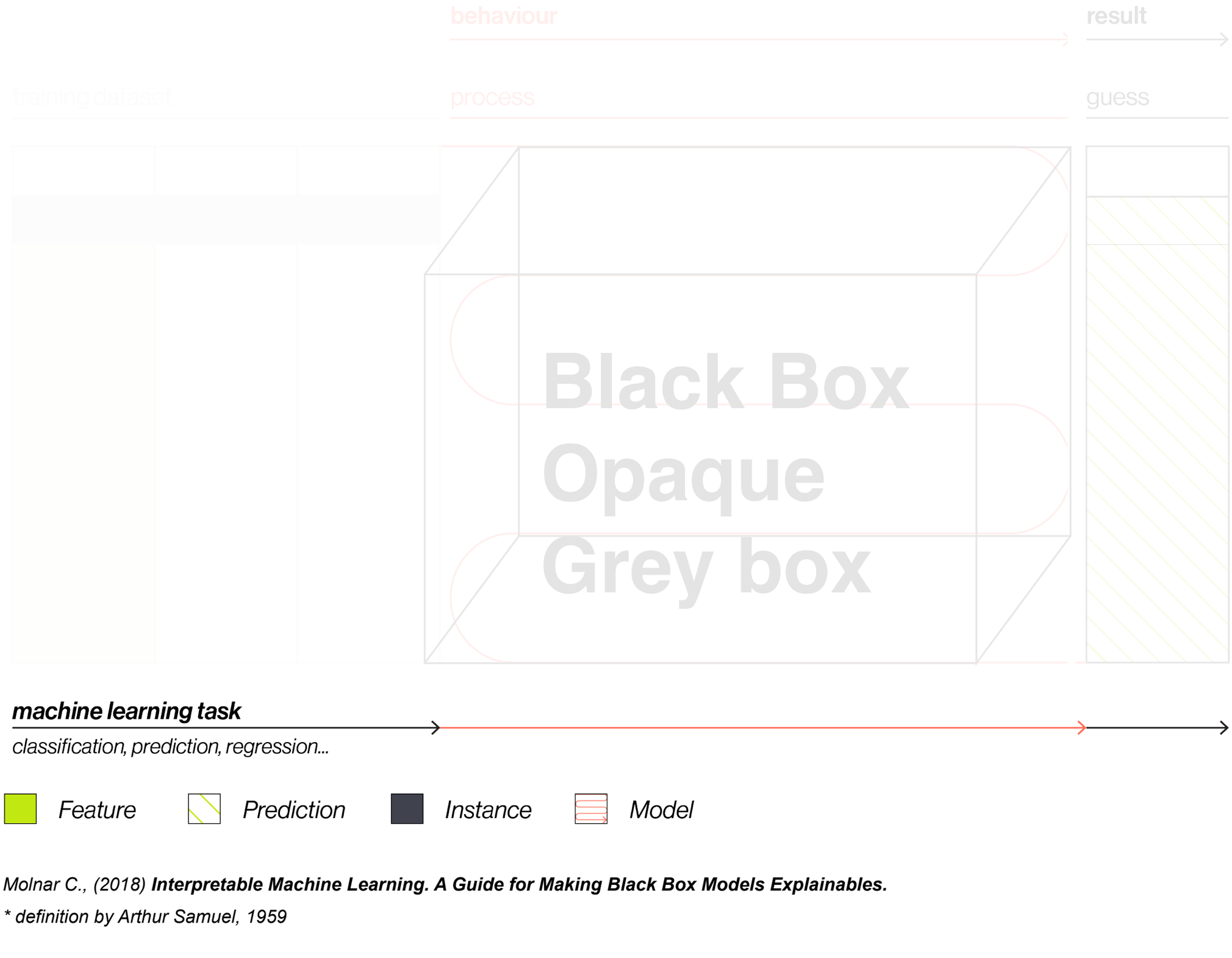
#information design

Gobbo B., Balsamo D., Mauri M., Bajardi P., Panisson A., Ciuccarelli P. (2019). **TopicTomographies: a visual approach to distil information from media streams**.

Machine Learning*

Pipeline example

Visualise it



Molnar C., (2018) ***Interpretable Machine Learning. A Guide for Making Black Box Models Explainables.***

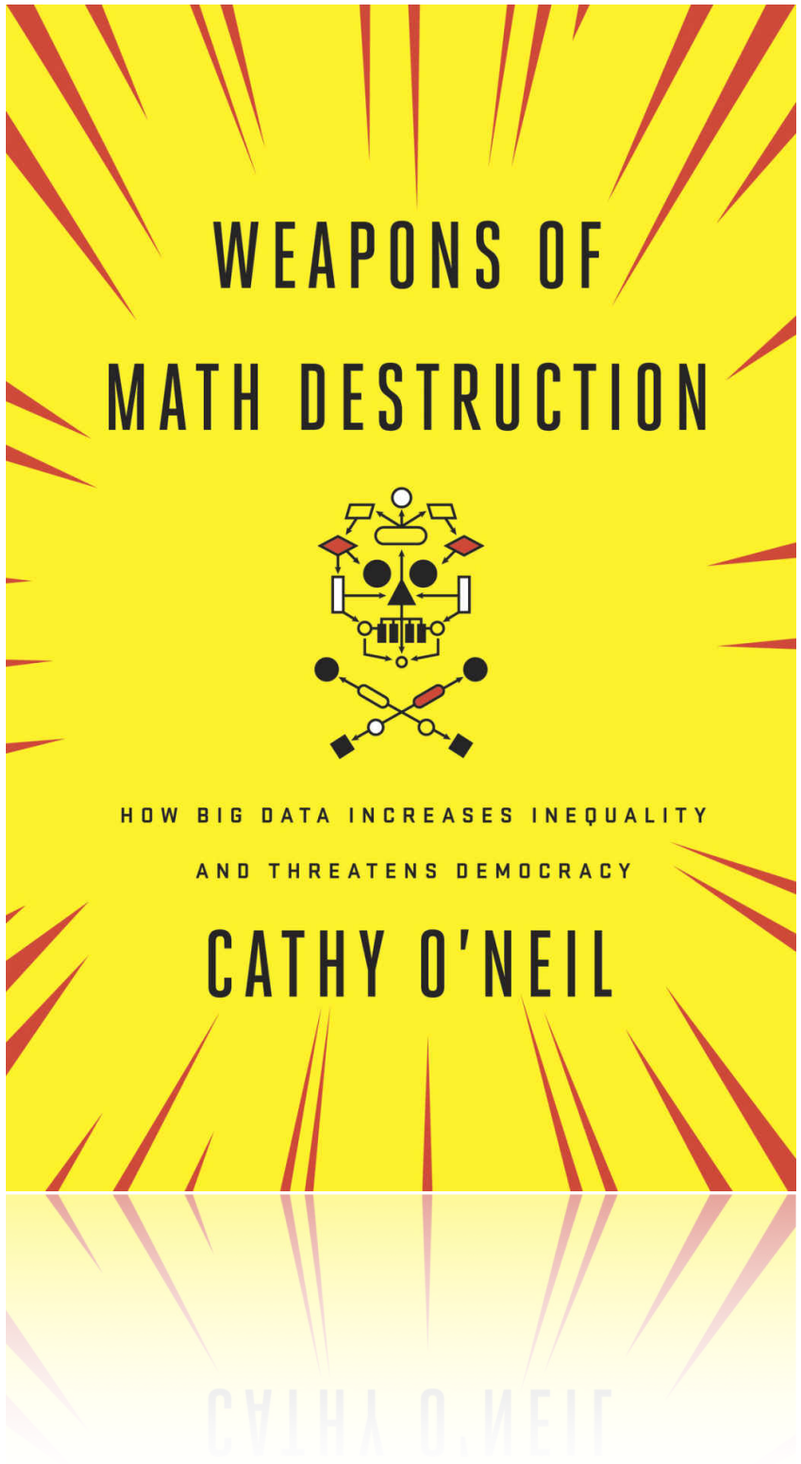
** definition by Arthur Samuel, 1959*

Buzzword or an actual turning point?

It explores how some *big data algorithms are increasingly used in ways that reinforce preexisting inequality.**

** especially machine learning algorithms*

Is a turning point because deals with training data generated by humans that carry on biases, prejudices and inequalities.



O' Neil, C. (2016)**, Weapons of Math Destruction**

What is Explainable Machine Learning

In the <u>context</u> of machine learning we define <u>interpretability</u> as the ability to <u>explain</u> or to present in understandable terms to a human.

—— *Doshi-Velez and Been Kim*

sender
ability to explain

message
the *interpretable* explanation

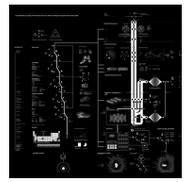
receiver
ability to interpret

Doshi-Velez, Kim B., (2017) ***Towards a Rigorous Science of Interpretable Machine Learning***
Jacobson, R. (1963) ***Saggi di linguistica generale***

Visual language of Explainable Machine Learning

Milestones: theory and practice

S



Anatomy of an AI System

Jun 2018

Share Lab

P

Anatomy of Intelligence

Oct 2018

Jhonatan Reus

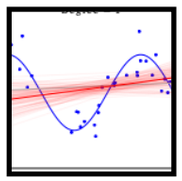
P

Serpentine Gallery

Mar 2019

Hito Steyerl

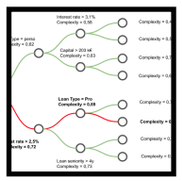
S



Random Forest I

Jul 2014

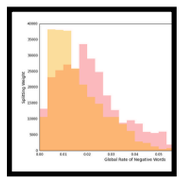
S



Random Forest II

Aug 2015

S

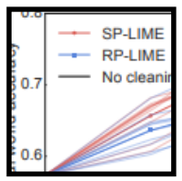


Decision Threshold

Oct 2015

AirBnb

S

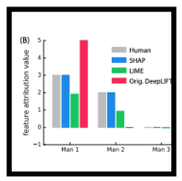


LIME

Feb 2016

Ribeiro

S



Shap

May 2017

Lundeberg

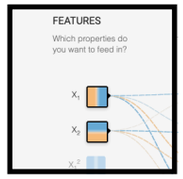
I

What if tool

July 2017

Google

I

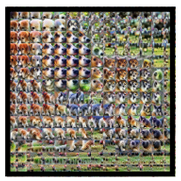


TensorFlow

Aug 2018

Google

I



The Building Blocks of Interpretability

Oct 2018

Olah

I

Manifold

May 2019

Uber

The Myth(o)s of Model Interpretability

June 2016

Lipton

Why Should I trust you

June 2017

Ribeiro

I know it when I see it"

July 2017

Offert

Towards XAI

June 2017

El-Assady

Theoretical Contributions / computer science

S *Static Visualisations*

I *Interactive Visualisations*

P *Performance*

Why explaining processes?

[…] accelerating the maturation of AI and ensuring its effectiveness has to rely on the study of relevant research domains that established well-studied processes to communicate information and knowledge.

—— Mennatallah *El Assady, 2019*

El-assady, M., Jentner, W., Kehlbeck, R., & Schlegel, U. (2019). ***Towards XAI : Structuring the Processes of Explanations Towards XAI : Structuring the Processes of Explanations****.*

A visual approach to explainable machine learning.
A design *framework*.

computer science

data science

visual analytics

semiotics

social sciences

ethics

sociology

design

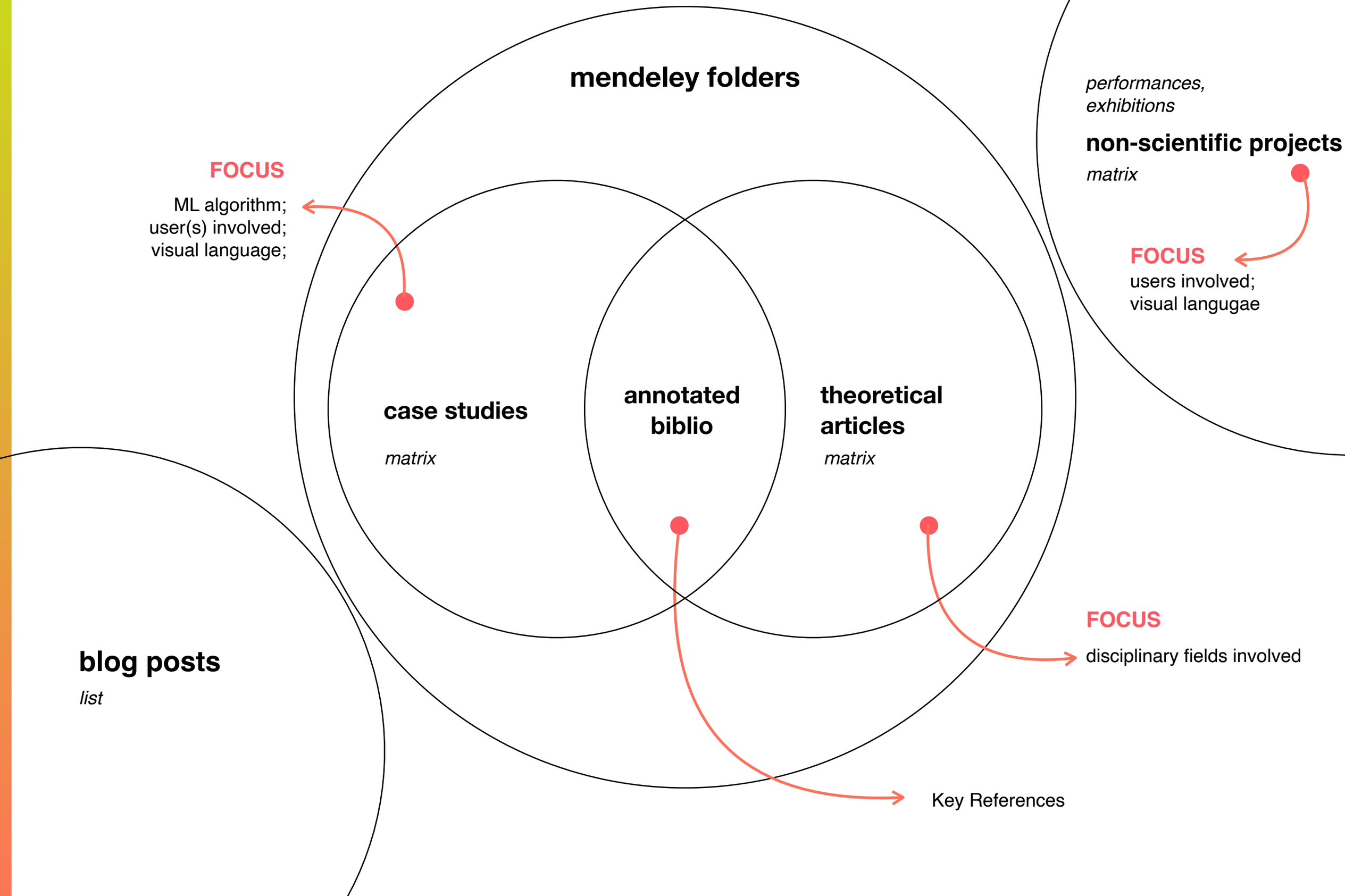
communication design

info design

semiotics

UI/UX

Literature review structure



computer science → **communication design**

visual analytics — information visualisation and design

explain behaviour

present results

classification, prediction, regression algorithms

explain to **interpret and explain**	*explain to* **compare**	*explain to* **tune**	*explain to* **teach**	*present to* **rise awareness**
Interpretable machine learning and explainable machine learning could be used both for the sake of interpretability and explainability or for comparing, teaching and tuning algorithms.	**Wang, 2019** *Manifold - UBER*	**Alsakallah, 2016** *Classification*	**Case, 2016** explorables	**Bouwlawmini, 2017** *Algorithmic JL*
Hohman, 2018 *translineation*	**Abadi, 2017** *Tensorflow*	**Abadi, 2017** *Tensorflow*	**Wattemberg, 2017** *t-sne*	**ShareLab, 2018** *Echo Anatomy translation*
visual analytics	*model builders*	*model users*	*non-experts*	**Chicau, 2018** **Information visualisation**
Yu, 2018 *model developers*	**Zhang, 2018** *Manifold*	**Olah, 2018** *CNN*	**Yosinki, 2017** *DeepViz Toolbox*	*Anatomies of intelligence*
Abdul, 2018 *Use of sub modular problems, trusting the prediction and the model*		**Khang, 2018** *model users* *GAN Lab*	*non-experts*	**El Assady, 2019** *Sheep, and Fruits: Metaphorical Narratives to Explain Artificial Intelligence and Build Trust.*
Ribeiro, 2016		**Offert, 2017** *If **intuitive** Interpretability is needed, representation of internal model states should be avoided*	**El Assady, 2018** *Towards XAI : Structuring the Processes of Explanation. **Building blocks** of explanations.*	

WHAT

WHO

HOW

Heer, 2018
Storytelling

■ *Theoretical Contributions*

computer science → **communication design**

WHAT

explain behaviour

classification, prediction, regression algorithms

explain behaviour

~~**present results**~~

WHY

explain to **interpret and explain**	*explain to* **compare**	*explain to* **tune**	*explain to* **teach**	*explain to* ~~*present to*~~ **rise awareness**

WHO

			budding users	***lay users***
				concerned users
				unaffected users
model developers	***experts***	***domain experts***	~~***non experts***~~	
			aim to C.S	*will never learn C.S.*

HOW

visual analytics					*information visualisation*
transliteration	***building blocks***	***storytelling***	***intuition***	***metaphorical narratives***	***translation***
					Information design

State of the art

Exploration space

Identification of the entry point

An increasing number of 'popular' visualisations have been developed, specifically designed to allow non-expert users to reach open-ended, reflective insights about data with personal relevance.

—— *Andrew Vande Moere and Helen Purchase*

Moere, A. Vande, & Purchase, H. (2011). ***On the role of design in information visualization****. Information Visualization 10(4) 356–371*

Identification of the entry point

[…] algorithmic miscalculations can disclose the hopes and fears, the expectations and the projections that users, institutions and companies have towards computational technologies.

—— *Public Data Lab - Not Equal Tech*

Jonathan Gray, Public Data LAB

Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). **The ethics of algorithms: Mapping the debate**. *Big Data & Society*, *3*(2)

Which is the role of information design in the field of explainable machine learning? Is Information design a good means for communicating machine learning' processes to a lay public, borrowing visual affordances from visual analytics?

Are algorithms' miscalculations the right entry point for raising awareness?

Which level of visual and conceptual complexity is needed?

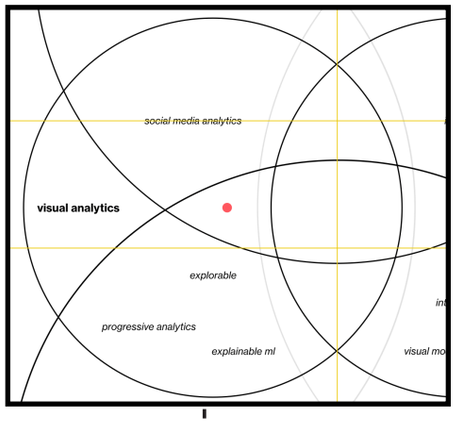
Which is the visual language of machine learning explainability for lay users?

Crossroads

Two distinct paths

HOW

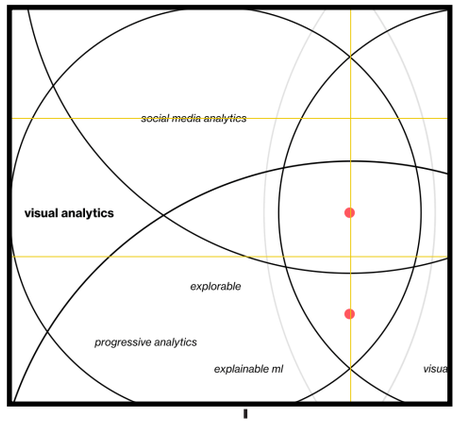
techniques



visual analytics

HOW, WHO

aim, objective, definition of the user



sociology, ethics

Visual Network Analysis of Visual Models

HOW

ures
n nature

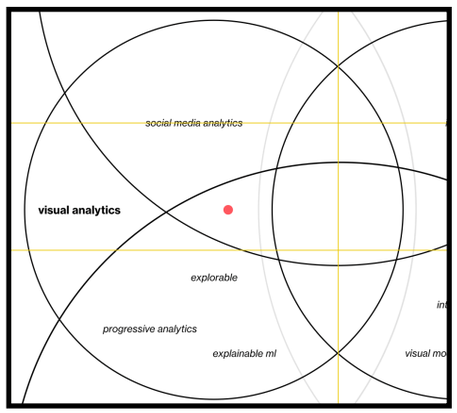
matrices

diagrams, line charts

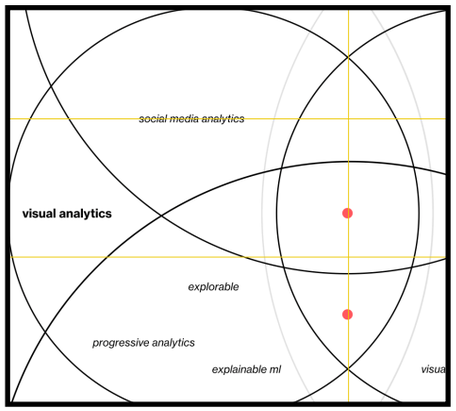
t-sne

maps

Identification of different visual languages and visual affordance.



The interface of the new AI



*** involved researchers**

Philosoper, Sociologist, Artist, Computer Scientist, Designer

Identification of emerging topics, entry points

Methodology

July 2019

Parallel, experimental and hybrid approach

Information design

Which is the role of information design in the field of explainable machine learning? Is Information design a good means for communicating algorithms' processes to a lay public, borrowing visual affordances from visual analytics?

Surveys

Visual surveys

Visual Network Analysis

In the field of interpretable machine learning which is the visual language adapted?

visual analytics

Applied Research

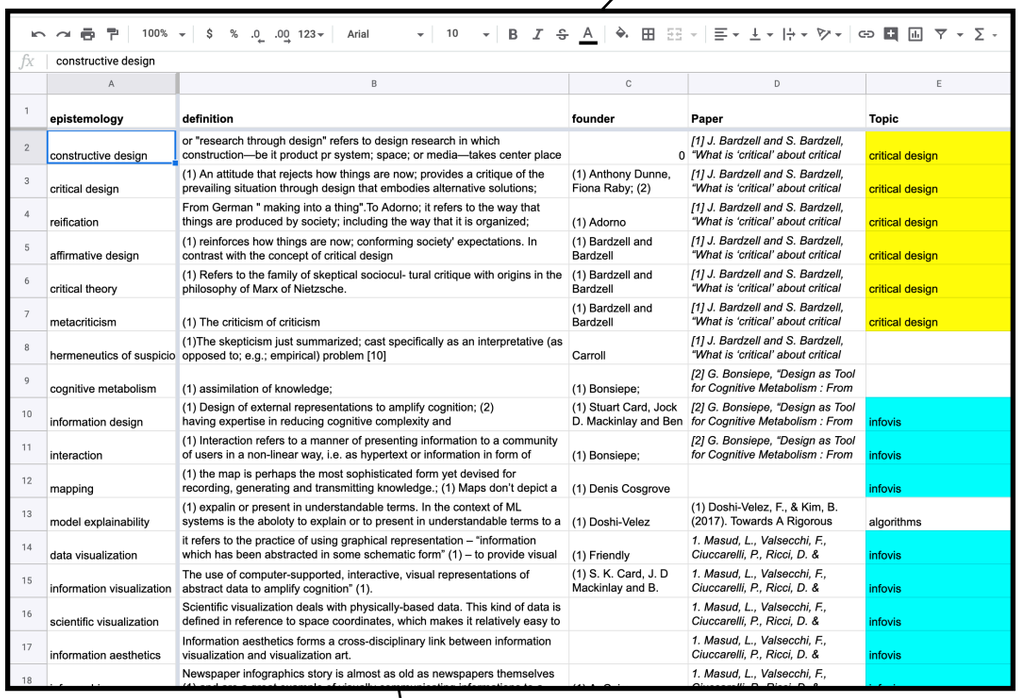
Research Through Design

Is machine learning explainability a good means for raising awareness of machine learning miscalculations among the lay public?

sociology, ethics

Research Glossary

Work in Progress



constructive design

	A	B	C	D	E
1	**epistemology**	**definition**	**founder**	**Paper**	**Topic**
2	constructive design	or "research through design" refers to design research in which construction—be it product pr system; space; or media—takes center place	0	*[1] J. Bardzell and S. Bardzell, "What is 'critical' about critical*	critical design
3	critical design	(1) An attitude that rejects how things are now; provides a critique of the prevailing situation through design that embodies alternative solutions;	(1) Anthony Dunne, Fiona Raby; (2)	*[1] J. Bardzell and S. Bardzell, "What is 'critical' about critical*	critical design
4	reification	From German " making into a thing".To Adorno; it refers to the way that things are produced by society; including the way that it is organized;	(1) Adorno	*[1] J. Bardzell and S. Bardzell, "What is 'critical' about critical*	critical design
5	affirmative design	(1) reinforces how things are now; conforming society' expectations. In contrast with the concept of critical design	(1) Bardzell and Bardzell	*[1] J. Bardzell and S. Bardzell, "What is 'critical' about critical*	critical design
6	critical theory	(1) Refers to the family of skeptical sociocul- tural critique with origins in the philosophy of Marx of Nietzsche.	(1) Bardzell and Bardzell	*[1] J. Bardzell and S. Bardzell, "What is 'critical' about critical*	critical design
7	metacriticism	(1) The criticism of criticism	(1) Bardzell and Bardzell	*[1] J. Bardzell and S. Bardzell, "What is 'critical' about critical*	critical design
8	hermeneutics of suspicio	(1)The skepticism just summarized; cast specifically as an interpretative (as opposed to; e.g.; empirical) problem [10]	Carroll	*[1] J. Bardzell and S. Bardzell, "What is 'critical' about critical*	
9	cognitive metabolism	(1) assimilation of knowledge;	(1) Bonsiepe;	*[2] G. Bonsiepe, "Design as Tool for Cognitive Metabolism : From*	
10	information design	(1) Design of external representations to amplify cognition; (2) having expertise in reducing cognitive complexity and	(1) Stuart Card, Jock D. Mackinlay and Ben	*[2] G. Bonsiepe, "Design as Tool for Cognitive Metabolism : From*	infovis
11	interaction	(1) Interaction refers to a manner of presenting information to a community of users in a non-linear way, i.e. as hypertext or information in form of	(1) Bonsiepe;	*[2] G. Bonsiepe, "Design as Tool for Cognitive Metabolism : From*	infovis
12	mapping	(1) the map is perhaps the most sophisticated form yet devised for recording, generating and transmitting knowledge.; (1) Maps don't depict a	(1) Denis Cosgrove		infovis
13	model explainability	(1) expalin or present in understandable terms. In the context of ML systems is the aboloty to explain or to present in understandable terms to a	(1) Doshi-Velez	(1) Doshi-Velez, F., & Kim, B. (2017). Towards A Rigorous	algorithms
14	data visualization	it refers to the practice of using graphical representation – "information which has been abstracted in some schematic form" (1) – to provide visual	(1) Friendly	*1. Masud, L., Valsecchi, F., Ciuccarelli, P., Ricci, D. &*	infovis
15	information visualization	The use of computer-supported, interactive, visual representations of abstract data to amplify cognition" (1).	(1) S. K. Card, J. D Mackinlay and B.	*1. Masud, L., Valsecchi, F., Ciuccarelli, P., Ricci, D. &*	infovis
16	scientific visualization	Scientific visualization deals with physically-based data. This kind of data is defined in reference to space coordinates, which makes it relatively easy to		*1. Masud, L., Valsecchi, F., Ciuccarelli, P., Ricci, D. &*	infovis
17	information aesthetics	Information aesthetics forms a cross-disciplinary link between information visualization and visualization art.		*1. Masud, L., Valsecchi, F., Ciuccarelli, P., Ricci, D. &*	infovis
18		Newspaper infographics story is almost as old as newspapers themselves		*1. Masud, L., Valsecchi, F.,*	

Google doc

WIKI

Page Discussion

Critical design

Definition [edit]

- Anthony Dunne

An attitude that rejects how things are now; provides a critique of the prevailing situation through design that embodies alternati
sensibility. Critical Design is not art!

- Barthes

Main page
Get MediaWiki
Get extensions
Tech blog
Contribute

Support
User help
FAQ
Technical manua
Support desk
Communication

Development
Bug tracker
Code repository
Code docs
Statistics
Wikimedia techn
Wikimedia audie

MediaWiki.org
Community porta
Recent changes
Translate conten
Random page
Current issues
Sandbox

Print/export
Create a book

Page Discussion

Model explainability

Definition [edit]

[1]Explain or present in understandable terms. In the context of machine learning systems is the a

References [edit]

[1] Doshi-Velez, F., & Kim, B. (2017). **Towards A Rigorous Science of Interpretable Machine L**

1. ↑ http://arxiv.org/abs/1702.08608

Main page
Get MediaWiki
Get extensions
Tech blog
Contribute

Support
User help
FAQ
Technical manual
Support desk
Communication

Development

This page was last edited on 10 July 2019, at 09:22.

Text is available under the Creative Commons Attribution-ShareAlike License; additional terms may apply. See Terms of Us

Privacy policy About MediaWiki.org Disclaimers Code of Conduct Developers Cookie statement Mobile view

Project opportunities

Real or **desired**

AlgoCount *— progetto Cariplo - UniMi - Social Sciences*

Making the **process** of information filtering and access **more accountable** to public scrutiny boy using interpretable machine learning.

CO DESIGN WORKSHOPS

Applied research

Research Through Design

Survey, Involving Users

***A field guide to Algorithms —** *Public Data Lab*

A *recipe book* for lay users concerned with algorithms miscalculations.

CO DESIGN WORKSHOPS

*** Visual experimentation on interpretable machine learning.**

Study visual analytics affordances in a ***computer science environment***

Visiting

September ‘20

social sciences

ethics

sociology

semiotics

Public Data Lab

visual analytics

information vis

info design

Konstanz University

MICROSOFT

GOOGLE

IBM

UBER

Georgia Tech
Media Studies

University of California

semiotics

dialogue

School of Poetic Computation
Brooklin

UI/UX

Bibliography

[1] Gray, C., & Malins, J. (2004). **Visualizing research: A guide to the research process in art and design**. Aldershot, Hants, England: Ashgate

[2] Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). **The ethics of algorithms: Mapping the debate**. *Big Data & Society*, *3*(2)

[3] Moere, A. Vande, & Purchase, H. (2011). ***On the role of design in information visualization****. Information Visualization 10(4) 356–371*

[4] Hohman, F. M., Kahng, M., Pienta, R., & Chau, D. H. (2018). ***Visual Analytics in Deep Learning: An Interrogative Survey for the Next Frontiers.*** *IEEE Transactions on Visualization and Computer Graphics, 1–20.*

[5] El-assady, M., Jentner, W., Kehlbeck, R., & Schlegel, U. (2019). ***Towards XAI : Structuring the Processes of Explanations Towards XAI : Structuring the Processes of Explanations****., CHI 2019*

July 2019 - Beatrice Gobbo

Publications

[1] Gobbo B., Balsamo D., Mauri M., Bajardi P., Panisson A., Ciuccarelli P. (2019). **TopicTomographies: a visual approach to distil information from media streams**. Eurovis Conference, 2019

[2] Bortoletti F., Gobbo B., Gerbino G., Elli T., Ciuccarelli P. (2018). **Venezia, la 'festa mobile': per un atlante in fieri Luoghi, figure e forme della favola antica nel primo Rinascimento.** Engramma, November 2018

July 2019 - Beatrice Gobbo

Thank you

Beatrice Gobbo

Supervisor - Paolo Ciuccarelli
Co-Supervisor - Michele Mauri

Coordinator - Paola Bertola

POLITECNICO MILANO 1863

DOTTORATO DI RICERCA IN DESIGN

DEN-SITY GN+

First Midterm Assessment / XXXIV Cycle