

Portland State University

PDXScholar

Online Northwest

Online Northwest 2017

Mar 31st, 10:15 AM - 11:00 AM

Ostriches, Minotaurs, Ghosts and Fossils in the Brave New Metadata World

Kelley McGrath

University of Oregon, kelleym@uoregon.edu

Follow this and additional works at: <https://pdxscholar.library.pdx.edu/onlinenorthwest>

Let us know how access to this document benefits you.

McGrath, Kelley, "Ostriches, Minotaurs, Ghosts and Fossils in the Brave New Metadata World" (2017). *Online Northwest*. 4.

<https://pdxscholar.library.pdx.edu/onlinenorthwest/2017/schedule/4>

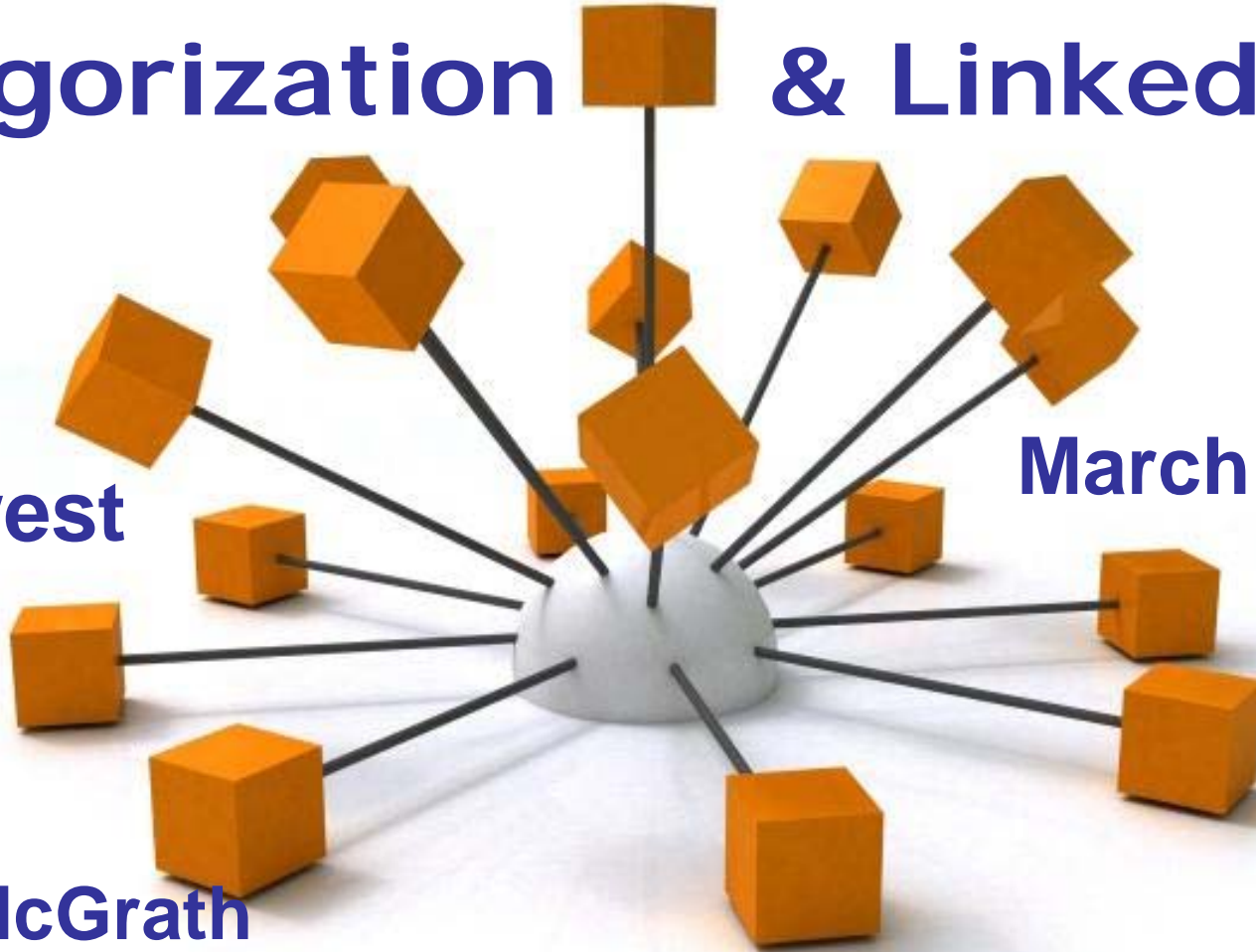
This Presentation is brought to you for free and open access. It has been accepted for inclusion in Online Northwest by an authorized administrator of PDXScholar. Please contact us if we can make this document more accessible: pdxscholar@pdx.edu.

Ostriches, Minotaurs, Ghosts and Fossils in the Brave New Metadata World: Categorization & Linked Data

**Online
Northwest**

March 31, 2017

**Kelley McGrath
University of Oregon**



Semantic Web

Combination of

- web technology (linking and identifying)
- artificial intelligence

Semantic in semantic web

not = meaning

= computable axiom (If $A=B$ and $B=C$
then $A=C$)



Inference to uncover new facts

Flipper is a dolphin
Dolphins are mammals

→ Flipper is a mammal



Inference to uncover new facts

Clint Eastwood directed *Million Dollar Baby*

Million Dollar Baby won the Academy Award for Best Picture

→ Clint Eastwood directed a film that won the Academy Award for Best Picture



Inference to uncover new facts

Library data is often free-text with minimal mark-up

245 Million Dollar Baby

586 Academy Award, 2005: Best achievement in directing, **Best motion picture of the year**, Best performance by an actor...



Inference to uncover new facts

Structured data with identifiers

Film identifier	Film title	Year
no2005028500	Million dollar baby (Motion picture : 2004)	2004

Award identifier	Award name
Q102427	Academy Award for Best Picture

Award identifier	Film identifier	Year
Q102427	no2005028500	2005



Inference to uncover new facts

How much do you want to record?

Best Picture
Best Director
Best Actor in a Leading Role
Best Actor in a Supporting Role
Best Actress in a Leading Role
Best Actress in a Supporting Role
Best Animated Feature
Best Animated Short Film
Best Cinematography
Best Costume Design

Best Documentary Feature
Best Documentary Short Subject
Best Film Editing
Best Foreign Language Film
Best Live Action Short Film
Best Makeup and Hairstyling
Best Original Score
Best Original Song
Best Production Design
Best Sound Editing
Best Sound Mixing
Best Visual Effects
Best Writing (Adapted Screenplay)
Best Writing (Original Screenplay)



Inference to uncover new facts

Linked data

http://id.loc.gov/authorities/names/no2005028500	http://rdaregistry.info/Elements/w/titleOfWork.en	Million dollar baby (Motion picture : 2004)
http://id.loc.gov/authorities/names/no2005028500	http://rdaregistry.info/Elements/w/dateOfWork.en	2004
http://id.loc.gov/authorities/names/no2005028500	https://www.w3.org/2002/07/owl#sameas	https://www.wikidata.org/wiki/Q10242

https://www.wikidata.org/wiki/Q10242	https://www.w3.org/2000/01/rdf-schema#label	Academy Award for Best Picture
https://www.wikidata.org/wiki/Q184255	https://www.w3.org/2000/01/rdf-schema#label	Million Dollar Baby
https://www.wikidata.org/wiki/Q184255	https://www.wikidata.org/wiki/Property:P166	https://www.wikidata.org/wiki/Q10242

Classical Theory of Categories

Does it have feathers?
Can it fly?

- clearly-defined
- mutually exclusive
- collectively exhaustive



Defined by necessary and jointly
sufficient conditions



The Unknowable



London After Midnight, one of the most sought after lost films, had its last surviving reel copy destroyed in a devastating studio vault fire and only exists in the form of film stills, as shown in an original 1927 theatrical release poster.



https://en.wikipedia.org/wiki/Lost_film

The Project

Directed by Clint
Eastwood

“Directed by”
“Clint Eastwood”

Directed by →
direction



Thank you to
Alden Lee,
Jaroslaw Szurek
and many others
for help with
categorization

Top 11 of 76 categories

production	14.4%
acting	13.6%
direction	10.8%
cinematography	10.1%
writing	9.2%
editing	8.4%
music	7.0%
presents	3.7%
composers	2.4%
other/unsure	2.0%
a film by /a ... film	2.0%
	83.5%

The Vague

responsible, Lê Mỹ Phương



The Vague

responsible, Lê Mỹ Phương

action, Shyam Kaushal

associate, Ved M Rao

Series proposed by Benoit Peeters

supervisor, Dr. Nurdin Perdana

Team works, Kartawijaya ... [et al.]



The Ambiguous

By Shakespeare [author of play]
By W.A. Mozart [composer of opera]
By John Cleese and Connie Booth

Bass [guitar or vocalist?]
Music [performer or composer?]
Songs [writer or singer?]



The Ambiguous

Musical adaptation and direction, Penella
[musical direction or direction?]

associate director and editor, Alexander
Hammid [associate editor or editor?]

author and singer of songs, Vladimir Vysotski
[author or author of songs?]

graphics and video editor, Michael Seibert
[graphics or graphics editor?]



The Incoherent

Winstar Cinema release, Shochiku Co.,
Ltd. presents a 3H Film Productions of
a film by Hou Hsiao-Hsien

an Old Photo film presents

A Sony Pictures Classics release of Vans
"Off the Wall" Productions presents an
ADP Productions

Villealfa Filmproductions esittää ; Aki
Kaurismäen elokuva



The Unique

a film by Wes Craven

This is not a film by Khvan

a média-stylo by Renée Gosson

the immortal classic by S. Anski

a Park Chan Wook film

a Spike Lee joint

a Danny Shechter dissection

色怠楸鰯



The Unique

a film by Wes Craven

This is not a film by Khvan

a média-stylo by Renée Gosson

the immortal classic by S. Anski

a Park Chan Wook film

a Spike Lee joint

a Danny Shechter dissection



色怠楸鰯

кино-разведчик Дзига Вертов*

*Thanks to Tom Dousa for researching this term. **Page 18**

The Unnamed

Showrunner

“responsible for all creative aspects of the show”

“makes all important decisions about the series' scripts, tone, attitude, look and direction... oversees casting, production design and budget... chooses directors and guest stars...”



The Unnamed

Showrunner

Usually credited as

- Executive producer
- Created by



The Unnamed

Cast:

Sicilian fishermen
Shaima and Hossein and families
the children of Biviers
a group of African students of the
University of Rome
inmates (of the State Prison in Ellis,
Texas)



The Untranslatable

realización = production? direction?

Dirección, realización y producción

总导演 = director-in-chief? executive director?

奚 = planning? development? production?

mise en scène = staging? direction?



The Long Long Tail

- Dream sequences based on designs by Salvador Dali.
- garden designer
- synchronization director
- tiger trainer
- spider web spinner
- floor drop painted by ...
 - body art & locations
 - movement coordinator



Not to Even Start on Instruments

From accordion to whistles
From clarinet to clavinet



From around the world:
bodhrán, conga,
kenkeni, mridangam,
tabla, taiko and so on



IMDb-Style Solution?

Does the Internet Movie Database (IMDb) provide a potential model?

- Limited list of broad role categories
- Very specific, free-text roles



IMDb-Style Solution?

Romeo and Juliet (1968)

Controlled vocabulary for Broad Terms:

Writing Credits

<u>William Shakespeare</u>	...	(play)
<u>Franco Zeffirelli</u>	...	(screenplay)

Grab Bag Category:

Other crew [=contributor?]

<u>Gabriella Bernardi</u>	...	production secretary
<u>Alberto Testa</u>	...	choreographer



The Variations

Assistant producer
Associate producer
Chief producer
Delegate producer
Deputy producer
Executive producer
DVD producer

Segment producer
Producer in Japan
Series producer
Series senior
producer
Producer for BBC
Producer of U.S.
release

Executive producer of
English version

Line producer



A case for narrower categories

Most popular Thai language films



Details

Official Sites: [Official Facebook](#) | [Official site](#)
Country: [USA](#) | [Germany](#)
Language: [English](#) | [Burmese](#) | [Thai](#)



- Primary or secondary
- Spoken, sung, signed or written (subtitles, captions or intertitles)

A case for narrower categories

Specific types of dates



Ivan the Terrible, Part II (1958)

- Copyright
- Production
- Recording / Filming
- Release / Distribution / Broadcast
- Unknown



It depends...

television producer

producer of *made-for-TV movie*

producer of *TV drama/comedy series*

- writer
- line producer
- no particular reason (vanity credit)

producer of *TV news*

- have editorial and supervisory roles
- select and order stories
- may write or edit stories



The Fickle

- technical director (some silent films)
- art director
- production designer



The Infinite Shades of Gray

Adaptation

“a continuum from works that are clearly adapted from Robinson Crusoe, to works that only belong to the genre known as “Robinsonade” or “desert island fiction,” and in which no more than the Robinson motif in a very general sense remains”



Wallheim, Henrik. “From Complex Reality to Formal Description: Bibliographic Relationships and Problems of Operationalization in RDA.” *Cataloging & Classification Quarterly* 54.7 (2016).

The Infinite Shades of Blue



Classical Theory of Categories

Does it have feathers?
Can it fly?

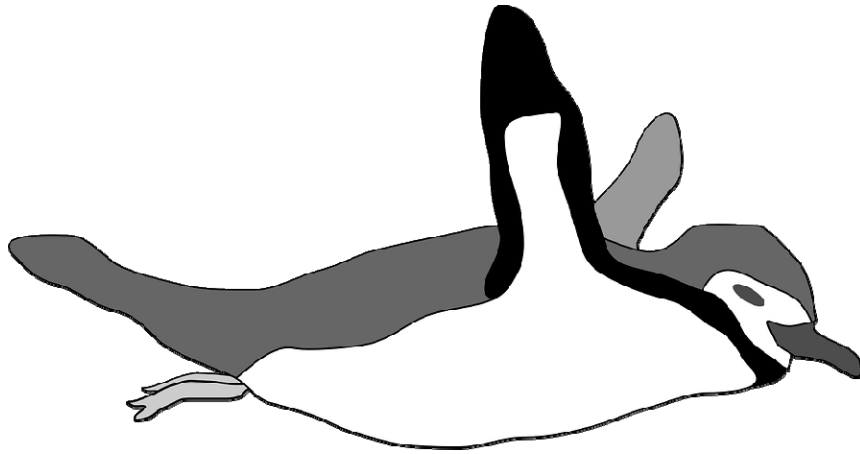
- clearly-defined
- mutually exclusive
- collectively exhaustive



Defined by necessary and jointly
sufficient conditions

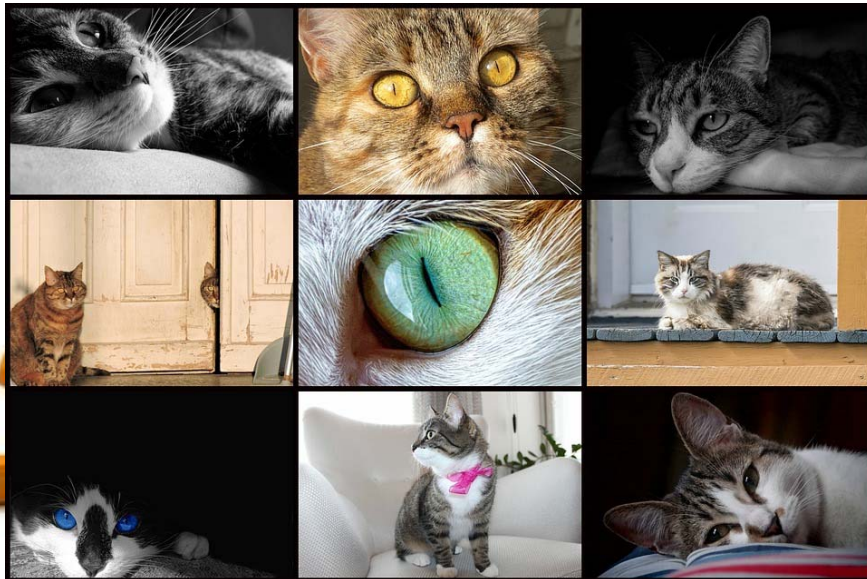


Prototype theory of categories



Machine Learning

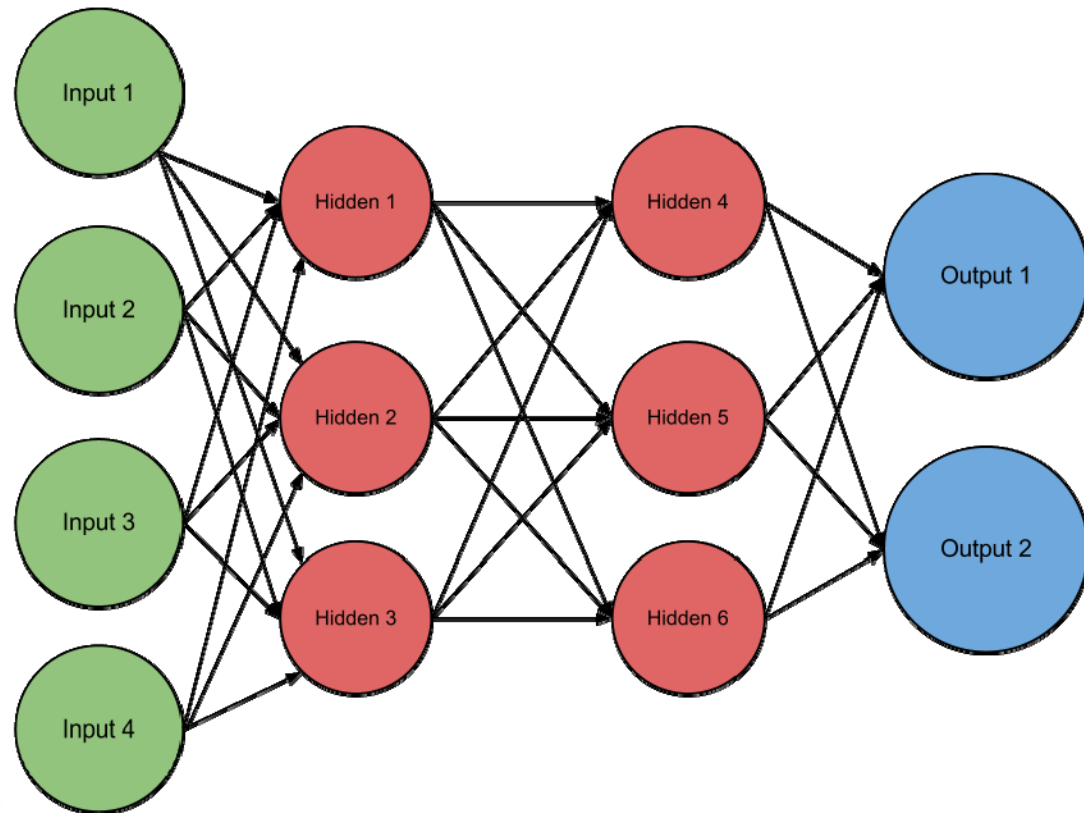
- Google Translate improvements
- Defeat of highly-ranked go players (10^{360} possible positions vs. 10^{123} for chess)



- Identification of cats in pictures

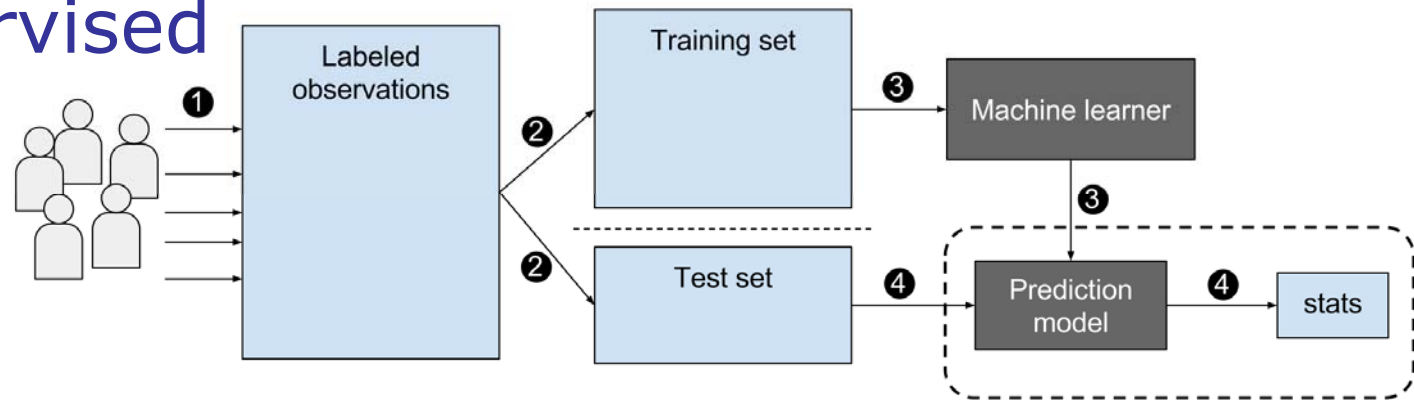
Machine Learning

Give raw data to an algorithm and let the algorithm make its own rules



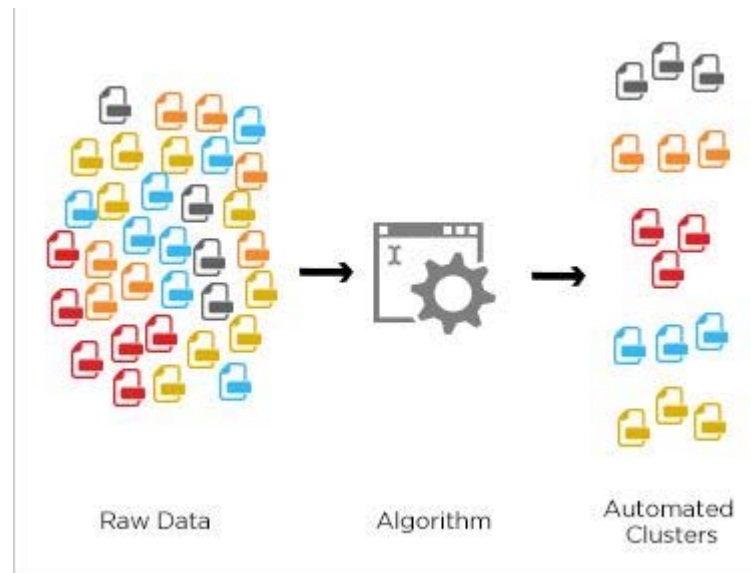
Machine Learning

- Supervised



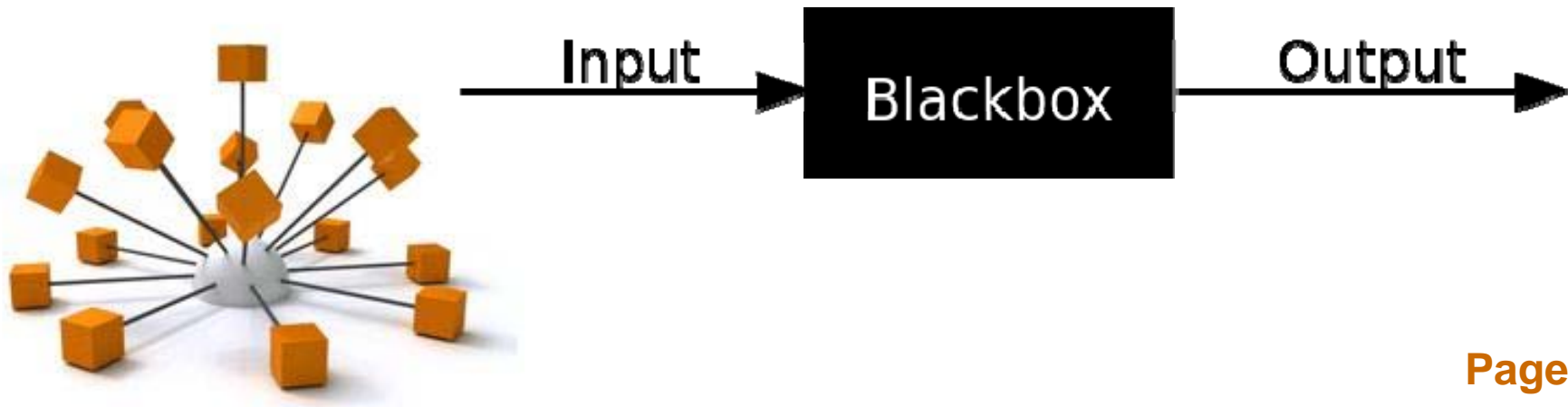
- Semi-supervised

- Unsupervised



Machine Learning

- Iterative, trial & error
- Statistical
- Fuzzy
- Black box



Machine Learning

“directed by Clint Eastwood”

Rust, Ronnie

<http://opaquenspace.org/ns/people/RustRonnie>

[Return to Vocabulary](#)

Type:	Personal Name
Label:	Rust, Ronnie English [en]
Alternate Name:	
Date:	
Comment:	University of Oregon Baseball team 2016-2017 English [en]



Related Links

Operationalizing relationship designators

<http://www.tandfonline.com/doi/abs/10.1080/01639374.2016.1200169?journalCode=wccq20>

Machine learning

<https://medium.com/@ageitgey/machine-learning-is-fun-80ea3ec3c471>



<https://www.nytimes.com/2016/12/14/magazine/the-great-ai-awakening.html>