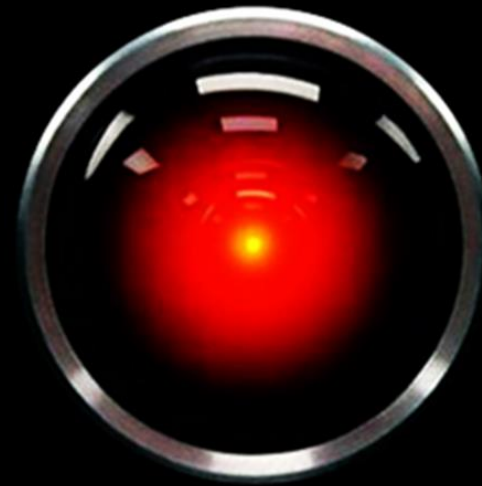


The Web We Mix

benevolent AIs for a resilient web

Fabien Gandon, <http://fabien.info>



I'm sorry, Dave.
I'm afraid I can't do that.

WIMMICS TEAM

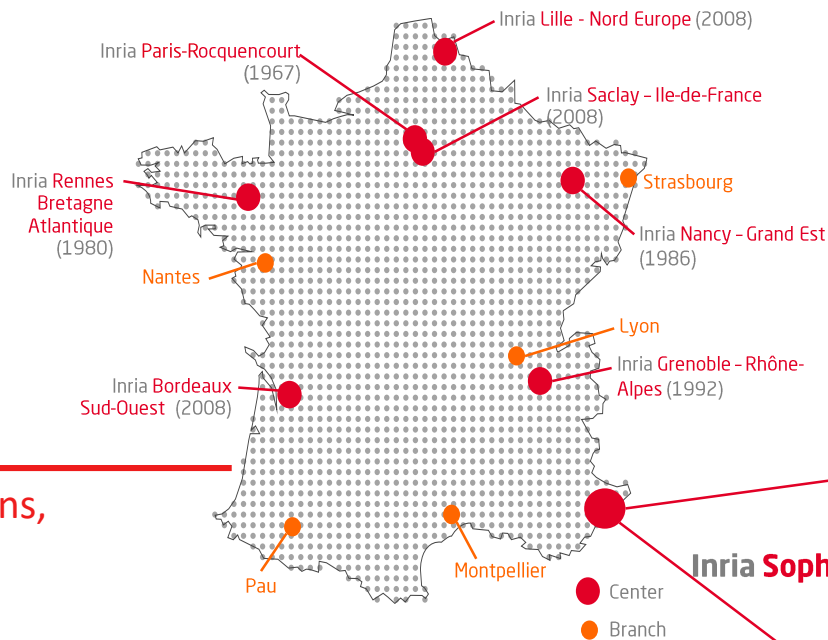
Web-Instrumented Man-Machine Interactions,
Communities and Semantics

■ Inria *Inria*

■ CNRS 

■ University Côte D'Azur (UCA)

UNIVERSITÉ
CÔTE D'AZUR 

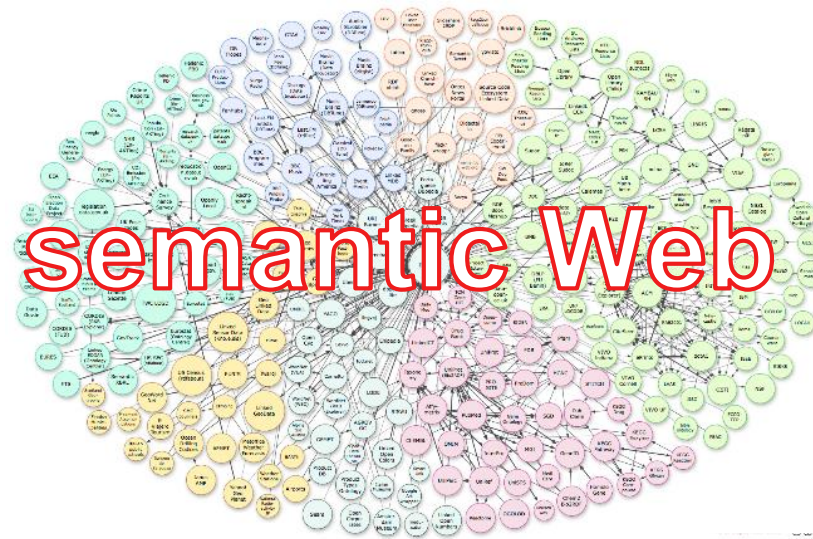


Inria Sophia Antipolis Méditerranée (1983) & UCA, I3S, CNRS - Nice, FRANCE

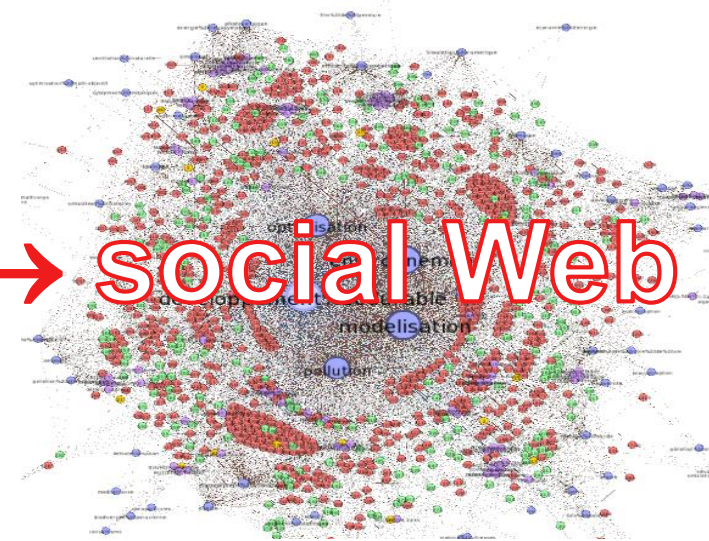


CHALLENGE

to bridge social semantics and formal semantics on the Web



social Web



MULTI-DISCIPLINARY TEAM

- 40~50 members
- ~15 nationalities
- 1 DR, 3 Professors
- 3CR, 4 Assistant professors



DR/Professors:

- Fabien GANDON, Inria, AI, KRR, Semantic Web, Social Web
- Nhan LE THANH, UCA, Logics, KR, Emotions, Workflows
- Peter SANDER, UCA, Web, Emotions
- Andrea TETTAMANZI, UCA, AI, Logics, Agents,

CR/Assistant Professors:

- Michel BUFFA, UCA, Web, Social Media
- Elena CABRIO, UCA, NLP, KR, Linguistics, Q&A, Text Mining
- Olivier CORBY, Inria, KR, AI, Sem. Web, Programming, Graphs
- Catherine FARON-ZUCKER, UCA, KR, AI, Semantic Web, Graphs
- Alain GIBOIN, Inria, Interaction Design, KE, User & Task models
- Isabelle MIRBEL, UCA, Requirements, Communities
- Serena VILLATA, CNRS, AI, Argumentation, Licenses, Rights

Research engineer: Franck MICHEL, CNRS, Linked Data, Integration, DB

External:

- Andrei Ciortea (University of St. Gallen) Agents, WoT, Sem. Web
- Nicolas DELAFORGE (Mnemotix) Sem. Web, KM, Integration
- Claude FRASSON (University of Montreal) Emotion, eLearning
- Freddy LECUE (Thales, Montreal) AI, Logics, Mining, Big Data, Sem. Web

AI for classical Web tasks





Dogs' UI metaphor
to Machine Learning



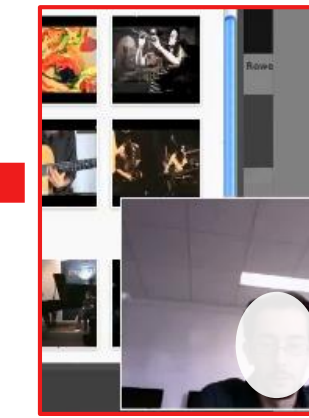
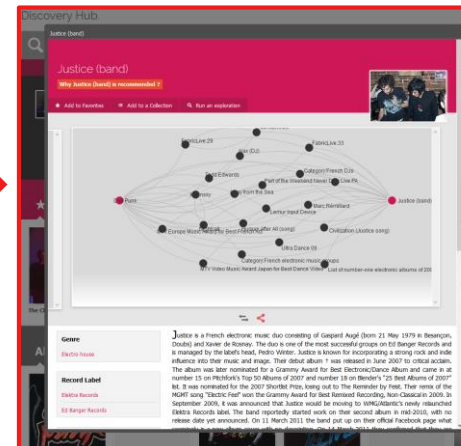
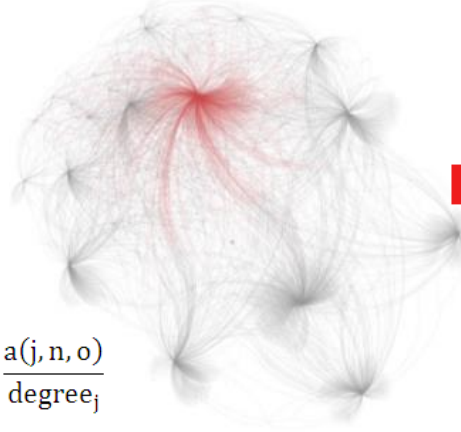
Sent to Web kennels
to work offline

Autonomy AgentWare Web Researcher, 1996

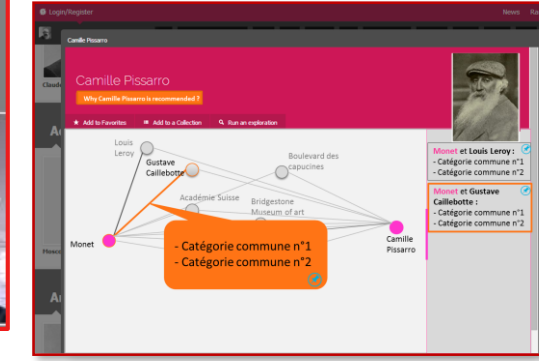
semantic spreading activation

$$a(i, n) = \prod_{o \in O} [a(i, n, o)] / \log(\text{degree}_i)$$

$$a(i, n + 1, o) = s(i, n, o) + \sum_j w(i, o) * \frac{a(j, n, o)}{\text{degree}_j}$$

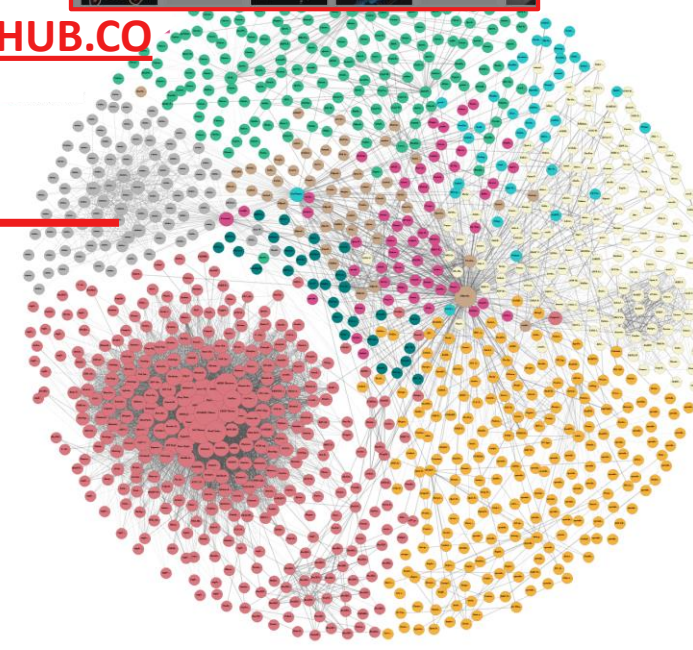
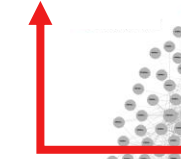


new evaluation protocol



[Marie, Giboin, Palagi et al.]

DISCOVERYHUB.CO



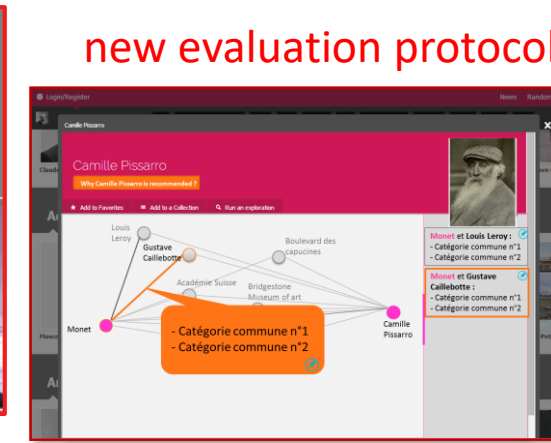
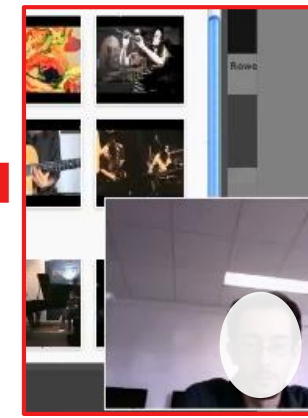
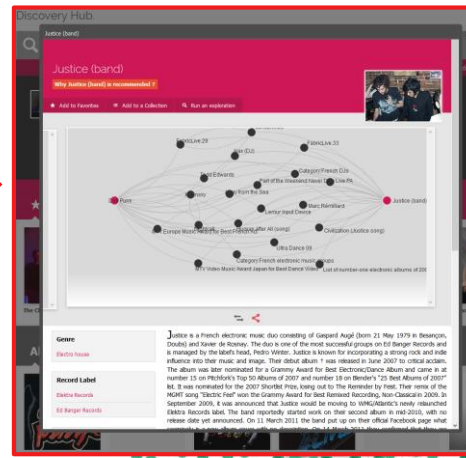
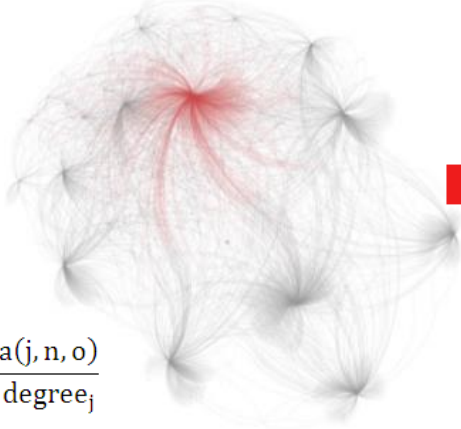
SEARCHING

- exploratory search
- question-answering

semantic spreading activation

$$a(i, n) = \prod_{o \in O} [a(i, n, o)] / \log(\text{degree}_i)$$

$$a(i, n + 1, o) = s(i, n, o) + \sum_j w(i, o) * \frac{a(j, n, o)}{\text{degree}_j}$$

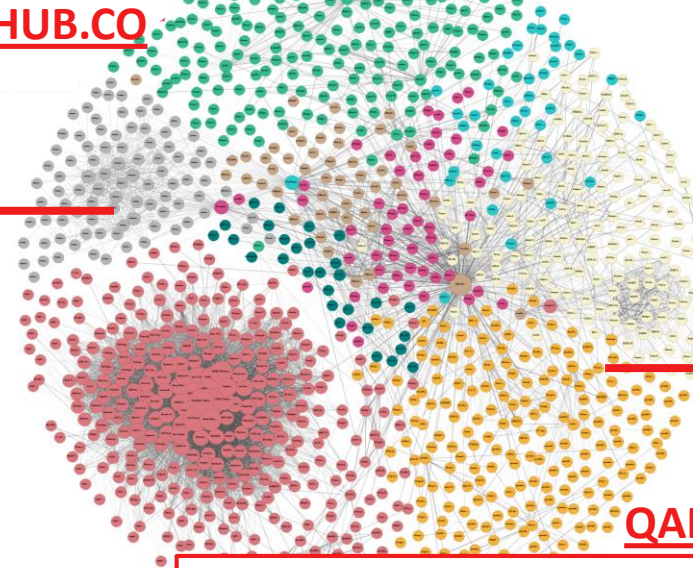


new evaluation protocol

[Marie, Giboin, Palagi et al.]

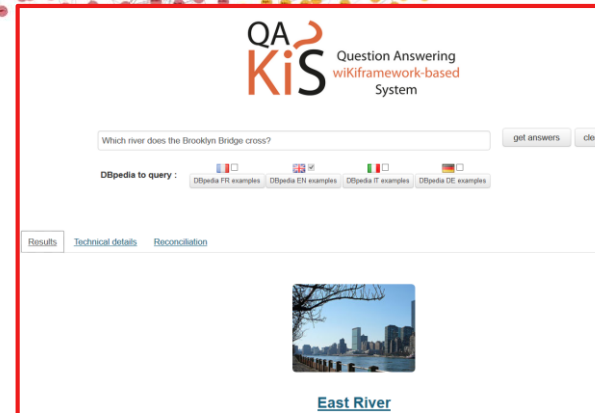
DISCOVERYHUB.CO

QAKIS.ORG



linguistic relational pattern extraction

starring(Work, Person)
 [D:Work], played by [R:Person]
 [D:Work] stars [R:Person]
 [D:Work] film stars [R:Person]



named entity recognition
 similarity based SPARQL
 generation

```
select * where {
  dbpr:Batman_Begins dbp:starring ?v .
  OPTIONAL {?v rdfs:label ?l
    filter(lang(?l)="en") } }
```

[Cabrio et al.]

SEARCHING

- exploratory search
- question-answering



BROWSING

e.g. SMILK plugin

[Lopez, Cabrio, et al.]

(AFP) - Robes à panier à motifs floraux, gilets longs portés avec des bermudas: dans la droite ligne de sa collection haute couture, Raf Simons a revisité la mode du 18ème siècle, avec un oeil futuriste, pour proposer une silhouette moderne au quatrième jour des défilés de prêt-à-porter parisiens.

Le show, auquel assistaient entre autres Carla Bruni-Sarkozy et Marion Cotillard, se tenait dans la Cour carrée du Louvre. Sur l'installation abritant le défilé, tout en miroirs, se réfléchissaient les façades du musée, au coeur du Paris historique.

Sur un podium en forme de navette spatiale, vêtements de cour et combinaisons de pilotes côtoient des robes blanches évoquant des uniformes scientifiques chic.

"J'ai commencé par me demander: qu'est-ce que la modernité? (...) Il me semblait plus contemporain d'aller vers un passé lointain plutôt que de moderniser l'esprit des dernières décennies", expose le couturier belge.

Pour le PDG de **Christian Dior**, Sidney Toledano, cette collection "s'adresse à une femme jeune, moderne, dans un style de vie totalement nouveau".

S'inspirer du passé permet selon lui à la mode de durer. "C'est ce que d'autres secteurs ne savent pas faire. Dans la technologie, on regarde toujours le produit de demain, futuriste, et finalement éphémère. En revanche la mode, à ce niveau-là, n'est pas dans l'éphémère, elle se réinspire du passé, sans passéisme, en se projetant", déclare-t-il à l'AFP après le show.

Contrairement à d'autres maisons parisiennes, **Dior** a décidé de retransmettre son défilé sur son site internet, s'adressant directement au public. "On lui parle directement mais avec des images

Informations supplémentaires

Intitulé: Christian Dior
Catégorie: Marque

DBpedia

Dior

Christian Dior, né le 21 janvier 1905 à Granville, dans la Manche, mort le 24 octobre 1957 à Montecatini Terme en Italie, est un grand couturier français. Il est le fondateur de la maison de couture qui porte son nom.

Ressources

- http://fr.dbpedia.org/resource/Catégorie:Naissance_en_1905
- http://fr.dbpedia.org/resource/Catégorie:Élève_de_l'Institut_d'études_politiques_de_Paris
- http://fr.dbpedia.org/resource/Catégorie:Mort_d'une_crise_cardiaque
- http://fr.dbpedia.org/resource/Catégorie:Personnalité_normande
- http://fr.dbpedia.org/resource/Catégorie:Décès_en_1957
- http://fr.dbpedia.org/resource/Catégorie:Naissance_à_Granville
- http://fr.dbpedia.org/resource/Catégorie:Haute_couture
- http://fr.dbpedia.org/resource/Catégorie:Couturier_français
- [http://fr.dbpedia.org/resource/Catégorie:Christian_Dior_\(entreprise\)](http://fr.dbpedia.org/resource/Catégorie:Christian_Dior_(entreprise))
- http://fr.dbpedia.org/resource/Catégorie:Décès_dans_la_province_de_Pistoia
- http://fr.dbpedia.org/resource/Catégorie:Wikipédia:Outil_de_retour_des_lecteurs
- http://fr.dbpedia.org/resource/Catégorie:Élève_du_collège_Stanislas_de_Paris

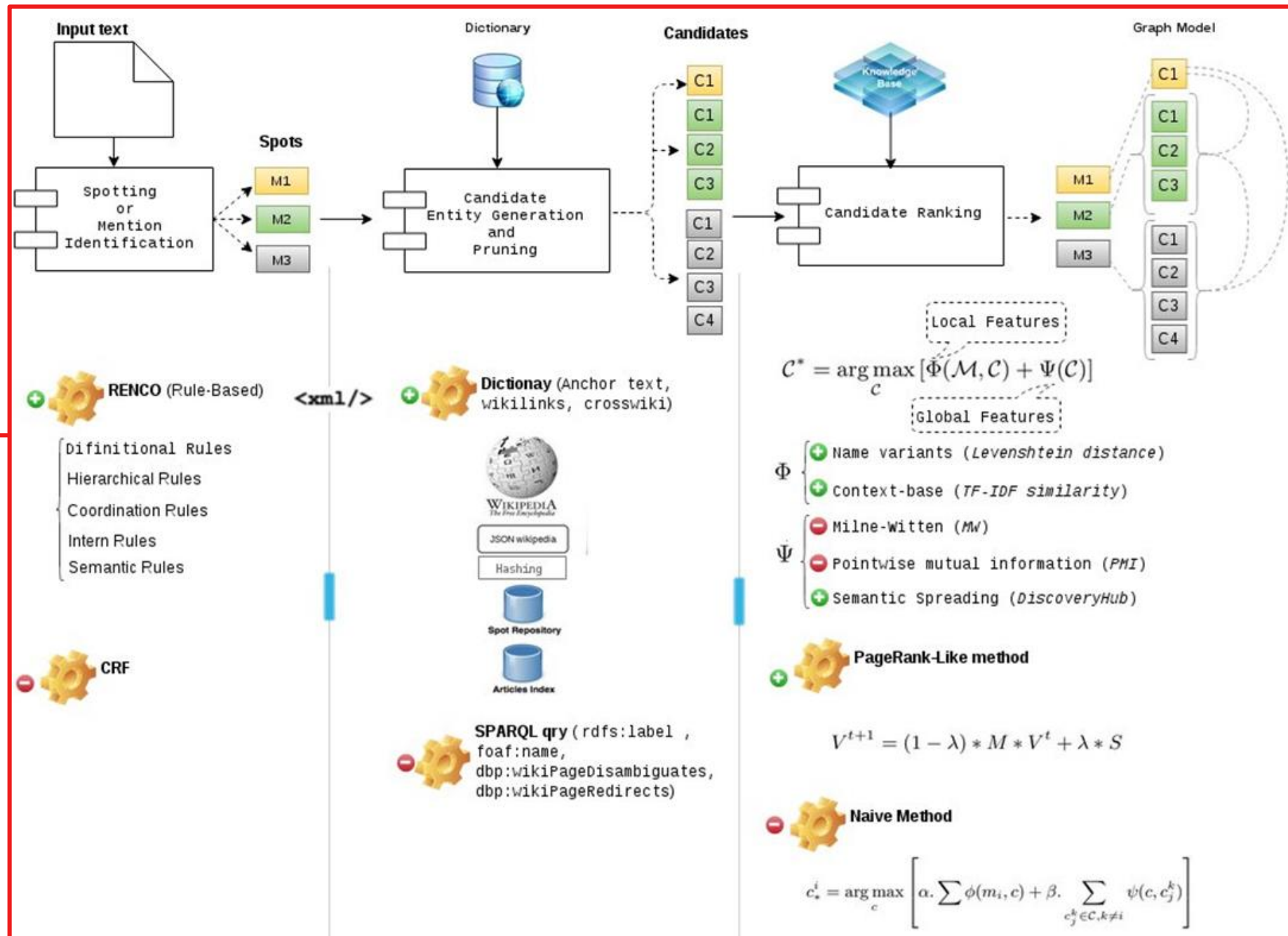
NetSent

Sentiment	Count
Positive	26
Negative	11

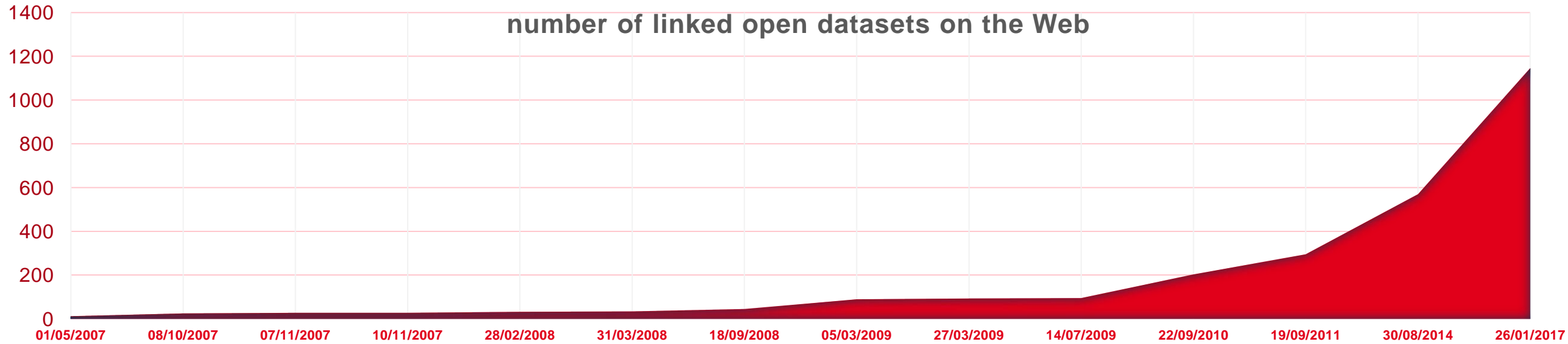
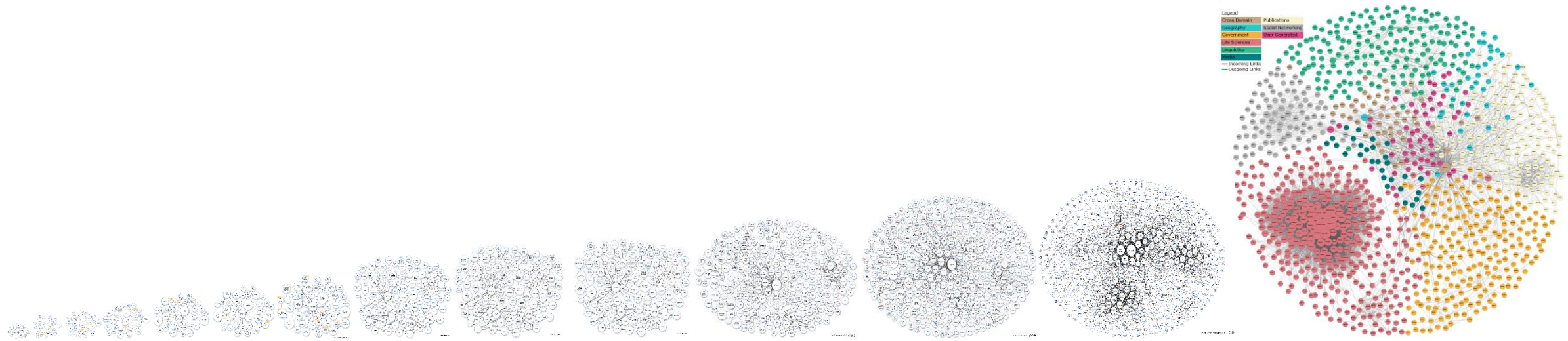
BROWSING

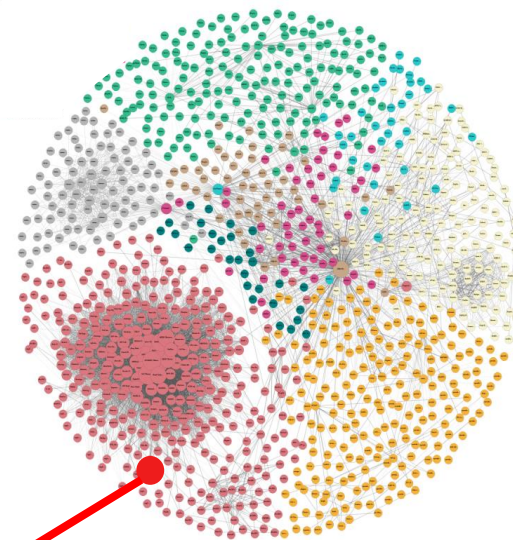
e.g. SMILK plugin

[Nooralahzadeh, Cabrio, et al.]



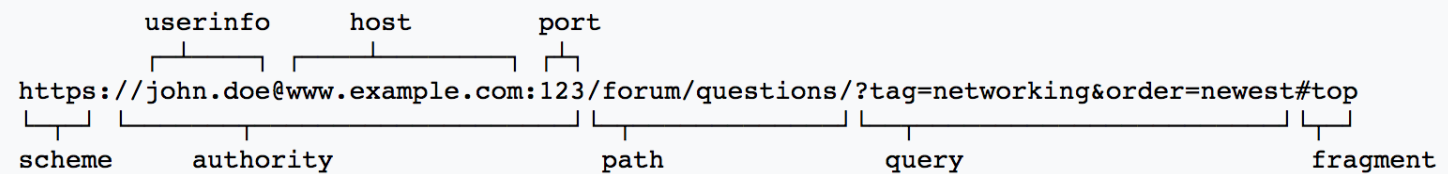
linked open data(sets) cloud on the Web

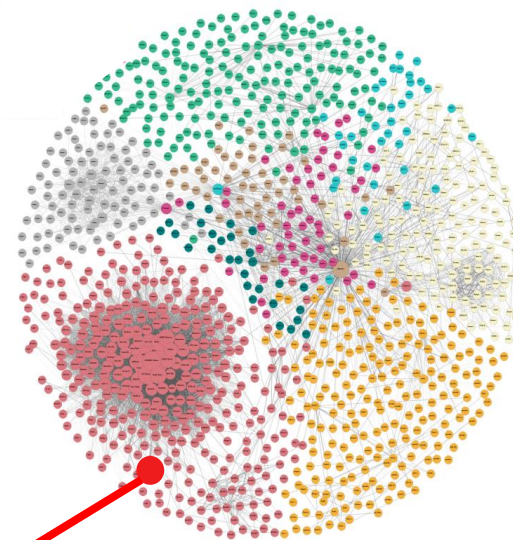




CRAWLING [Huang, Gandon 2019]

- Predict data availability
- Select features of URIs





CRAWLING [Huang, Gandon 2019]

- Predict data availability
- Select features of URIs
- Learn crawling selection
- Online learning w. crawling

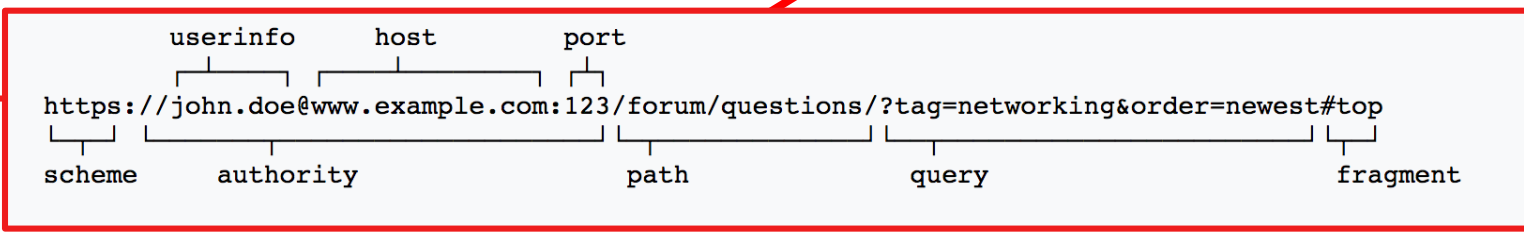
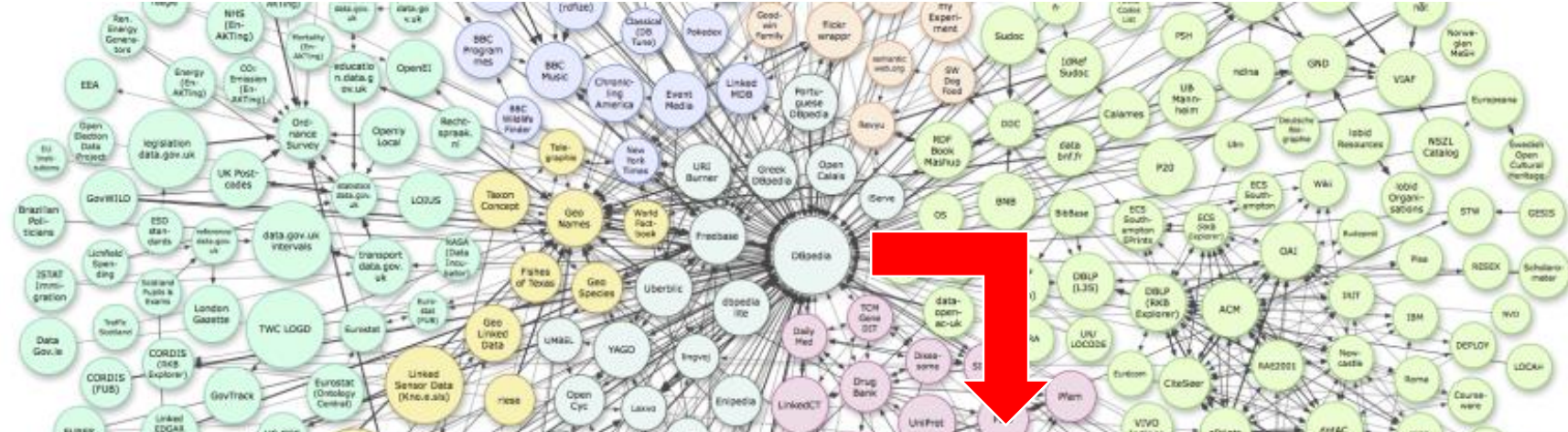


Table 1. Performance of the combinations of feature sets

Combination of Feature Sets	KNN		Naive Bayes		SVM	
	F-measure	Accuracy	F-measure	Accuracy	F-measure	Accuracy
$F_{+t+r+p+x}$	0.6951	0.7407	0.7154	0.7462	0.7944	0.7722
F_{+t+p+x}	0.6261	0.6832	0.7094	0.7413	0.7801	0.7643
F_{+t+r+x}	0.6773	0.7121	0.7111	0.7448	0.7829	0.7650
F_{+t+r+p}	0.7592	0.7731	0.7660	0.7701	0.8216	0.7902
F_{+r+p+x}	0.6015	0.7010	0.6328	0.7075	0.6839	0.7074
F_{+t+x}	0.5582	0.6912	0.6012	0.6172	0.6828	0.6277
F_{r+p}	0.3953	0.5810	0.4874	0.6097	0.6790	0.6424
F_{p+x}	0.4392	0.5739	0.6086	0.6238	0.6689	0.6269

Crawler	Percentage
BFS	0.302
crawler_NB (20K)	0.341
crawler_NB (40K)	0.345
crawler_SVM (20K)	0.402
crawler_SVM (40K)	0.413
crawler_KNN (20K)	0.331
crawler_KNN (40K)	0.324
LDCOC ($\tau = 0.5, \epsilon = 0.17$)	0.655



EDTECH & WEB

for e-learning & serious games

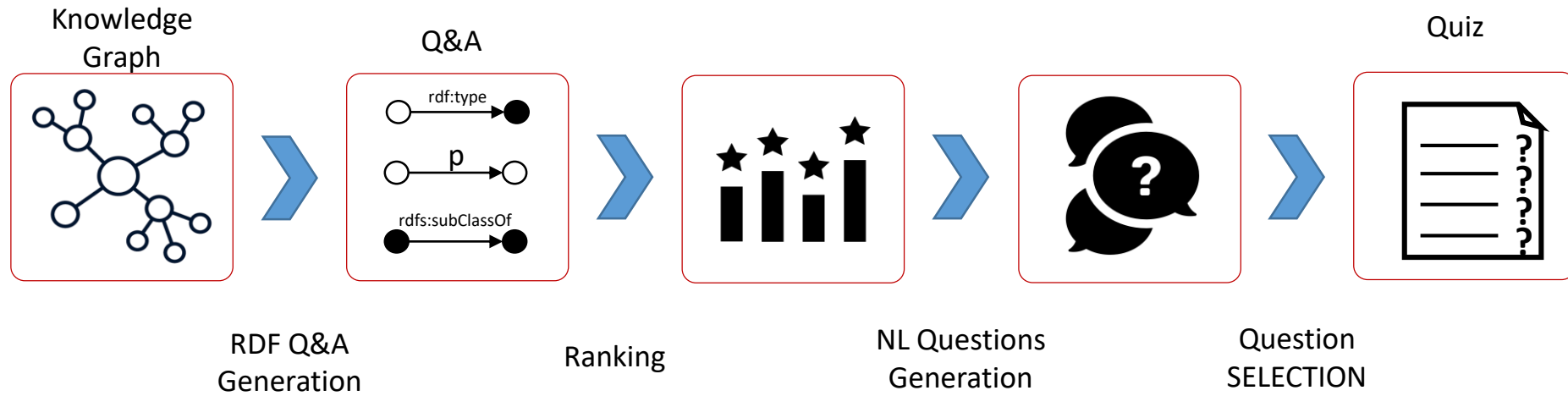
[Rodriguez-Rocha, Faron-Zucker et al.]

Player: John Doe Current game score: 0 Total score: 0 Errors left: 3 Game level: 1

Start	100	300	100	End
	300		300	
200				300
200				100
200				50
50		200		
50	300	300		

For 200 points:

What is the capital of Austria?



QUIZZES

Automated generation of quizzes [Rodriguez-Rocha, Faron-Zucker et al.]

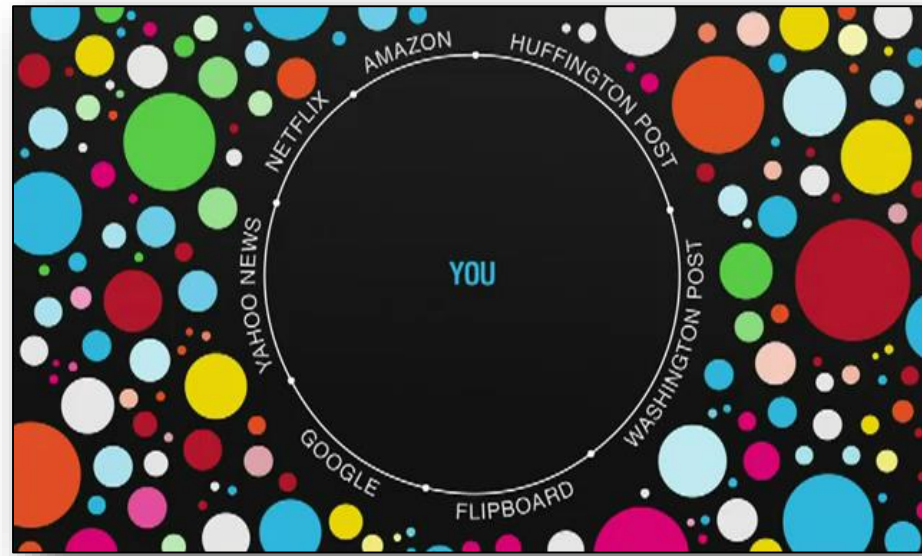
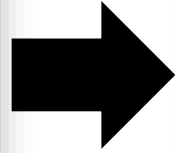


but...

Google

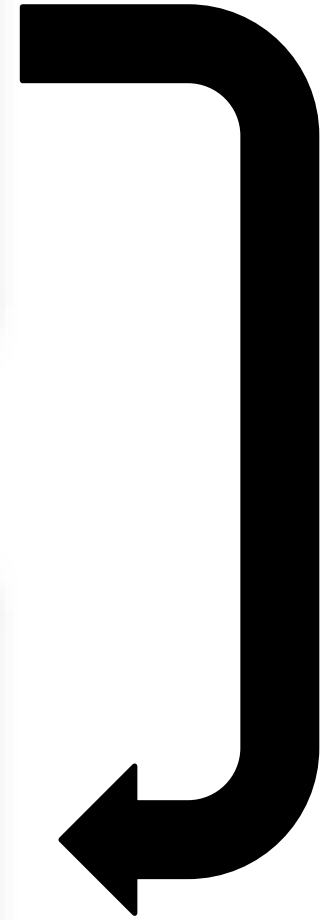
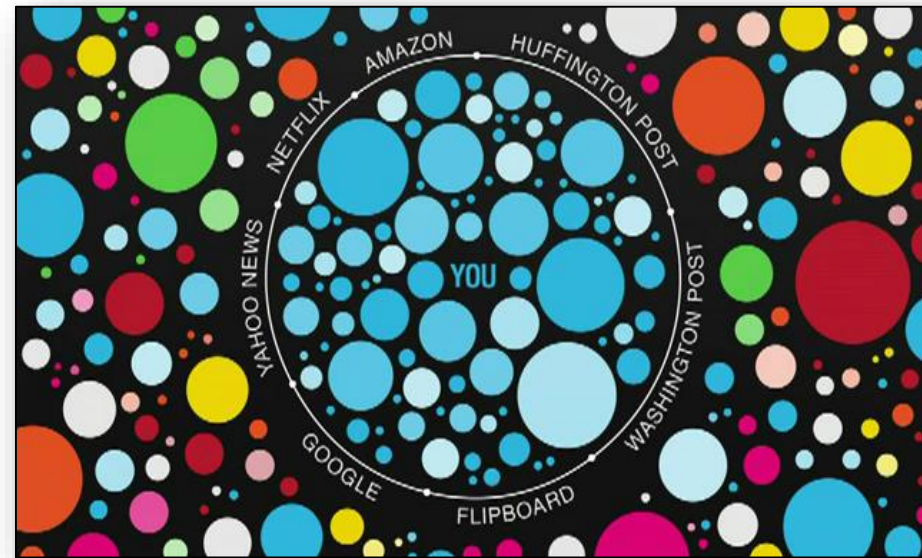
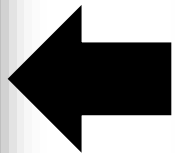
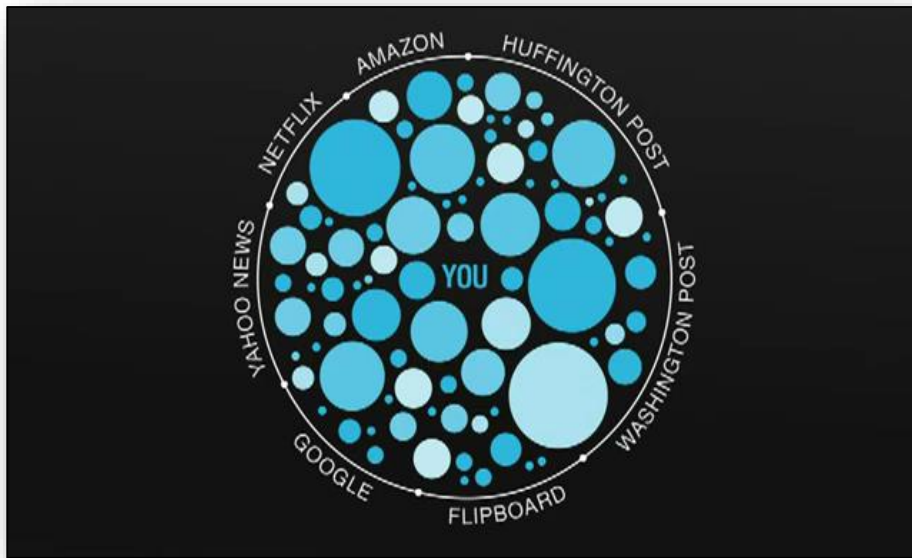
404. That's an error.





over adaptation & filter bubble

(c.f. Eli Pariser)



AITO BURST BUBBLES





Translate

French English Slovenian Detect language ▾



English German French ▾

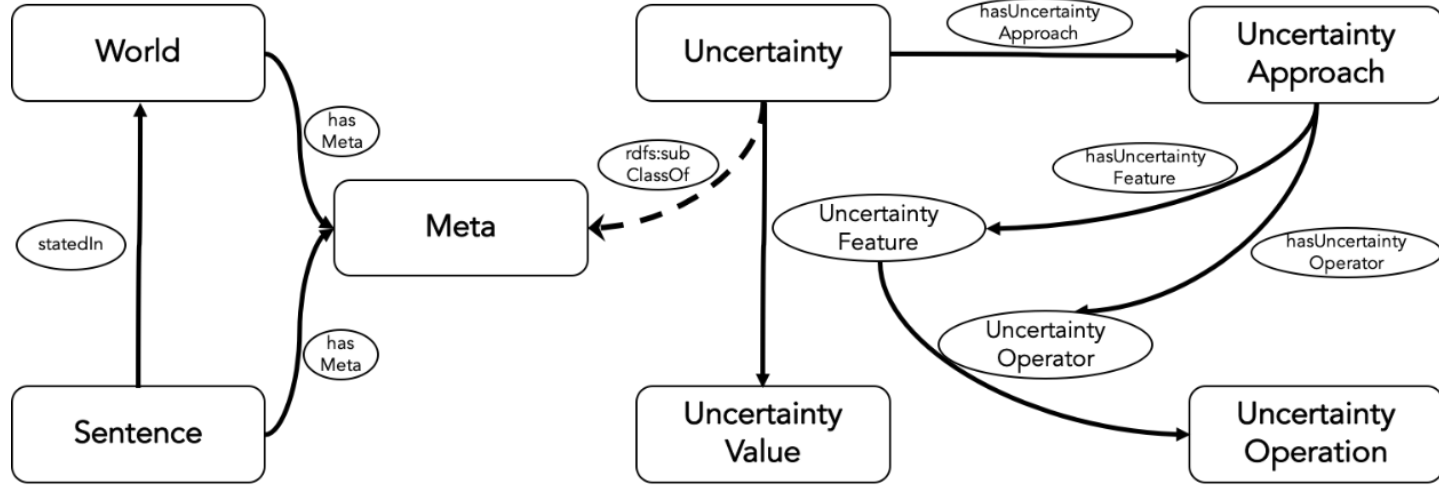
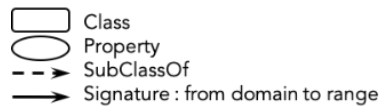
Translate

world-wide web |



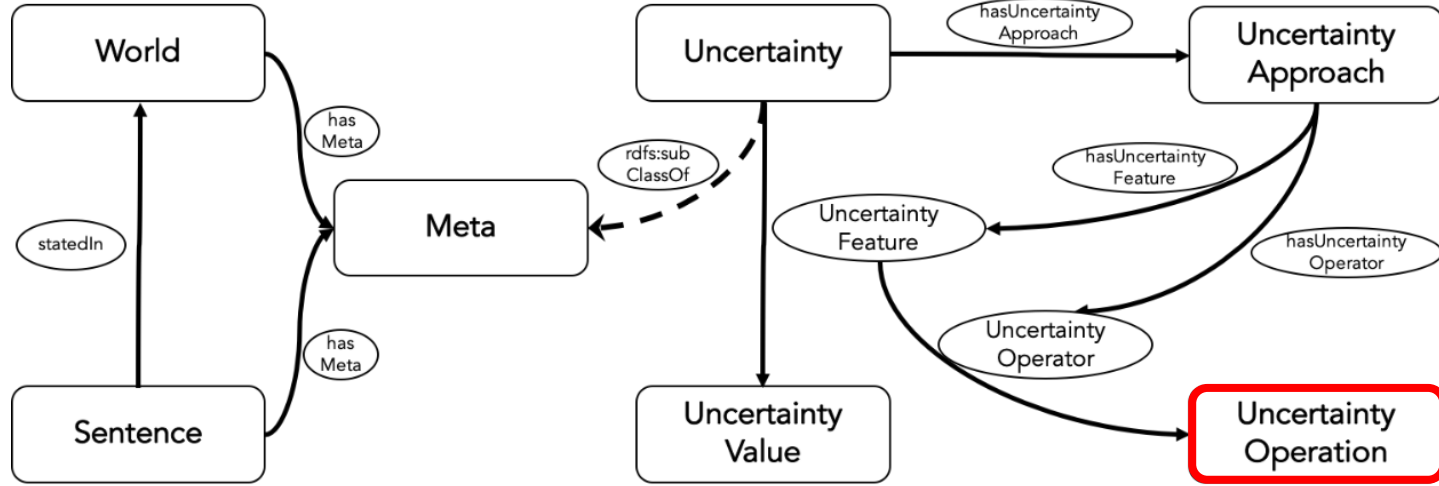
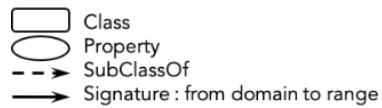
Internet

bias...



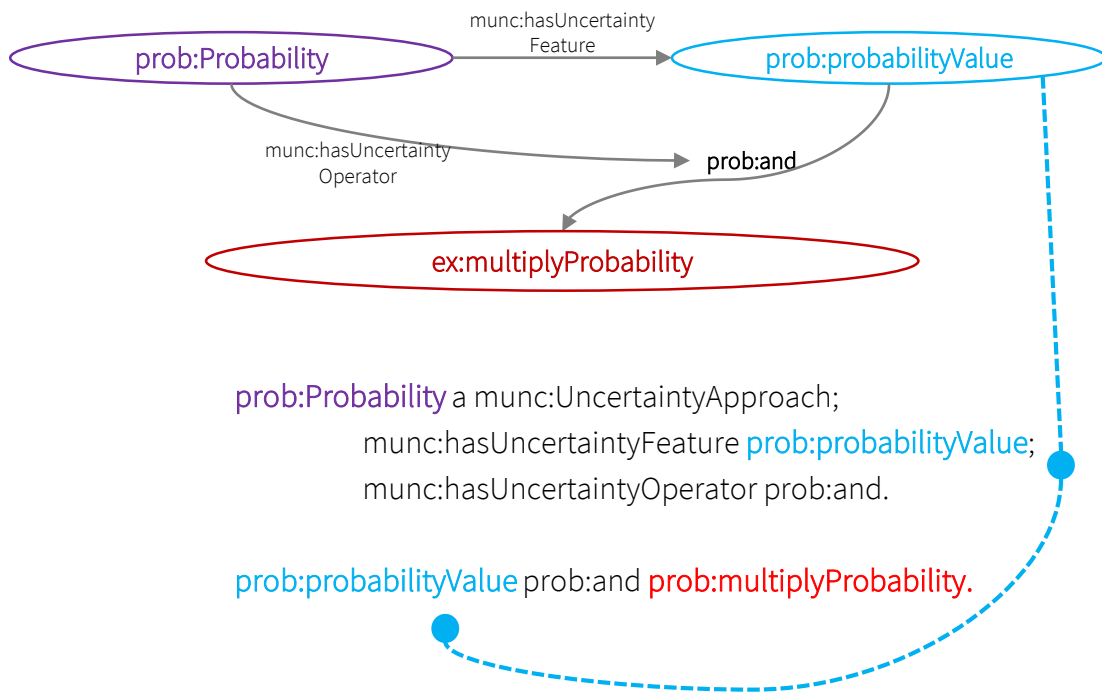
UNCERTAINTY [Djebri et al 2019]

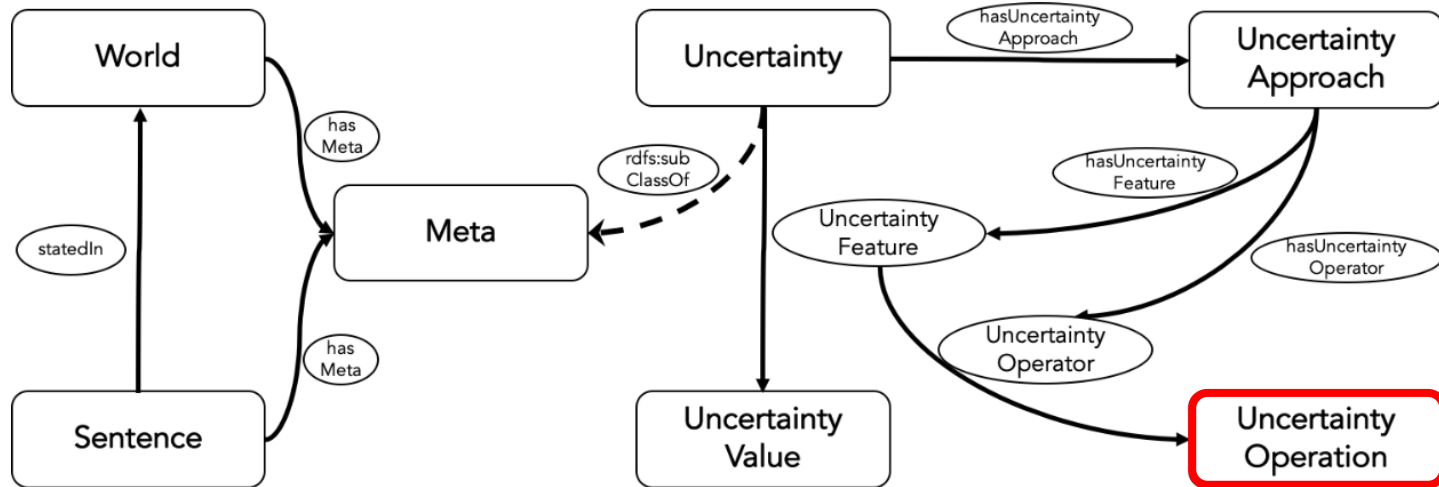
Can uncertainty be formalized on top of the standards of the Semantic Web, to be published on the Web ?



UNCERTAINTY [Djebri et al 2019]

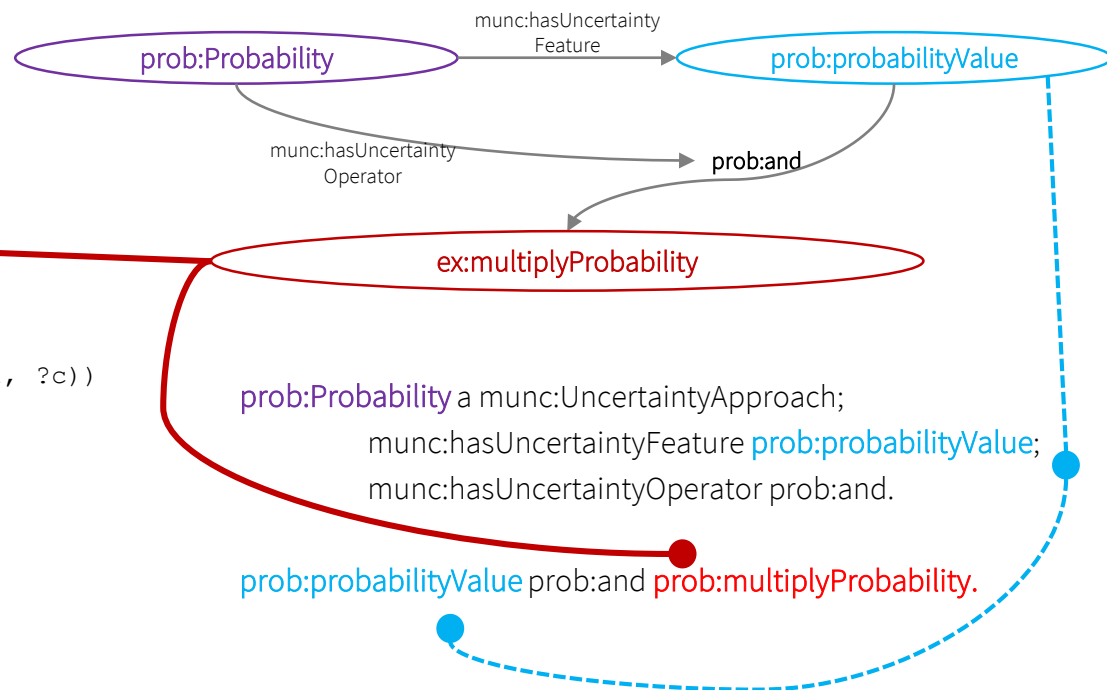
Can uncertainty be formalized on top of the standards of the Semantic Web, to be published on the Web ?





UNCERTAINTY [Djebri et al 2019]

LINKED CALCULI

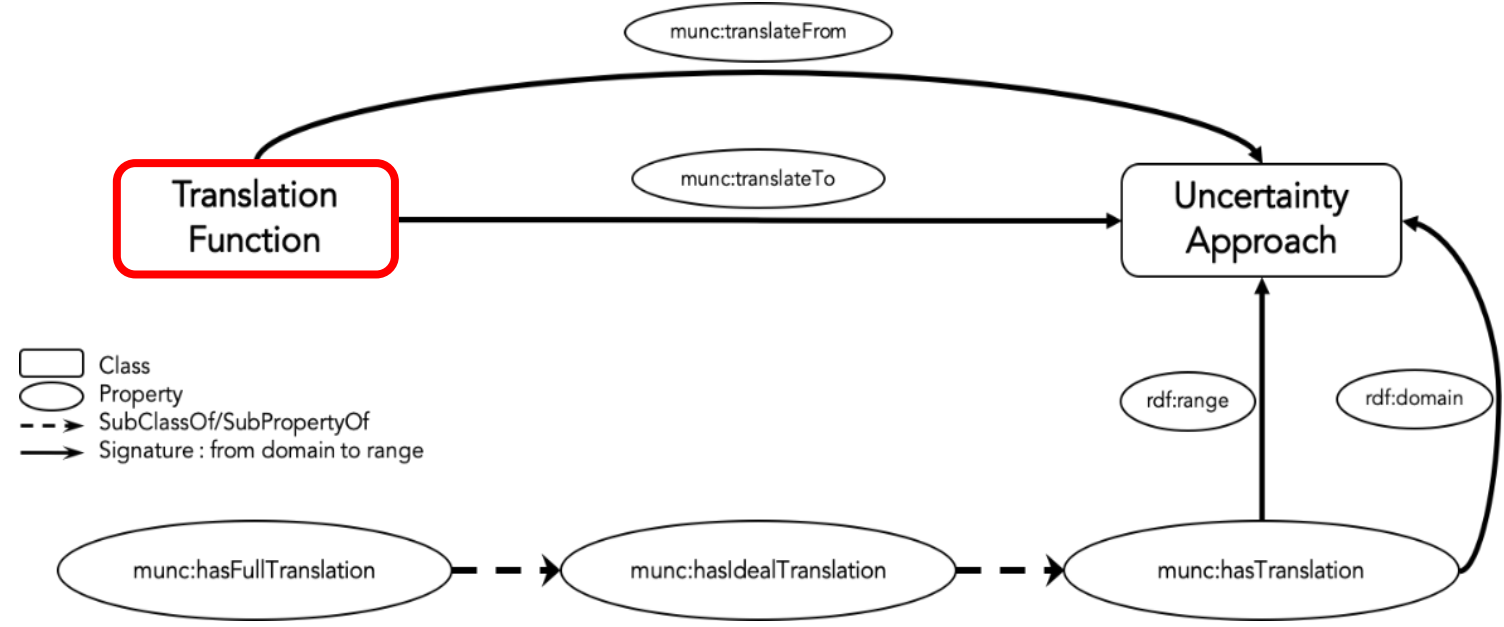


```

function prob:multiplyProbability(?s1, ?s2, ?c) {
  let(?v1 = munc:getMeta(?s1, prob:Probability)){
    if(prob:verifyIndependent(?s1, ?s2) == true)
      ?v2 = munc:getMeta(?s2, prob:Probability, xt:list(?s1, ?c))
      return (?v1 * ?v2)
    } else {
      ?v2 = munc:getMeta(?s2, prob:Probability)
      return (?v1 * ?v2)
    }
  }
}
  
```

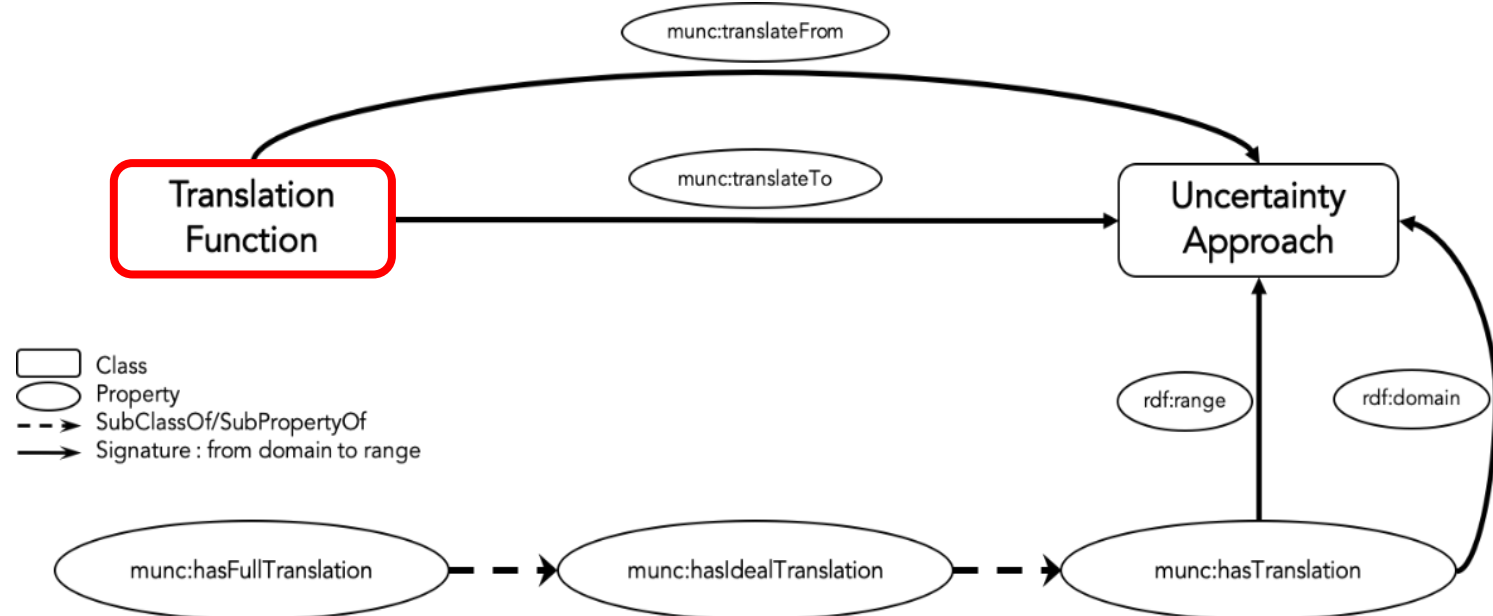


UNCERTAINTY [Djebri et al 2019]



Can uncertainty be translated and negotiated?

UNCERTAINTY [Djebri et al 2019]



Can uncertainty be translated and negotiated?

- Specify uncertainty in parameter linked to the format
 - GET /some/resource HTTP/1.1
 Accept: text/turtle;uncertainty="http://example.com/Probability";q=0.8,
 text/turtle;uncertainty="http://example.com/Possibility";q=0.2;

Content Negotiation by Profile
 W3C Working Draft 30 April 2019

This version:
<https://www.w3.org/TR/2019/WD-dx-prof-conneg-20190430/>

Latest published version:
<https://www.w3.org/TR/dx-prof-conneg/>

Latest editor's draft:
<https://w3c.github.io/dxwg/conneg-by-ap/>

Test suite:
<https://github.com/CSIRO-enviro-informatics/prof-conneg-test-suite>

Implementation report:
<https://github.com/CSIRO-enviro-informatics/prof-conneg-implementations>

Previous version:
<https://www.w3.org/TR/2018/WD-dx-prof-conneg-20181218/>

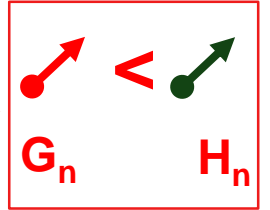
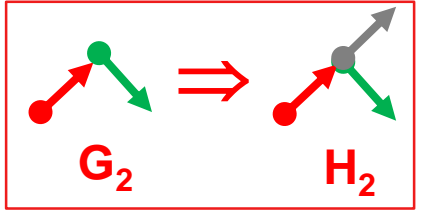
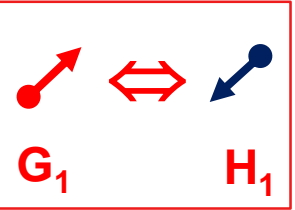
Latest Recommendation:
<https://www.w3.org/TR/2018/WD-dx-prof-conneg-20181218/>

- Use uncertainty as a profile : [prof-Conneg](#)
 - GET /some/resource HTTP/1.1
 Accept: text/turtle;q=0.8;profile="prob:Probability",
 text/turtle;q=0.2;profile="poss:Possibility"
 - HEAD /some/resource HTTP/1.1
 Accept: text/turtle;q=0.9,application/rdf+xml;q=0.5
 Link: <http://example.com/Probability>; rel="profile" (RFC 6906)
 - GET /some/resource HTTP/1.1
 Accept: text/turtle
 Prefer: profile="prob:Probability" (RFC 7240)

Web for classical AI tasks



AI for classical Web tasks



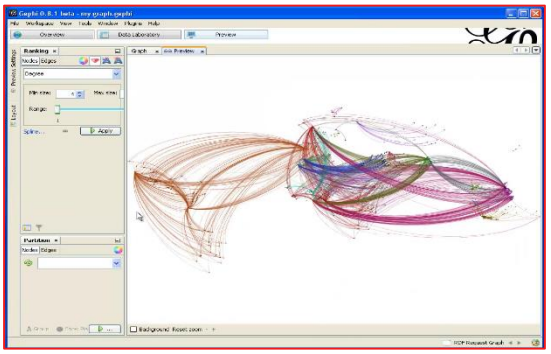
&

&

[Corby, Faron-Zucker et al.]

abstract graph machine

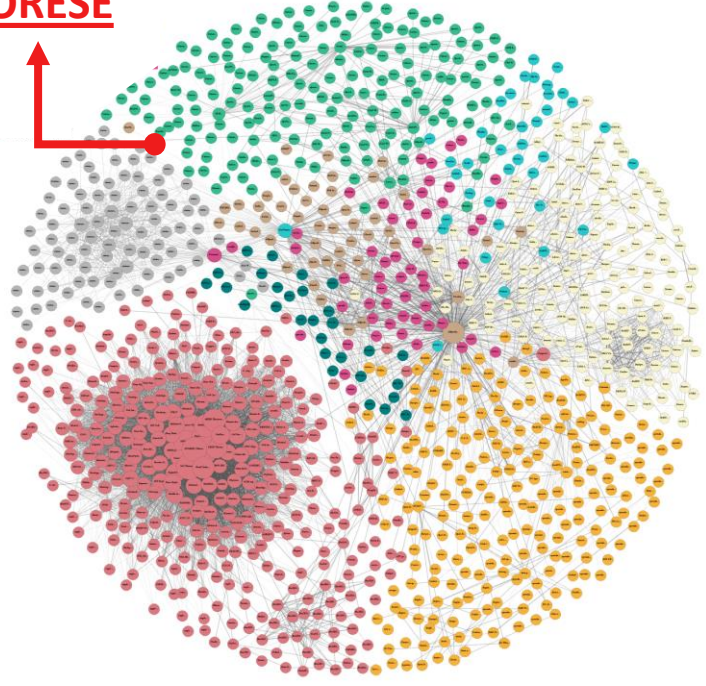
STTL

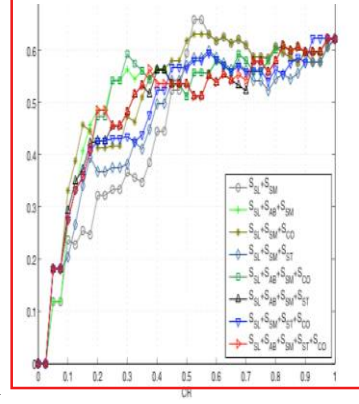
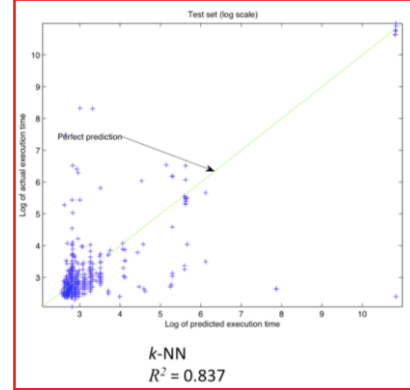
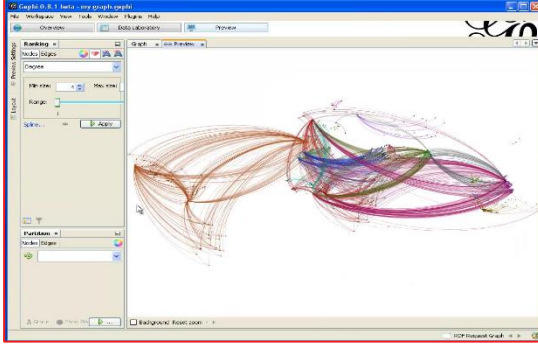
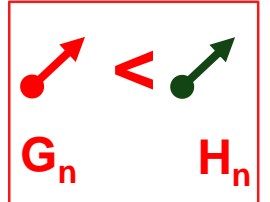
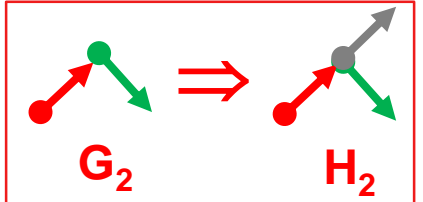
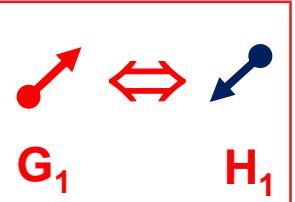


QUERY & INFER

- graph rules and queries
- deontic reasoning
- induction

CORESE





[Corby, Faron-Zucker et al.]

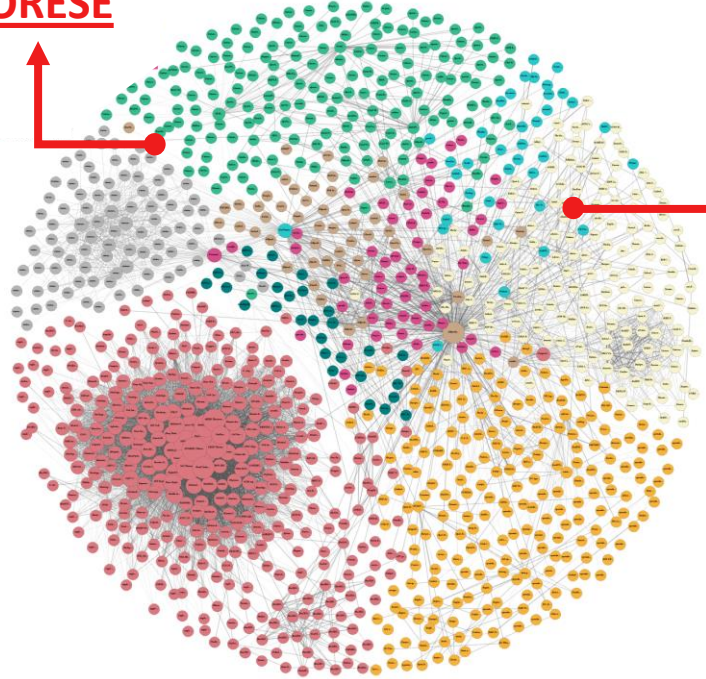
abstract graph machine

CORESE

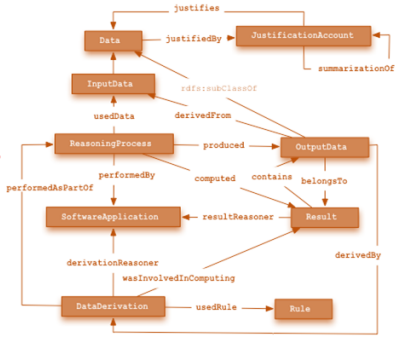
RATIO4TA

QUERY & INFER

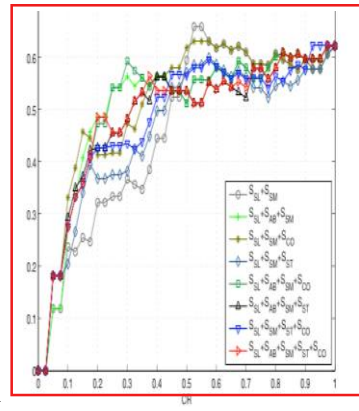
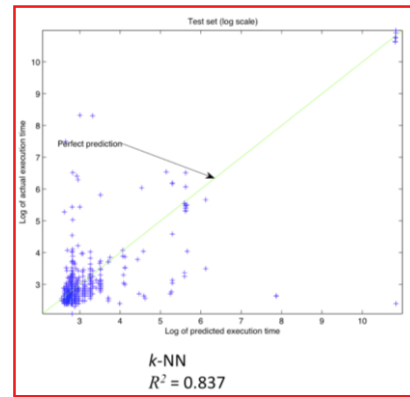
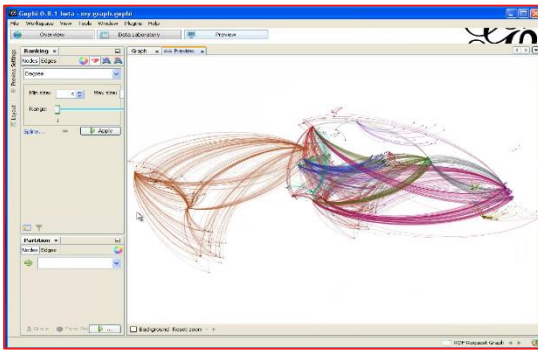
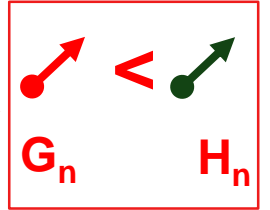
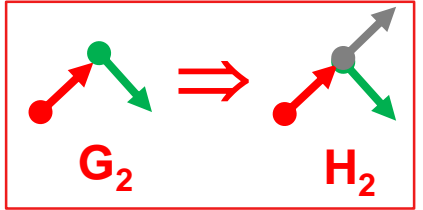
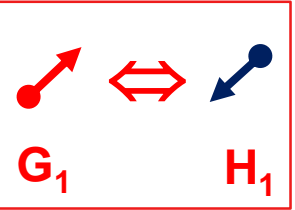
- graph rules and queries
- deontic reasoning
- induction



predict & explain



[Hasan et al.]



[Corby, Faron-Zucker et al.]

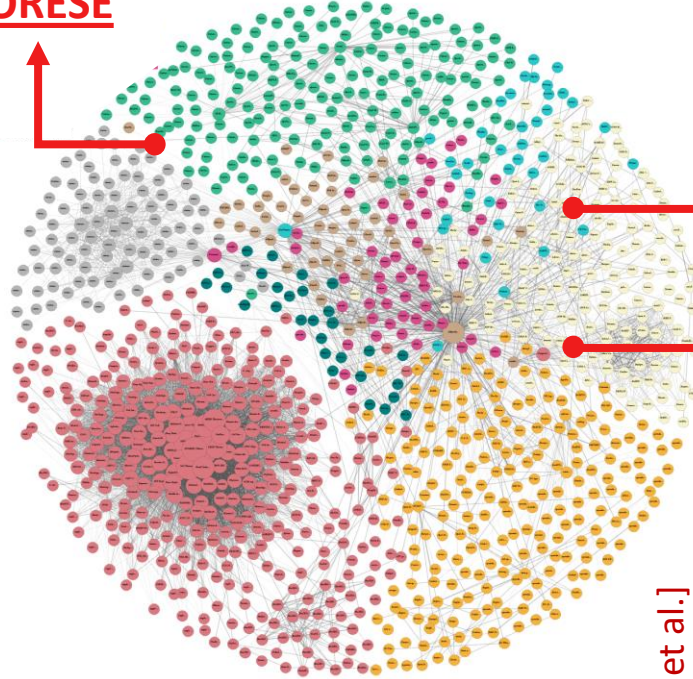
abstract graph machine
STTL

QUERY & INFER

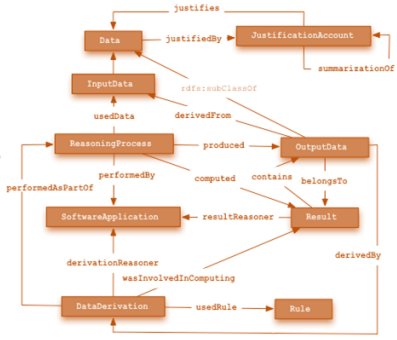
- graph rules and queries
- deontic reasoning
- induction

CORESE

RATIO4TA



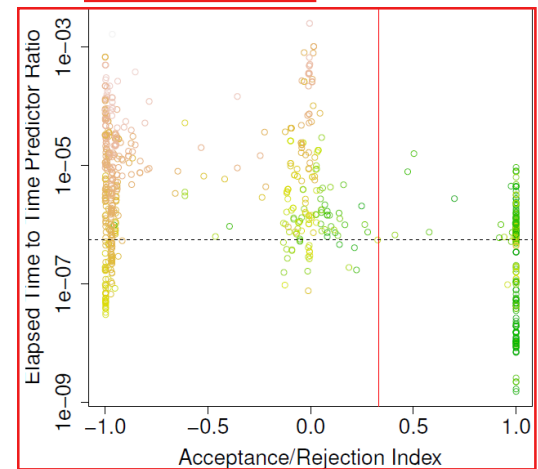
predict & explain



[Hasan et al.]

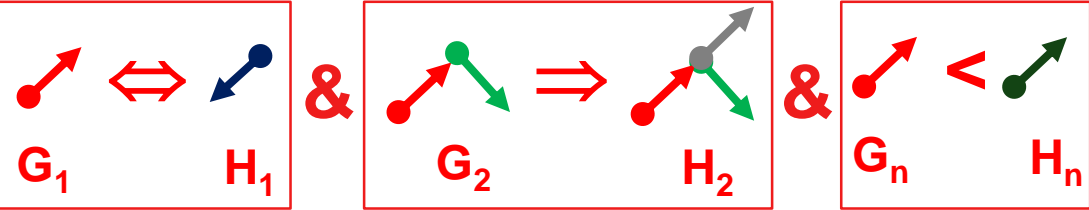
INDUCTION

[Tettamanzi et al.]



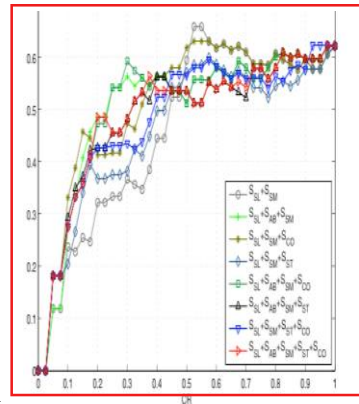
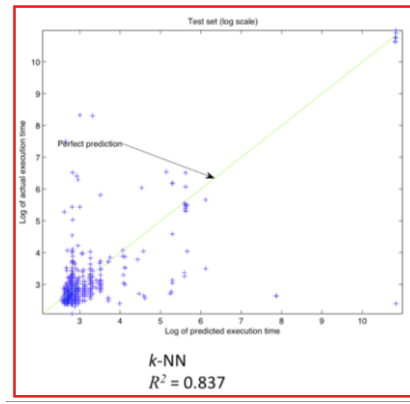
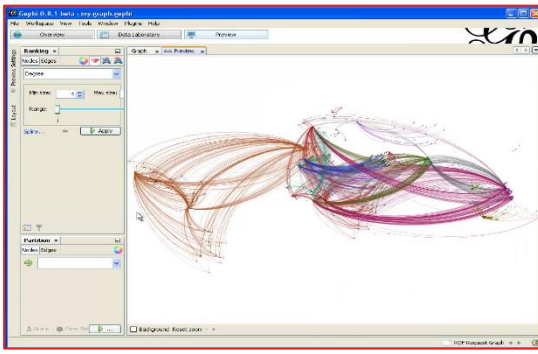
find missing knowledge

$$\phi = \text{SubClassOf}(\text{dbo:LaunchPad} \text{ dbo:Infrastructure})$$



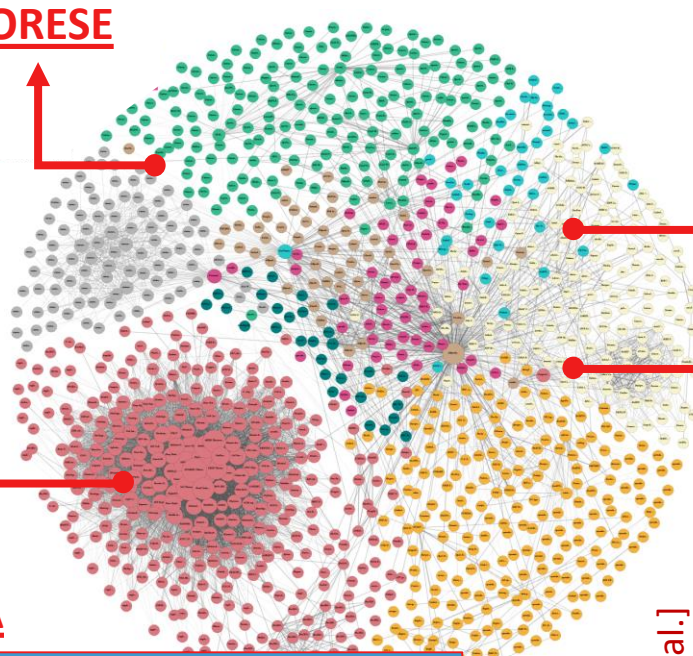
[Corby, Faron-Zucker et al.]

abstract graph machine
STTL



QUERY & INFER

- graph rules and queries
- deontic reasoning
- induction



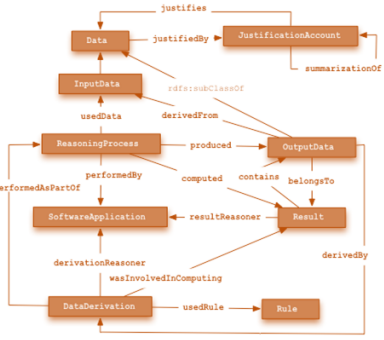
CORESE

RATIO4TA

LICENTIA

INDUCTION

predict & explain



[Hasan et al.]

License your Data

Licentia is a suite of services to support you in looking for a suitable license for your data. Select the service you need from the list below!

Find a license for your Data 	Check if a license is compatible with your needs 	Visualize & download licenses
----------------------------------	--	-----------------------------------

deontic reasoning, license compatibility and composition

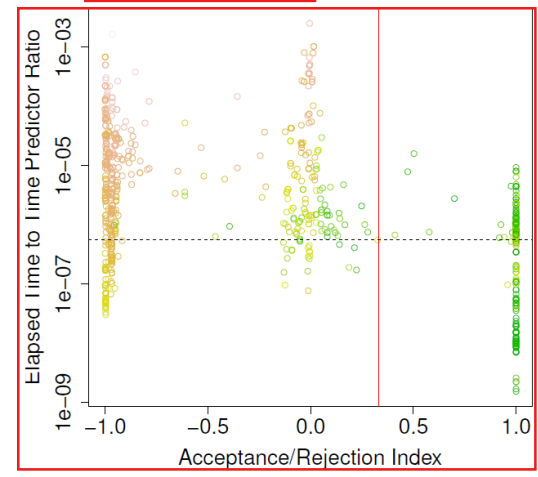
[Villata et al.]

$$L = \{l_1, l_2\}$$

$$R^{O1} = \{r_1 := \Rightarrow_0^l Attribution, \quad r_2 := \rightsquigarrow_0^l Commercial\}$$

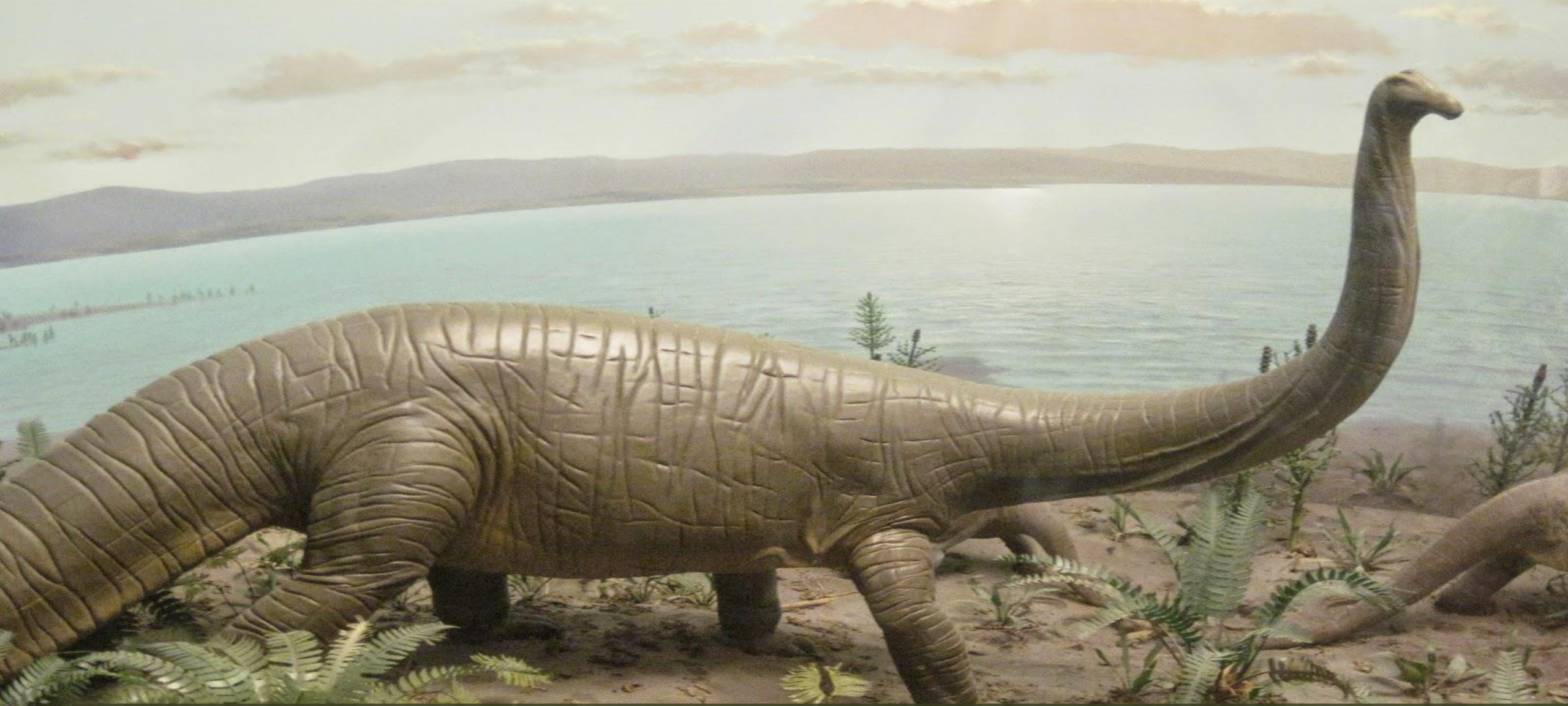
$$R^{O2} = \{r_3 := \Rightarrow_0^l \sim Commercial, \quad r_4 := \Rightarrow_0^l ShareAlike, \quad r_5 := \rightsquigarrow_0^l Derivative\}$$

[Tettamanzi et al.]



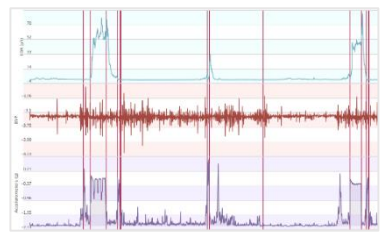
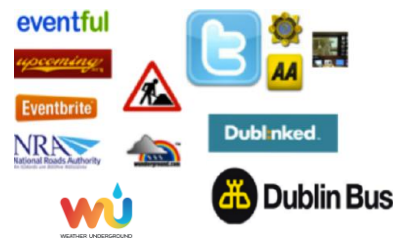
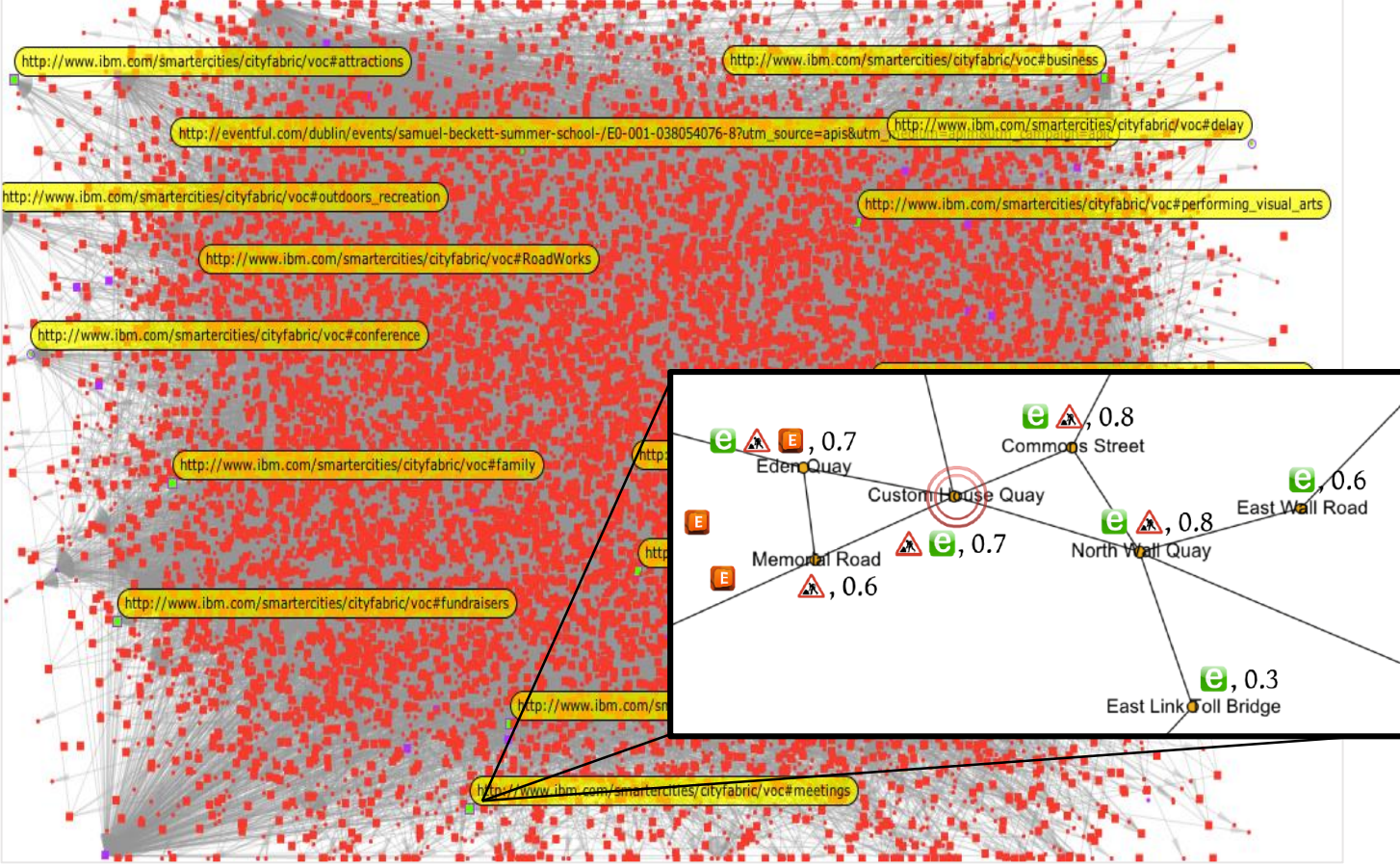
find missing knowledge

$$\phi = \text{SubClassOf}(\text{dbo:LaunchPad} \text{ dbo:Infrastructure})$$



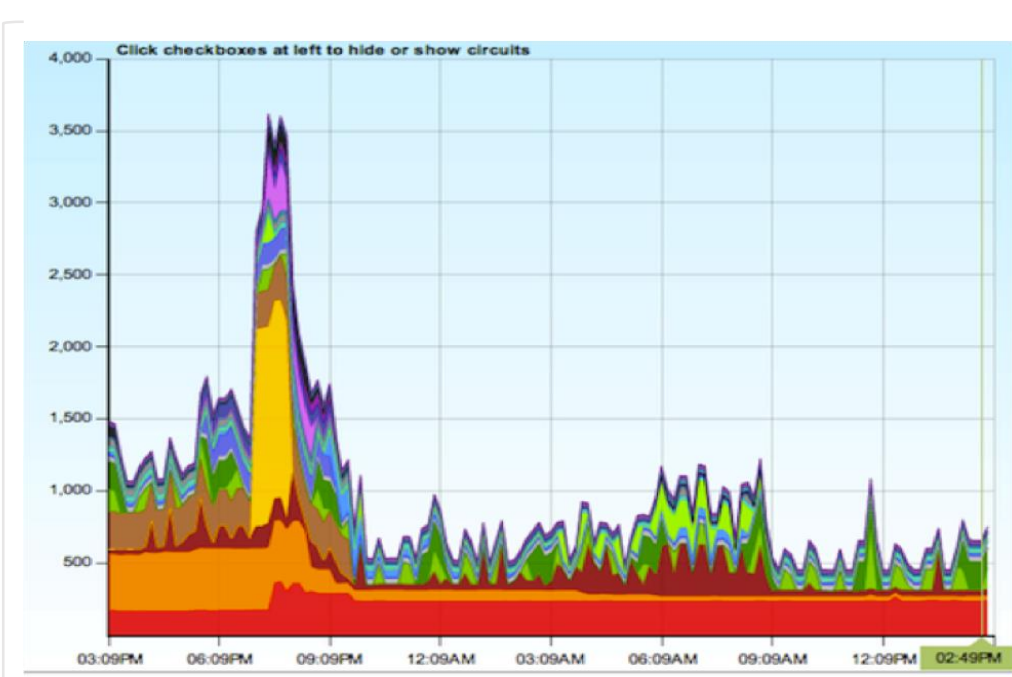
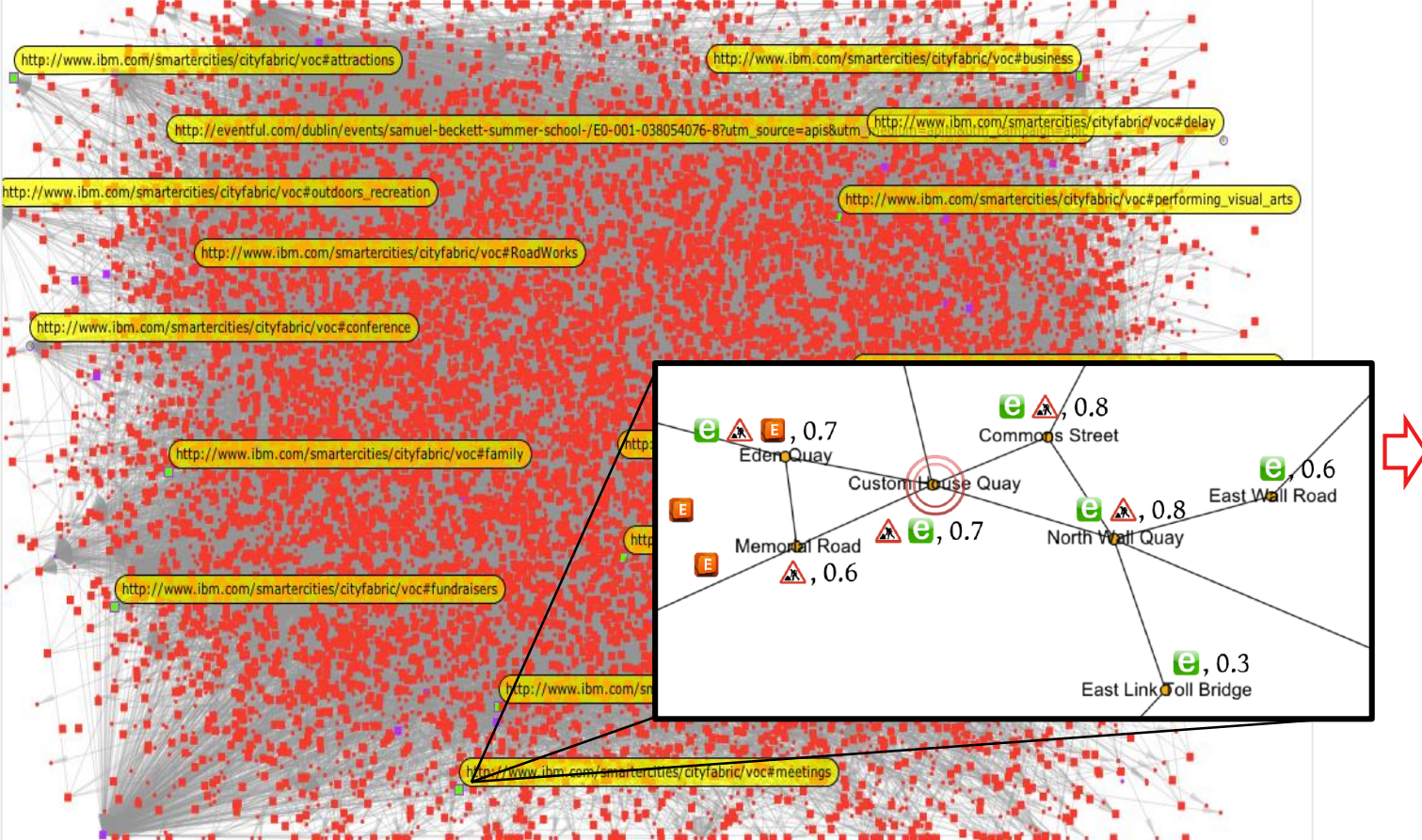
the Web: this place where invisible *brontobytes* graze furiously

10^{27} bytes



Smarter Cities – IBM Dublin

[Lécué, 2015]



eventful
 Eventbrite
 NRA National Roads Authority
 WU WEATHER UNDERGROUND
 Dublin Bus
 Twitter
 AA
 DBpedia

IBM Research - STAR CITY: Semantic Traffic Analytics and Reasoning for CITY

Close Weather Context

Close Selected Context

Open Control Panel

Events: Road Incidents, Events: Road Works, Events: Social

Diagnosis: Road Incidents, Diagnosis: Road Works, Diagnosis: Social

obstruction:40.625, closures:9.090909, music:65.35088

accident:33.53204, luas:26.528385, music:28.47962

Smarter Cities – IBM Dublin

[Lécué, 2015]

Sexe	Date	Cause	CISP2	...	History	Observations
H	25/04/2012	vaccin-antitétanique	A44	...	Appendicite	EN CP - Bon état général - auscult pulm libre; bdc rég sans souffle - tympan ok-

Element	Number
Patients	55 823
Consultations	364 684
Past medical history	187 290
Biometric data	293 908
Semiotics	250 669
Diagnosis	117 442
Row of prescribed drugs	847 422
Symptoms	23 488
Health care procedures	11 850
Additional examination	871 590
Paramedical prescription	17 222
Observations/notes	56 143

PREDICT HOSPITALIZATION

[Gazzotti, Faron et al. 2017]

- Physician's records classification in order to predict hospitalization

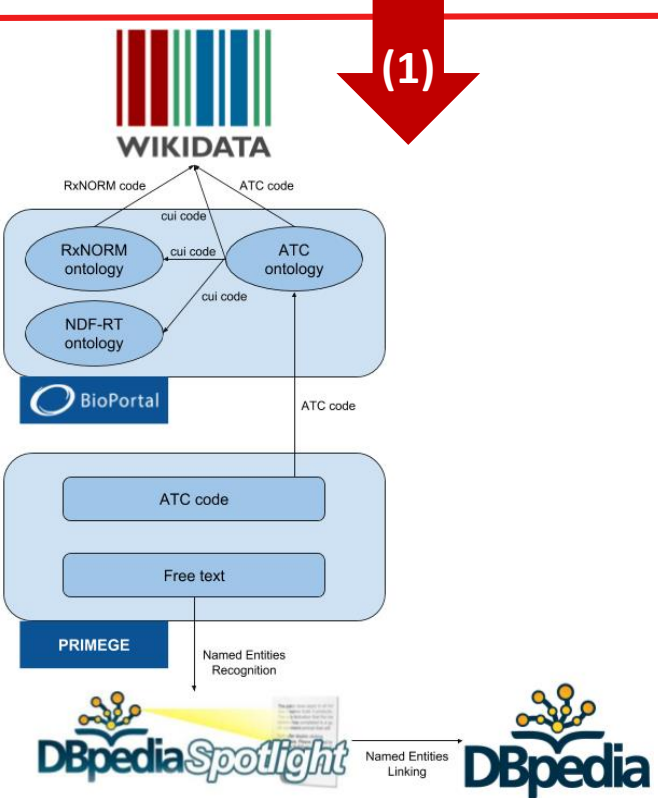
Sexe	Date	Cause	CISP2	...	History	Observations
H	25/04/2012	vaccin-antitétanique	A44	...	Appendicite	EN CP - Bon état général - auscult pulm libre; bdc rég sans souffle - tympan ok-

Element	Number
Patients	55 823
Consultations	364 684
Past medical history	187 290
Biometric data	293 908
Semiotics	250 669
Diagnosis	117 442
Row of prescribed drugs	847 422
Symptoms	23 488
Health care procedures	11 850
Additional examination	871 590
Paramedical prescription	17 222
Observations/notes	56 143

PREDICT HOSPITALIZATION

[Gazzotti, Faron et al. 2017]

- Physician's records classification in order to predict hospitalization
- Augment data with structured knowledge



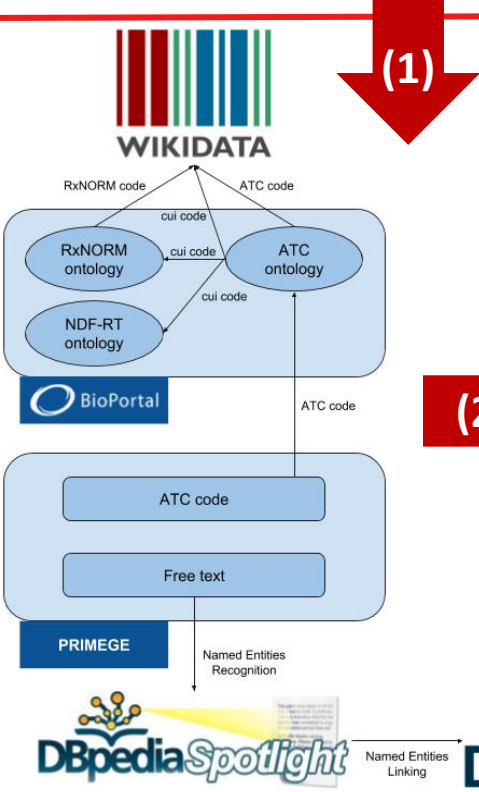
Sexe	Date	Cause	CISP2	...	History	Observations
H	25/04/2012	vaccin-antitétanique	A44	...	Appendicite	EN CP - Bon état général - auscult pulm libre; bdc rég sans souffle - tympan ok-

Element	Number
Patients	55 823
Consultations	364 684
Past medical history	187 290
Biometric data	293 908
Semiotics	250 669
Diagnosis	117 442
Row of prescribed drugs	847 422
Symptoms	23 488
Health care procedures	11 850
Additional examination	871 590
Paramedical prescription	17 222
Observations/notes	56 143

PREDICT HOSPITALIZATION

[Gazzotti, Faron et al. 2017]

- Physician's records classification in order to predict hospitalization
- Augment data with structured knowledge and study impact on different prediction methods
- Study enrichment and features impact and combinations on different ML methods



Représentation	SGD	SVC	RF	Log	Moyenne
<i>référence</i>	0.7877	0.8270	0.8533	0.8491	0.8293
+t	0.8054	0.8239	0.8522	0.8545	0.8340
+s	0.7889	0.8221	0.8522	0.8485	0.8279
+s*	0.7985	0.8339	0.8449	0.8514	0.8322
+c ₁	0.7859	0.8235	0.8433	0.8453	0.8245
+c ₁₋₂	0.7871	0.8254	0.8480	0.8510	0.8279
+c ₂	0.8209	0.8348	0.8522	0.8505	0.8396
+d _{prevent}	0.7796	0.8254	0.8506	0.8479	0.8259
+d _{treat}	0.7925	0.8338	0.8472	0.8481	0.8304
+d _{C1}	0.8108	0.8281	0.8498	0.8460	0.8337
+wa	0.8065	0.8223	0.8468	0.8545	0.8325
+wi	0.8137	0.8149	0.8484	0.8501	0.8318
+wm	0.7991	0.8221	0.8453	0.8458	0.8281
+t + s + c ₂ + wa + wi	0.7782	0.8258	0.8486	0.8547	0.8268
+t + s* + c ₂ + wa + wi	0.7801	0.8239	0.8494	0.8543	0.8269
+t + c ₂ + wa + wi	0.7707	0.8140	0.8531	0.8571	0.8237

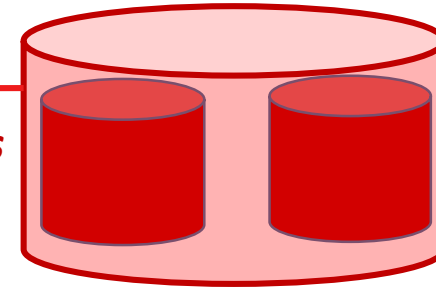
MonaLIA

[Bobasheva et al. 2017]

- reason & query on RDF metadata to build balanced, unambiguous, labelled training sets.

Joconde database from French museums

350 000 *images*
of artworks



RDF *metadata* based
on external thesauri

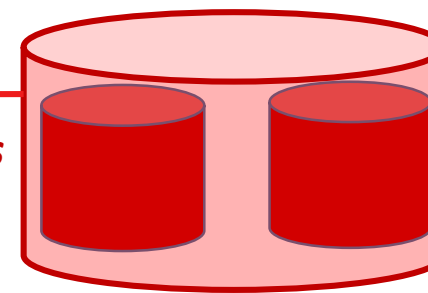
Joconde database from French museums

MonaLIA

[Bobasheva et al. 2017]

- reason & query on RDF metadata to build balanced, unambiguous, labelled training sets.
- transfer learning & CNN classifiers on targeted categories (topics, techniques, etc.)

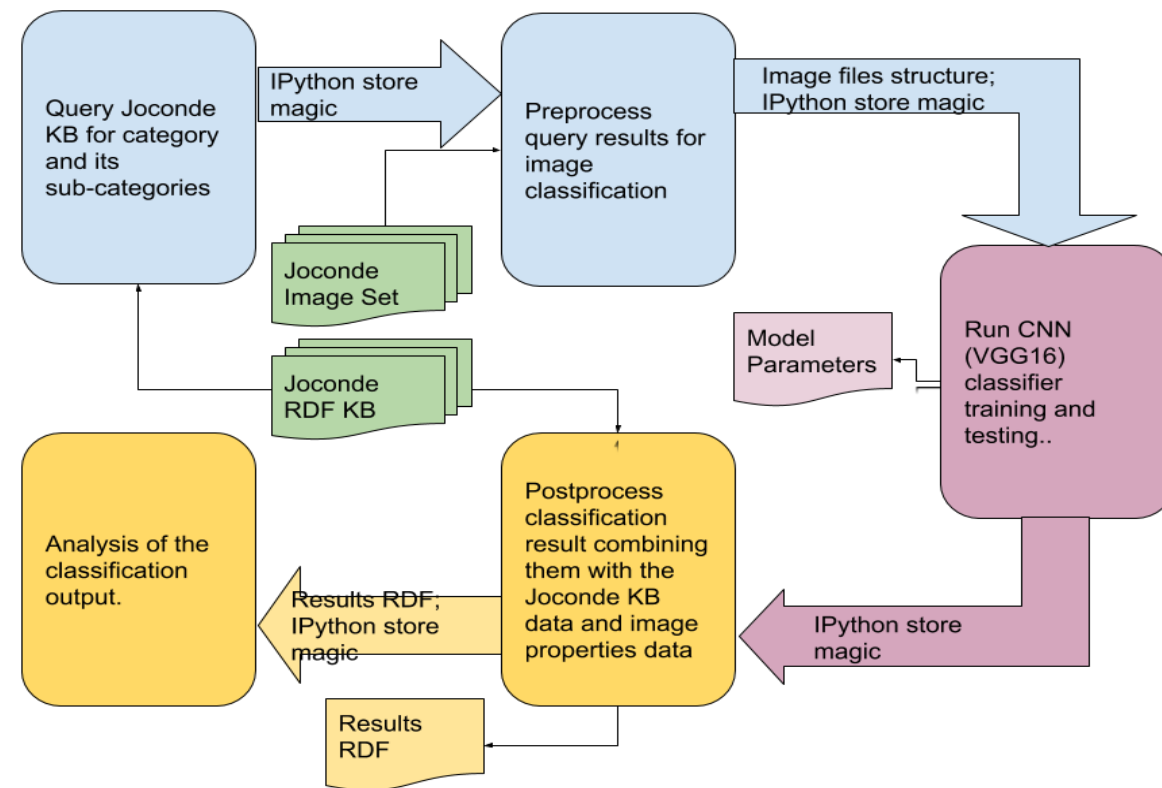
350 000 *images*
of artworks



RDF *metadata* based
on external thesauri



MonaLIA Data Pipeline

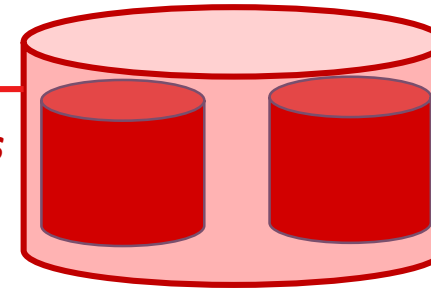


MonaLIA

[Bobasheva et al. 2017]

- reason & query on RDF metadata to build balanced, unambiguous, labelled training sets.
- transfer learning & CNN classifiers on targeted categories (topics, techniques, etc.)
- reason & query RDF metadata of results to address silence, noise and explain

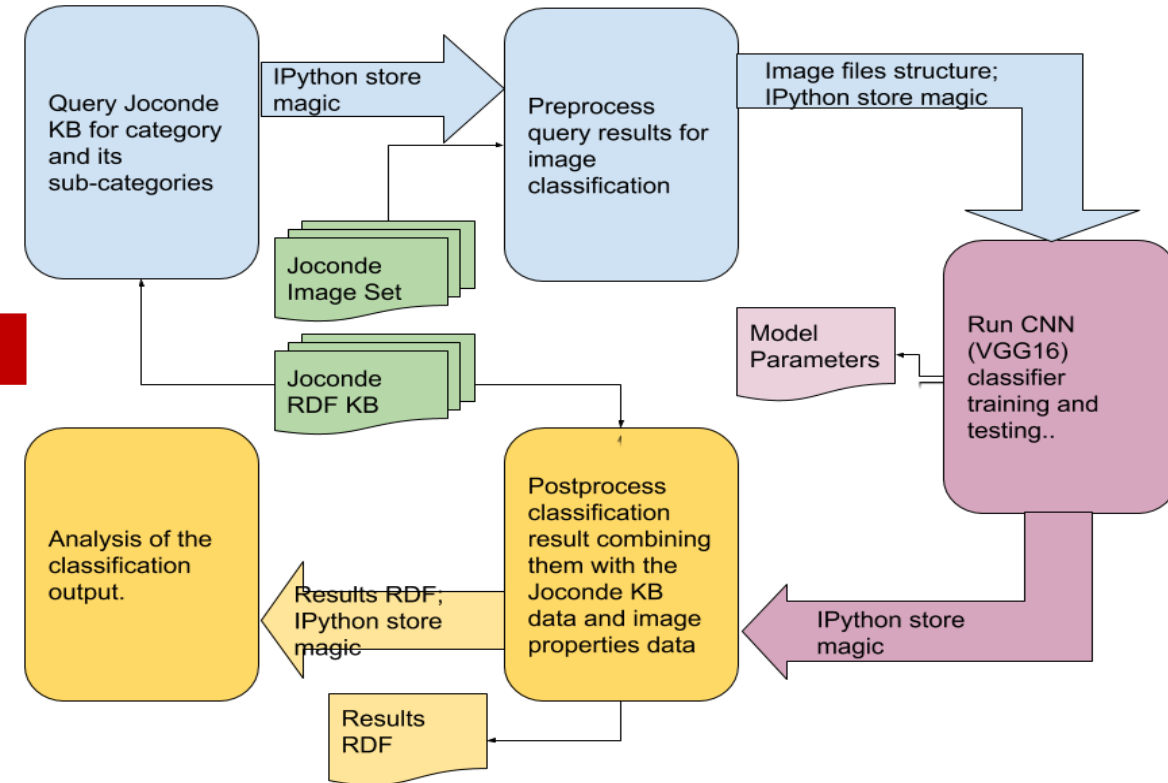
350 000 *images* of artworks



RDF *metadata* based on external thesauri



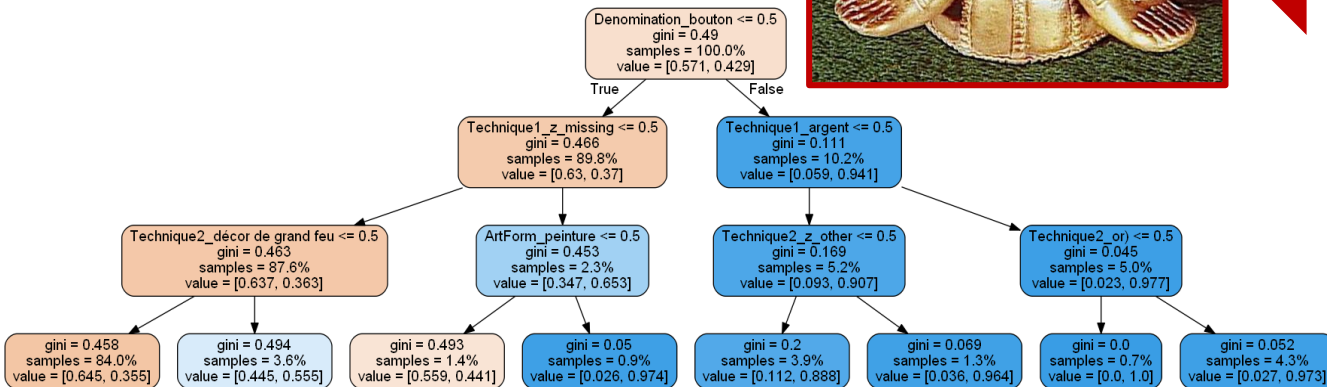
MonaLIA Data Pipeline



animal ✓

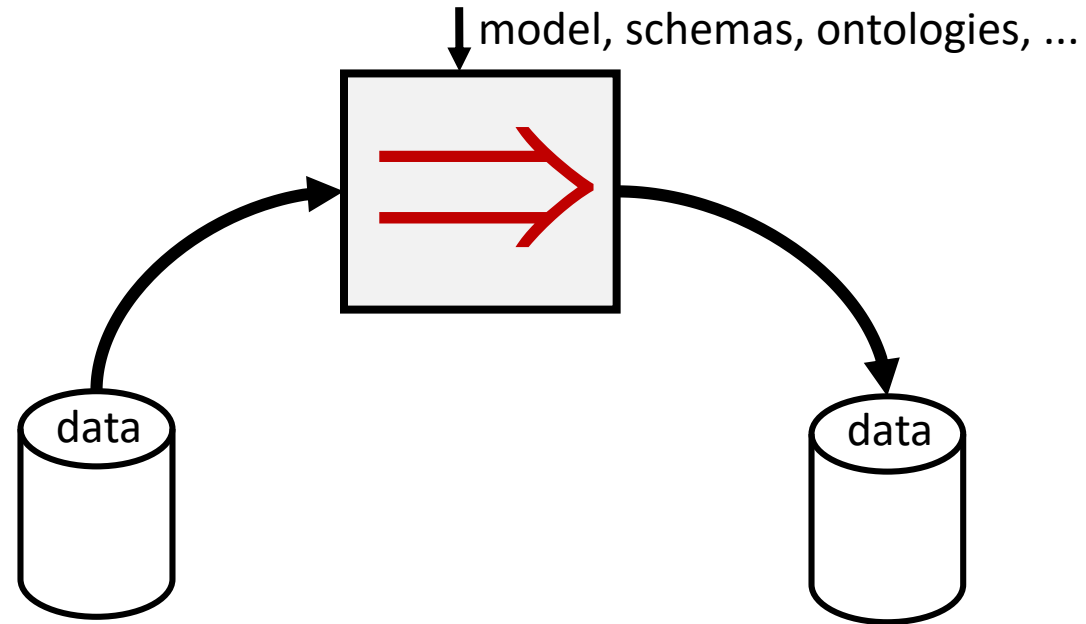
bird ?

painting ✗



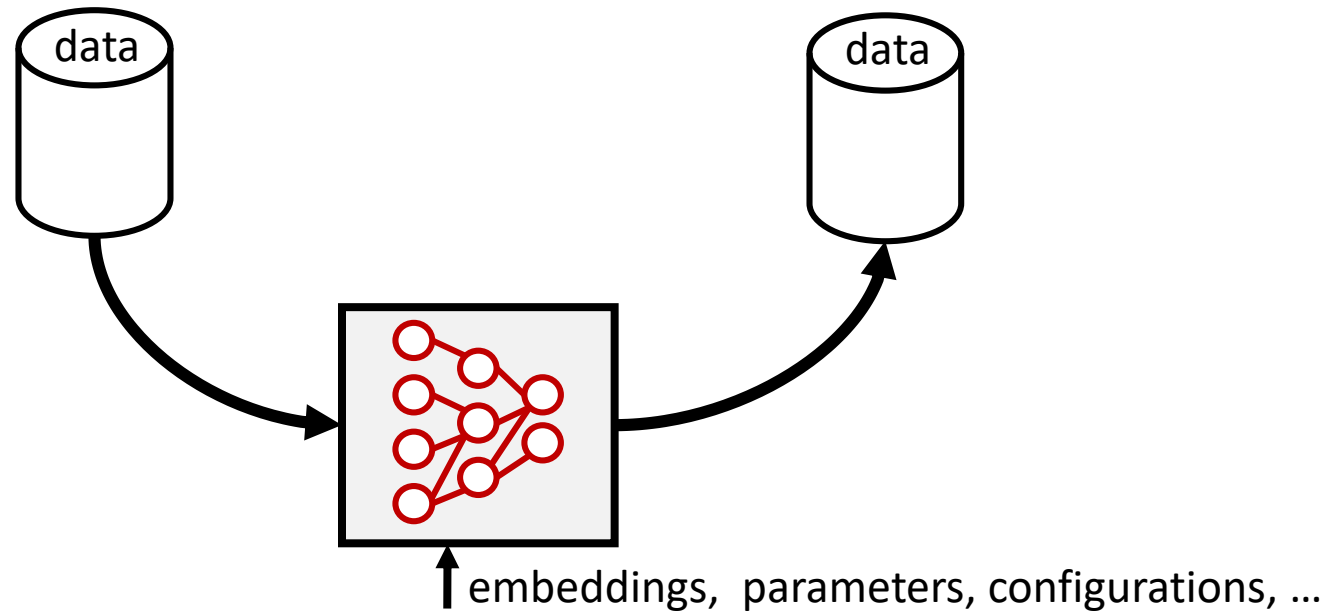
deduce data

15% progress



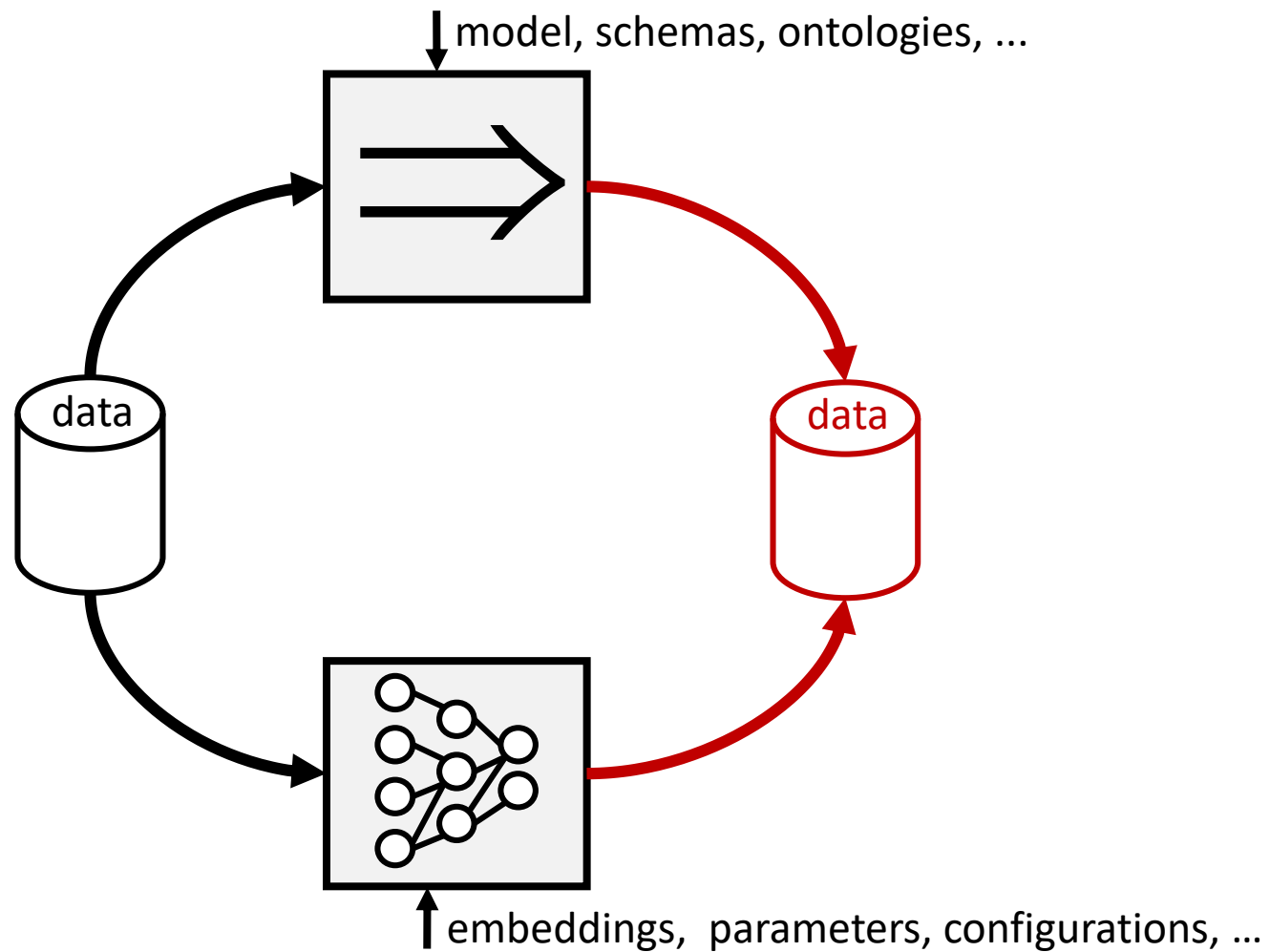
learn data

30% progress



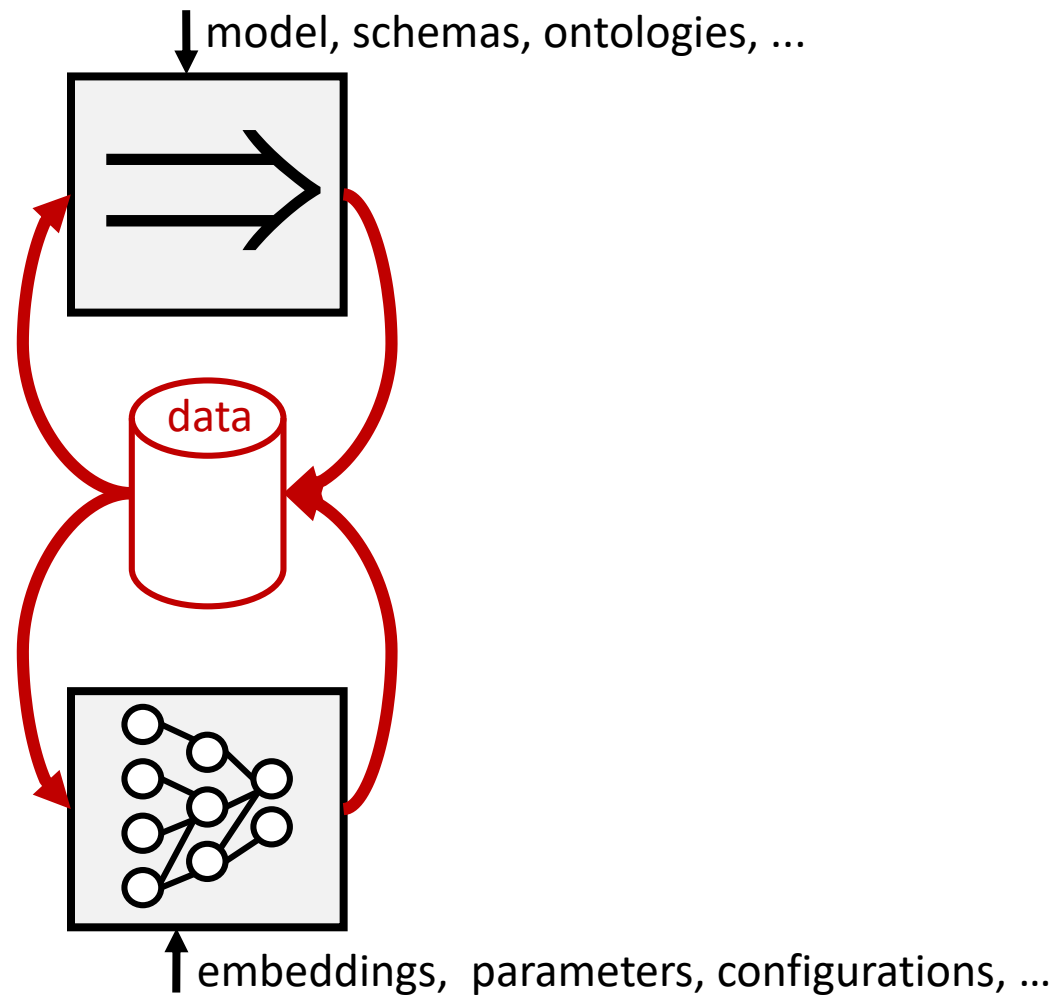
sum intelligence

45% progress



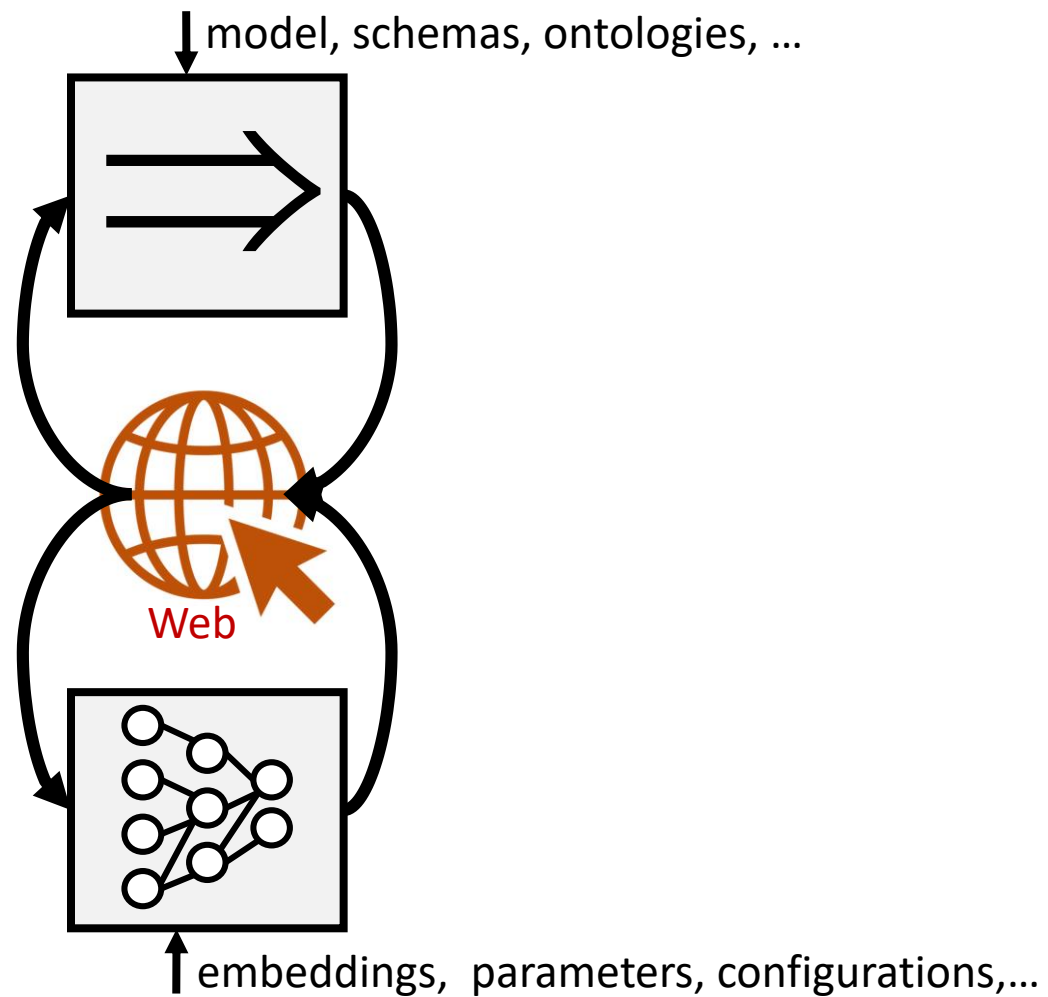
combine intelligence

60% progress



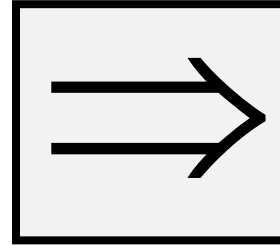
remotely combine

75% progress



deeply combine

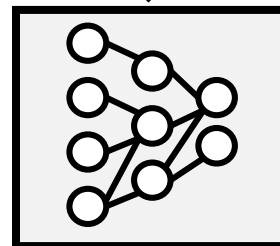
90% progress



↕ data, knowledge, model, schemas, ontologies, ...

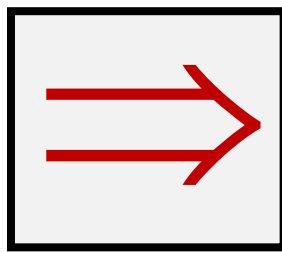


↕ data, knowledge, embeddings, parameters, configurations,...



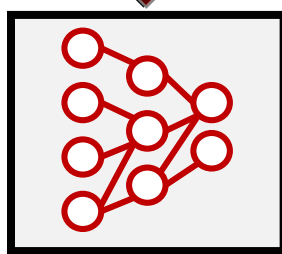
combining AIs on the Web

100% progress



data, knowledge, model, schemas, ontologies, ...

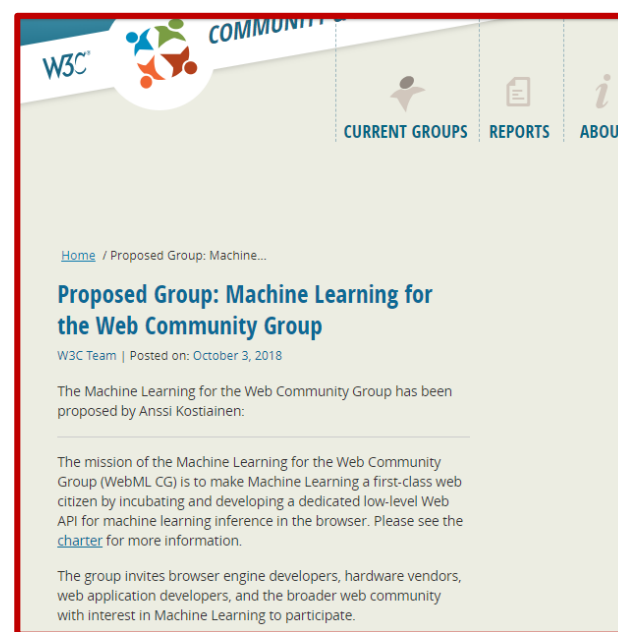
data, knowledge, embeddings, parameters, configurations, ...



WEB EDGE AI

[WebML @ W3C]

- Edge AI directly in the browser
- Web APIs, models, protocols,...



W3C COMMUNITY

CURRENT GROUPS REPORTS ABOUT

[Home](#) / Proposed Group: Machine...

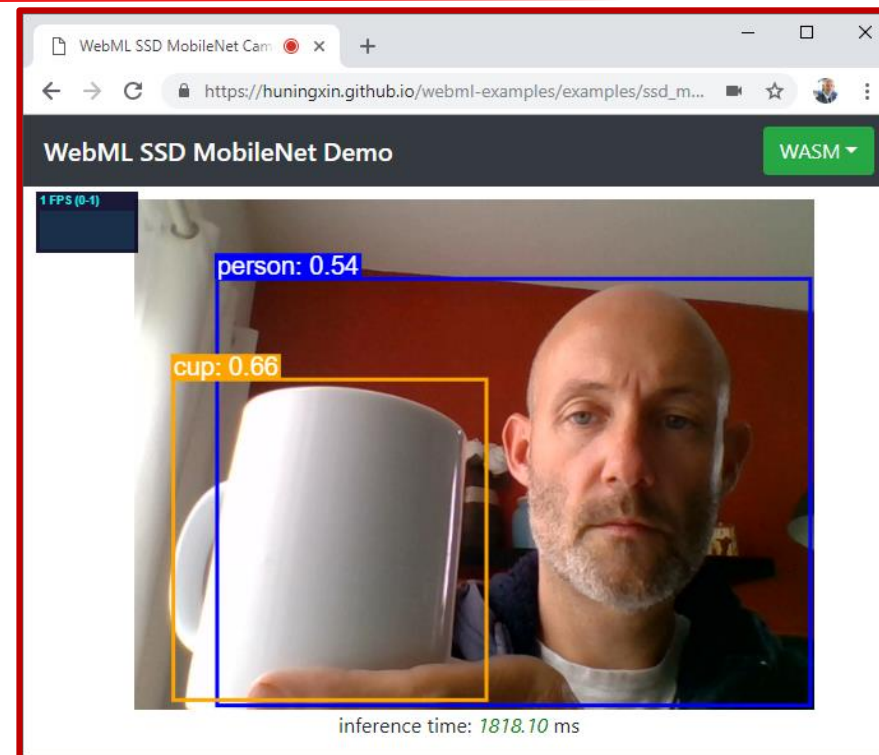
Proposed Group: Machine Learning for the Web Community Group

W3C Team | Posted on: October 3, 2018

The Machine Learning for the Web Community Group has been proposed by Anssi Kostainen:

The mission of the Machine Learning for the Web Community Group (WebML CG) is to make Machine Learning a first-class web citizen by incubating and developing a dedicated low-level Web API for machine learning inference in the browser. Please see the [charter](#) for more information.

The group invites browser engine developers, hardware vendors, web application developers, and the broader web community with interest in Machine Learning to participate.



WebML SSD MobileNet Cam x +

https://huningxin.github.io/webml-examples/examples/ssd_m...

WebML SSD MobileNet Demo

WASM

1 FPS (0-1)

person: 0.54

cup: 0.66

inference time: 1818.10 ms

WebML Examples

SSD MobileNet Image Demo

SSD MobileNet Camera Demo

MobileNet Image Demo

MobileNet Camera Demo

SqueezeNet Image Demo

SqueezeNet Camera Demo

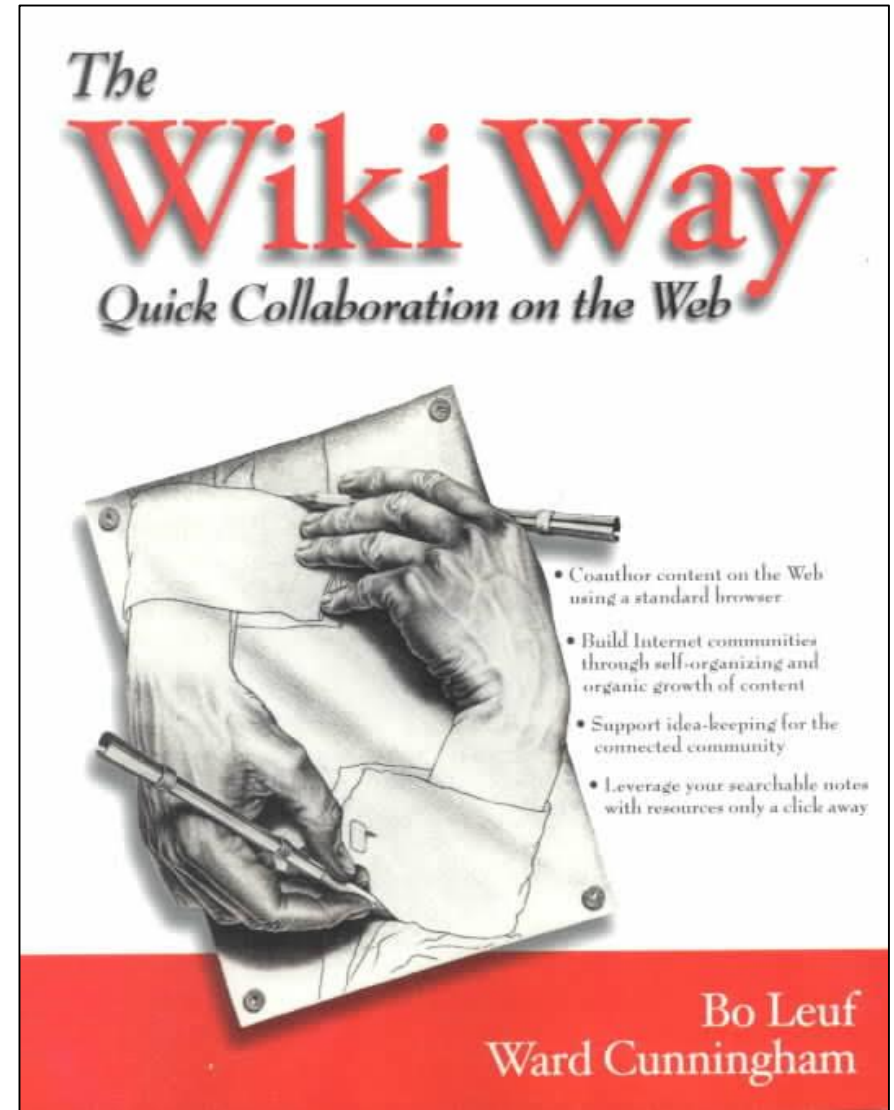
PoseNet Image Demo

PoseNet Camera Demo

[Source code](#)

Web ways applied to AI

e.g. copy-paste based reuse and contribution to create Web AIs



but...



404

... Oops! Something is missing

[HOME](#)

all birds can fly
all penguins are birds
so ...



automated deduction

PIPE : 0.9143



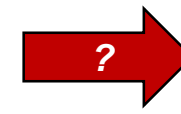
Ceci n'est pas une pipe.

automated Classification



Assertions
(evidence)

e




ϕ

Axiom
(hypothesis)



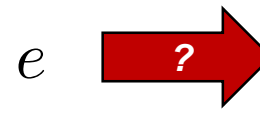
$$\Pr(\phi \mid e) = \frac{\Pr(e \mid \phi) \Pr(\phi)}{\Pr(e \mid \phi) \Pr(\phi) + \Pr(e \mid \neg\phi) \Pr(\neg\phi)}$$

 how to evaluate $\Pr(e \mid \phi)$ in an open-world?

PROBABILITY



Assertions
(evidence) e



ϕ

Axiom
(hypothesis)



$$\Pr(\phi | e) = \frac{\Pr(e | \phi) \Pr(\phi)}{\Pr(e | \phi) \Pr(\phi) + \Pr(e | \neg\phi) \Pr(\neg\phi)}$$

⚠ how to evaluate $\Pr(e | \phi)$ in an open-world?

PROBABILITY

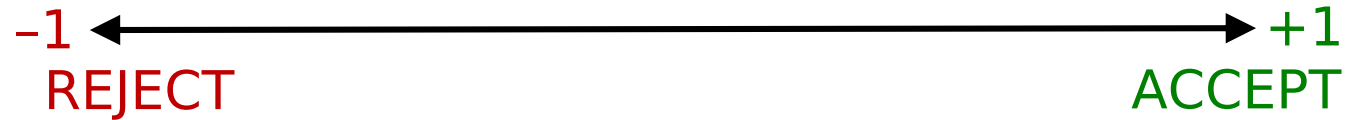
POSSIBILITY [Tettamanzi et al.]

- Possibilistic Axiom Scoring
- Possibility and Necessity Measures

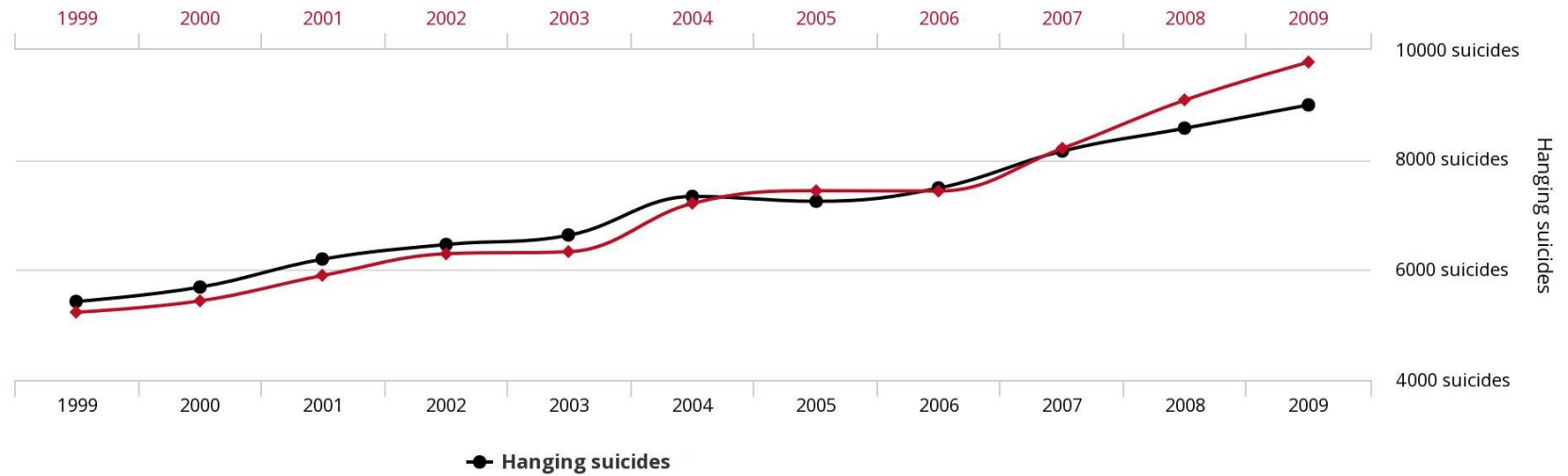
$$\Pi(A) = \max_{\omega \in A} \pi(\omega);$$

$$N(A) = 1 - \Pi(\bar{A}) = \min_{\omega \in \bar{A}} \{1 - \pi(\omega)\}.$$

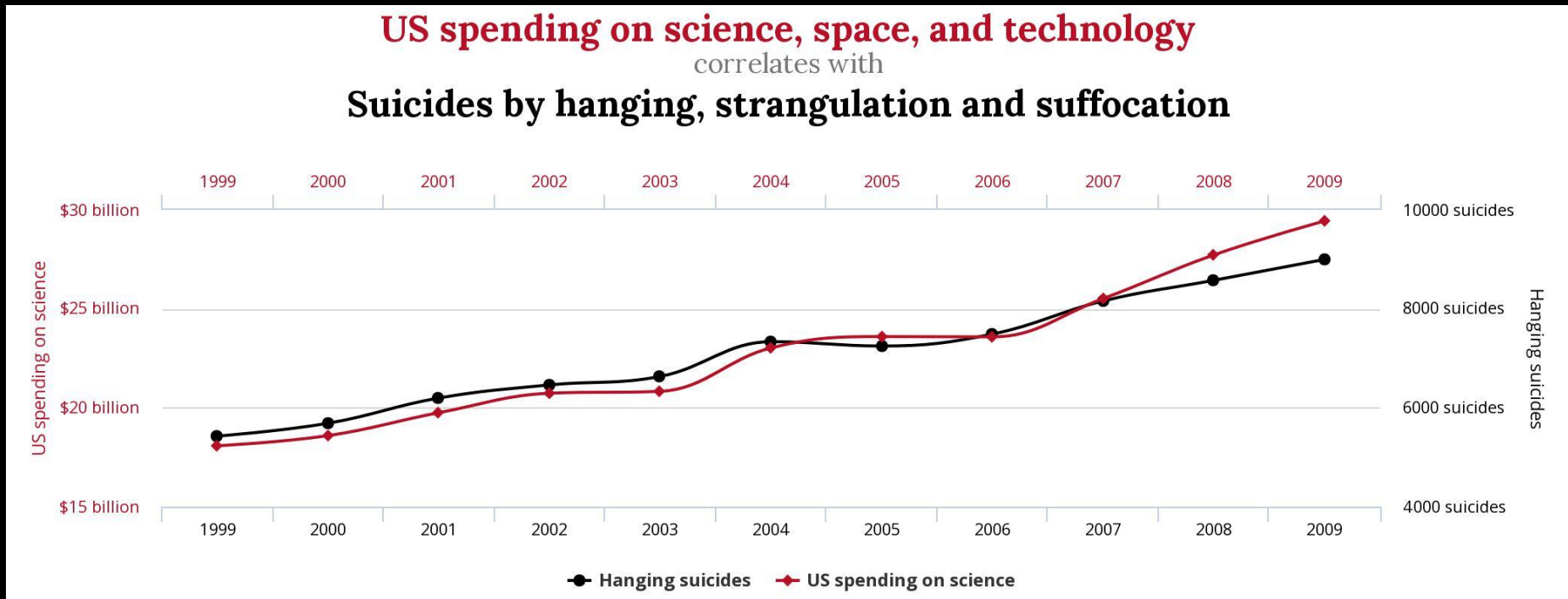
$$\text{ARI}(\phi) = N(\phi) - N(\neg\phi) = N(\phi) + \Pi(\phi) - 1$$



Suicides by hanging, strangulation and suffocation



correlation



<http://tylervigen.com/spurious-correlations>

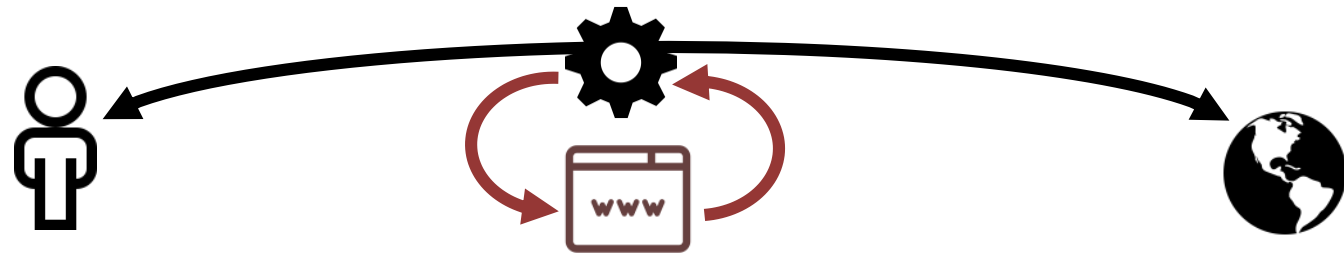
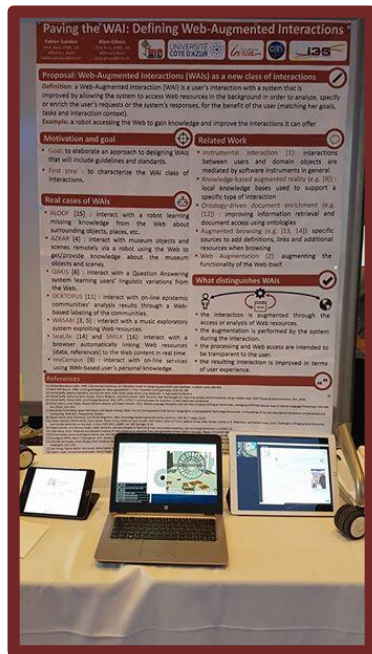
correlation vs. causation

AI and Web augmentation



Web for classical AI tasks

AI for classical Web tasks



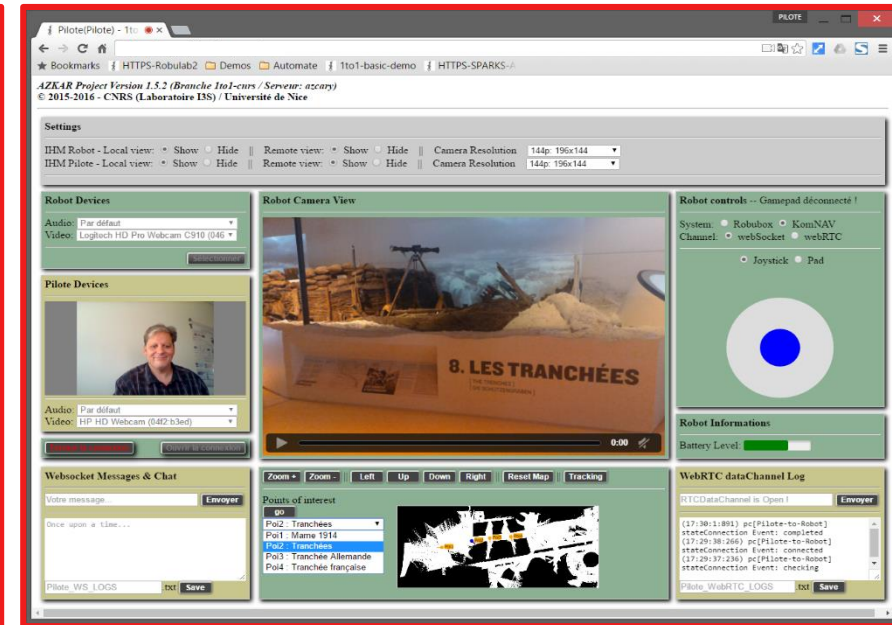
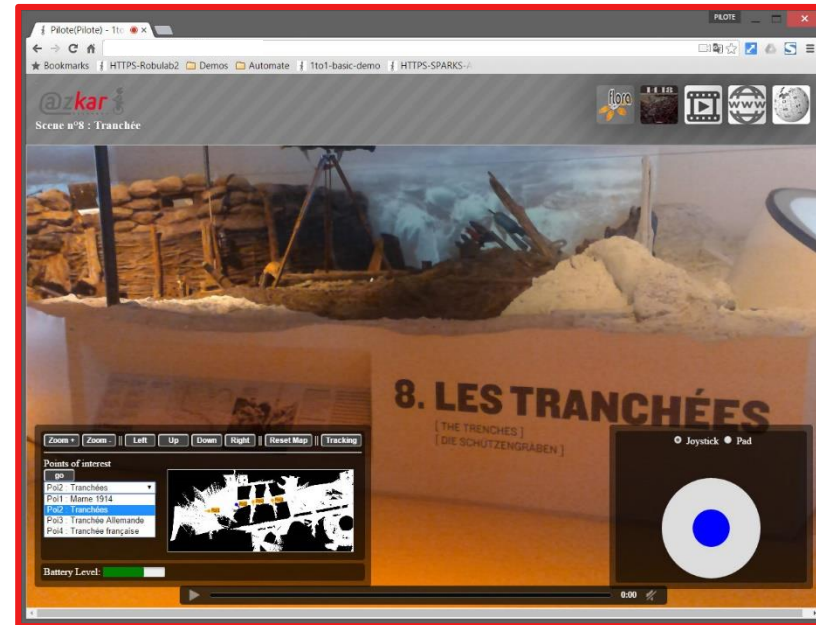
« a *Web-Augmented Interaction (WAI)* is a user's interaction with a system that is improved by allowing the system to access Web resources »



[Gandon, Giboin, WebSci17]

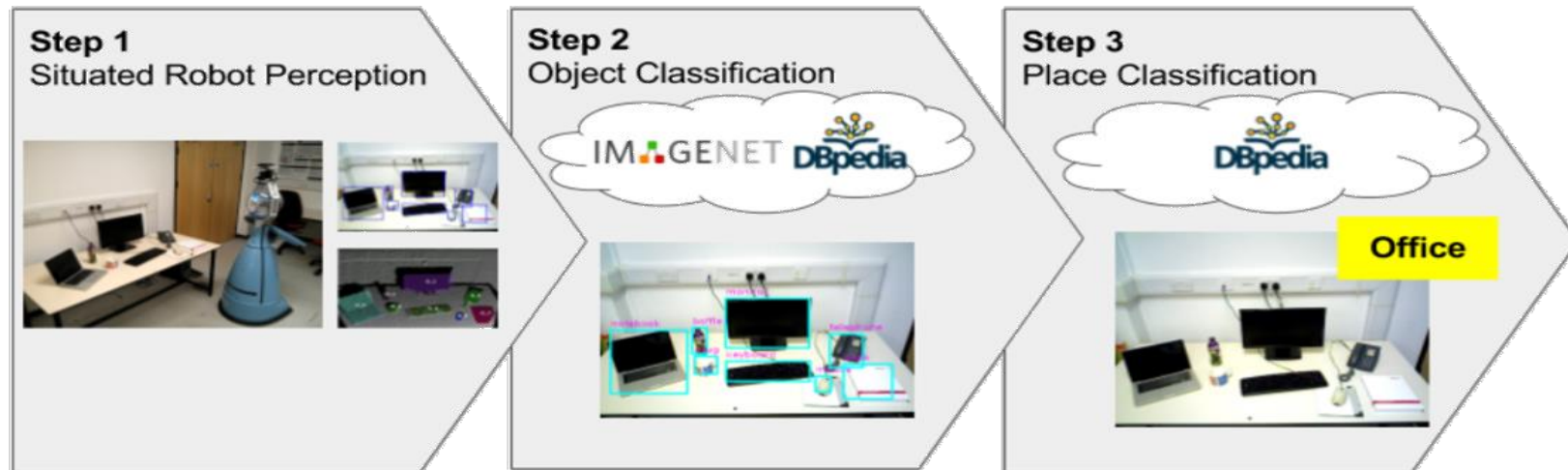
AZKAR [Buffa et al.]

remotely visit and interact with a museum through a robot and via the Web

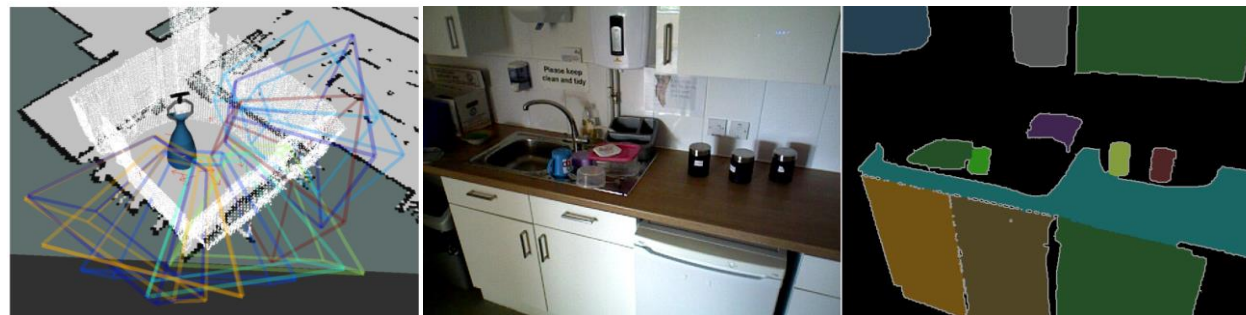




[Cabrio, Basile et al.]

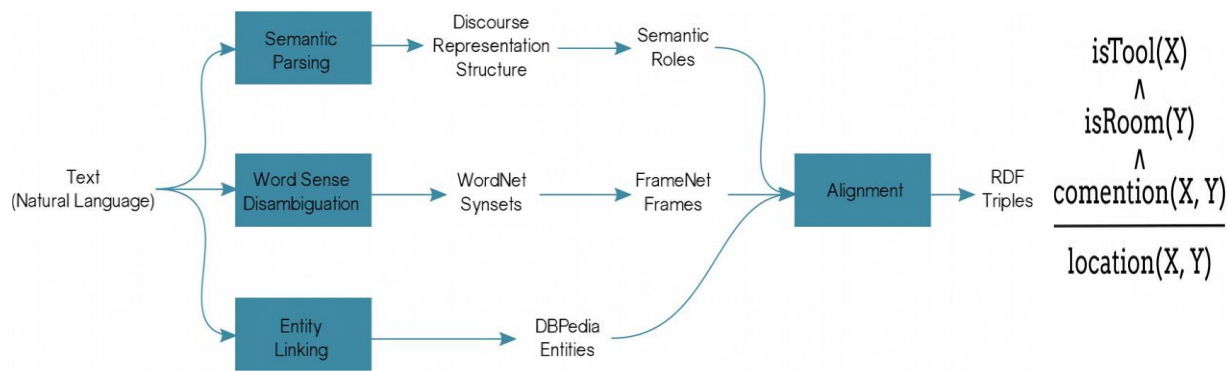


ALOOF: Web and Perception



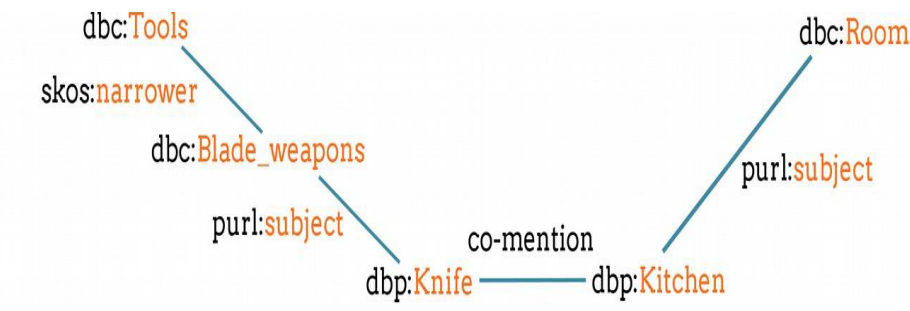
Semantic Web-Mining and Deep Vision for Lifelong Object Discovery (ICRA 2017)

Making Sense of Indoor Spaces using Semantic Web Mining and Situated Robot Perception (AnSWeR 2017)



isTool(X)
 \wedge
 isRoom(Y)
 \wedge
 comention(X, Y)

 location(X, Y)

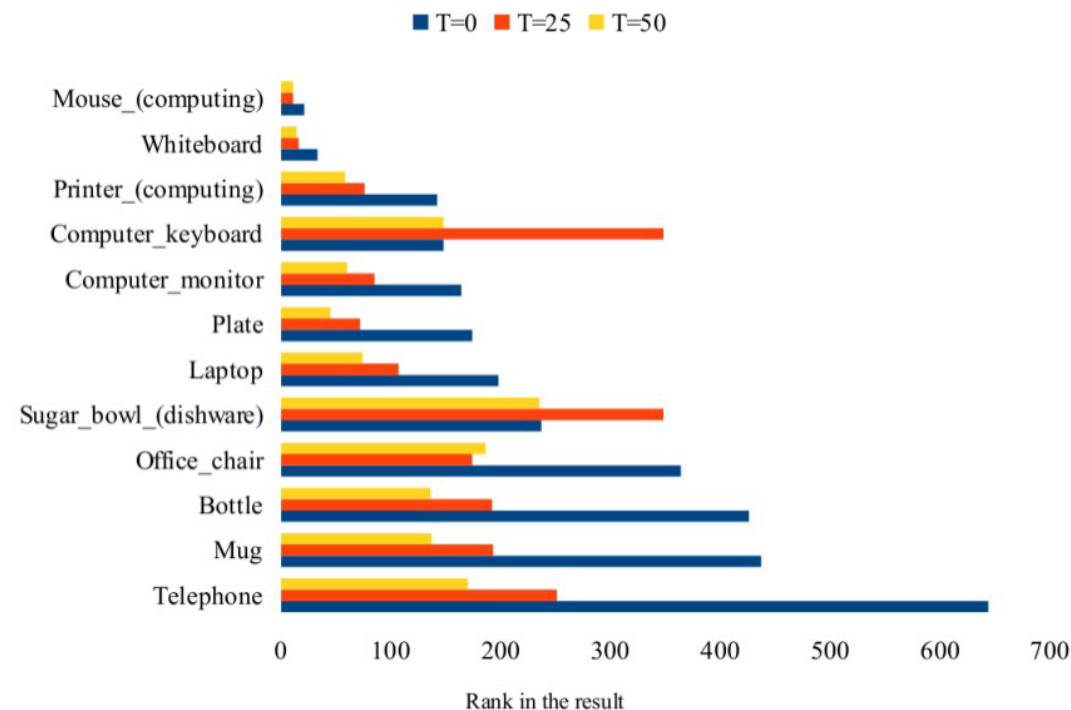


Annie cuts the bread in the kitchen with her knife \rightarrow dbp:Knife alsoof:Location dbp:Kitchen

ALOOF: robots learning by reading on the Web

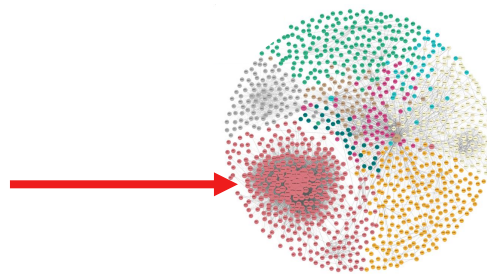
First Object Relation Knowledge Base:
 46.212 co-mentions gave 49 tools, 14 rooms, 101 "possible location" relations,

[Cabrio, Basile et al.]





DeKO

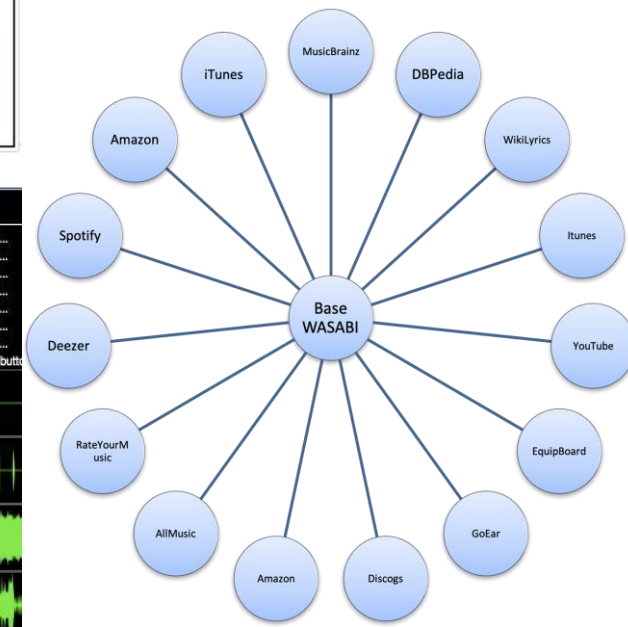
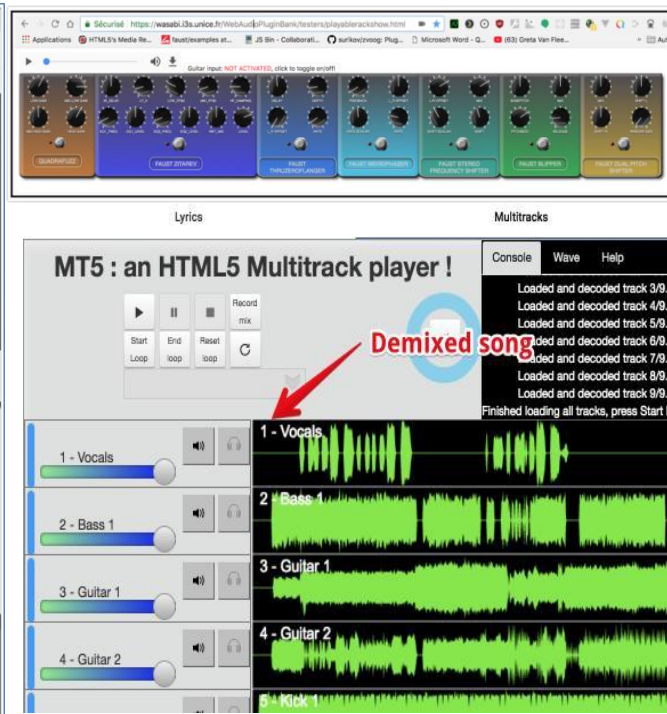
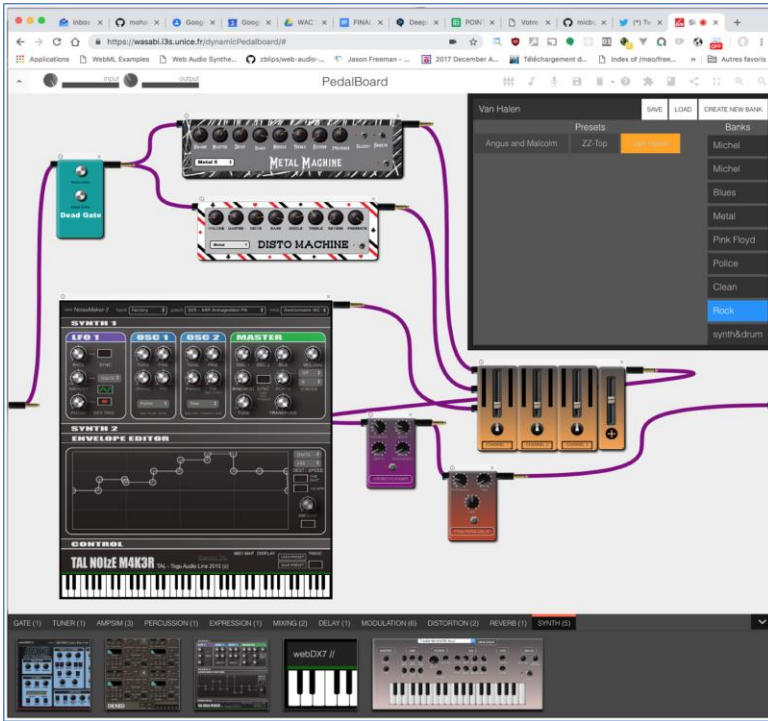


ALOOF: RDF dataset about objects

- common sense knowledge about objects: classification, prototypical locations and actions
- knowledge extracted from natural language parsing, crowdsourcing, distributional semantics, keyword linking, ...

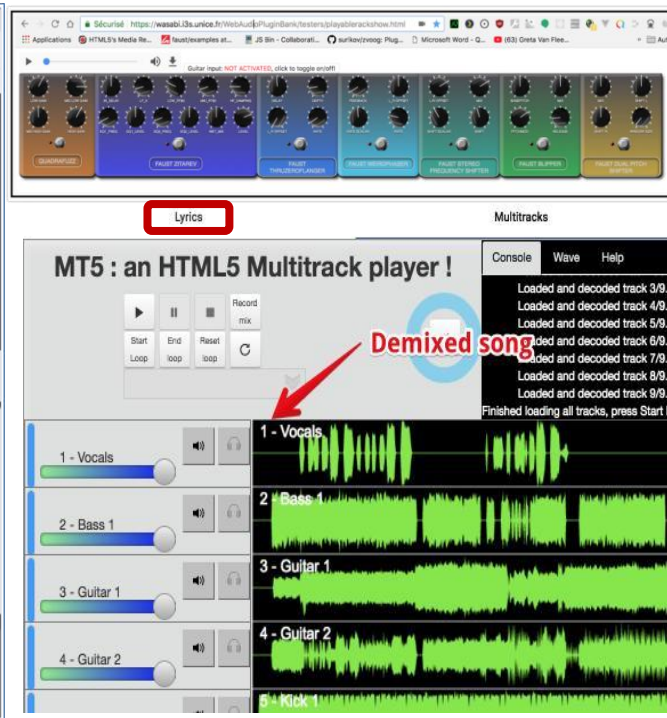
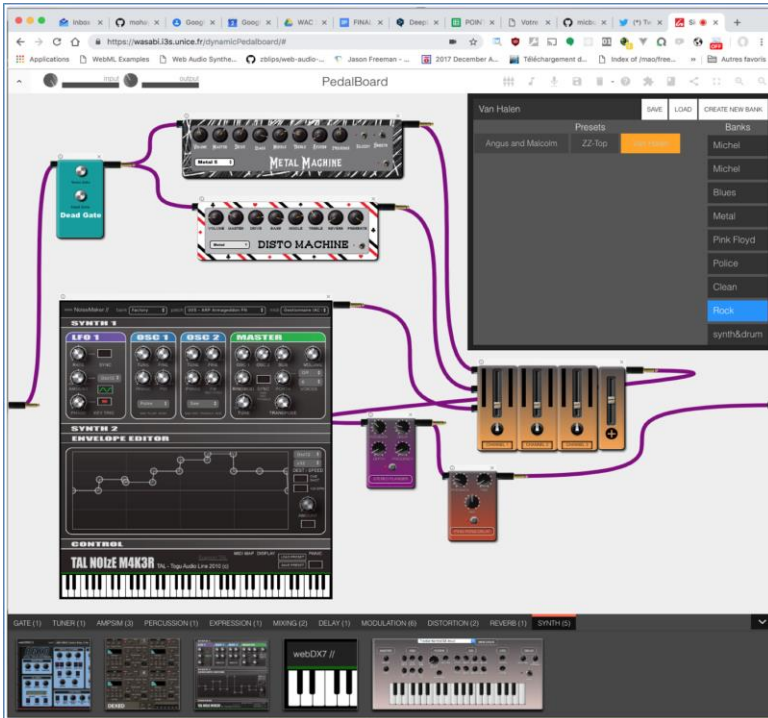
[Cabrio, Basile et al.]

[Buffa et al.]



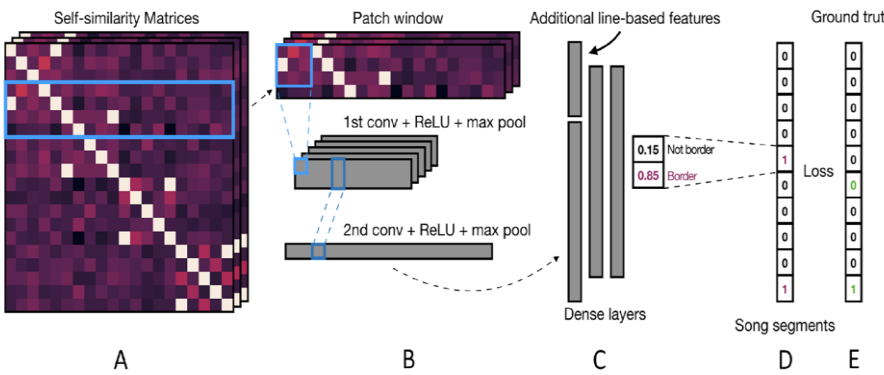
WASABI

Web-augmented music interactions



WASABI

Web-augmented music interactions



- V₁ you took me by surprise
just like a rainbow in the night
when i looked in your eyes
i could see diamonds shining bright
- V₂ i never realized
you'd be the one to make it right
you got me hypnotized
before my life was black and white
- B₁ you and me make a team
i want you near me every day
last night i had a dream
you went away
- C₁ dont break my heart
dont let me down
dont break my heart
dont make me frown
- V₃ were getting serious
life by your side can be so nice
you're so mysterious
yesterday you were cold as ice
- B₂ i wnder how you feel
if all your feelings are the same
when are your smiles for real
or is it a game
- C₂ dont break my heart
dont let me down
dont break my heart
dont make me frown
- C₃ dont break my heart
dont let me down
dont break my heart
dont make me frown
- C₄ dont break my heart
dont let me down
dont break my heart
dont make me frown
- O dont break my heart
dont let me down

repetitive lyrics structure

repetitive melody structure

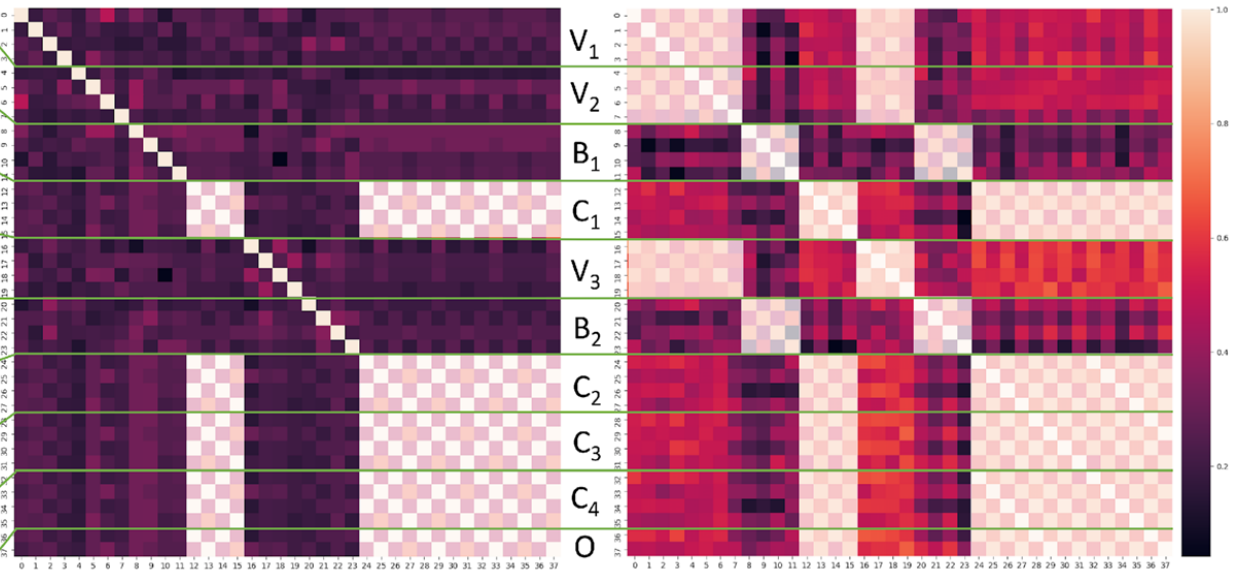
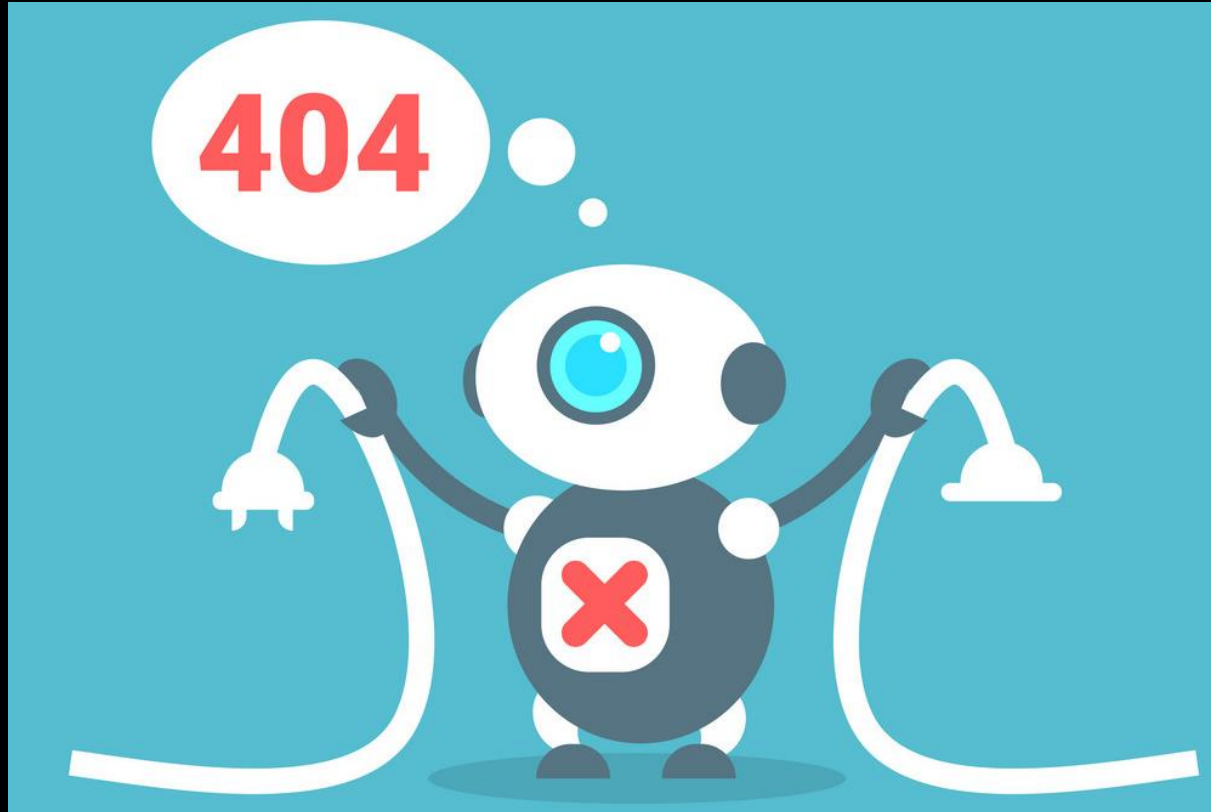


Figure 2: Convolutional Neural Network-based model inferring lyrics segmentation (Fell et al. 2018).

but...



ELIZA: You're not really talking about me -- are you ?
YOU: What?
ELIZA: What do you think ?
YOU: I mean, I feel happy to be writing right now, to be typing this as we speak. Lately, I've had writer's block.
ELIZA: You say you mean ?
YOU: That's what I said.
ELIZA: Can you elaborate on that ?
YOU: I only really feel happy when I'm writing.
ELIZA: How long have you been writing ?
YOU: Always, but I've been writing seriously since I was eighteen.
ELIZA: What incident are you thinking of ?
YOU: Um.
ELIZA: What does that suggest to you ?
YOU: I don't think it's enough, to only be happy when you write.
ELIZA: Does that trouble you ?
YOU: ...Um, because you can't write all the time. And then there's the real world, and you still have to deal with real people.
ELIZA: What are your feelings now ?
YOU: I've been thinking a lot of that poem by Yeats: "The Man and the Echo."
Do you know it?
ELIZA: Do you say you've been thinking a lot of that poem by yeats: "the man and the echo for some special reason ?
YOU: Yes



VOX

digital VT100
digivox



Lenny!

608 vidéos • 99 774 vues • Dernière modification le 13 juin 2018



Toao.net

S'ABONNER

"Lenny" is a hilarious set of recorded messages designed to waste telemarketers' time.

From Eliza to Lenny, talking to machines

- 22 Toao.net
- 23 Anna from Sirius XM wants Lenny Toao.net
- 24 Jim Evans is a licensed painter wh Toao.net
- 25 [NSFW] Jaeden says a different sc Toao.net
- 26 John Lopez from Save Today Mark Toao.net
- 27 [NSFW] Windows Technical Depart Toao.net
- 28 Cameron, Amy, and Cameron want Toao.net
- 29 Lenny - George with Medicare Hea Toao.net
- 30 Joshna and Dominic want explain Toao.net

barely visible evolution of the place of humans in
complex web applications

 =  user

 =  data

 =  processor

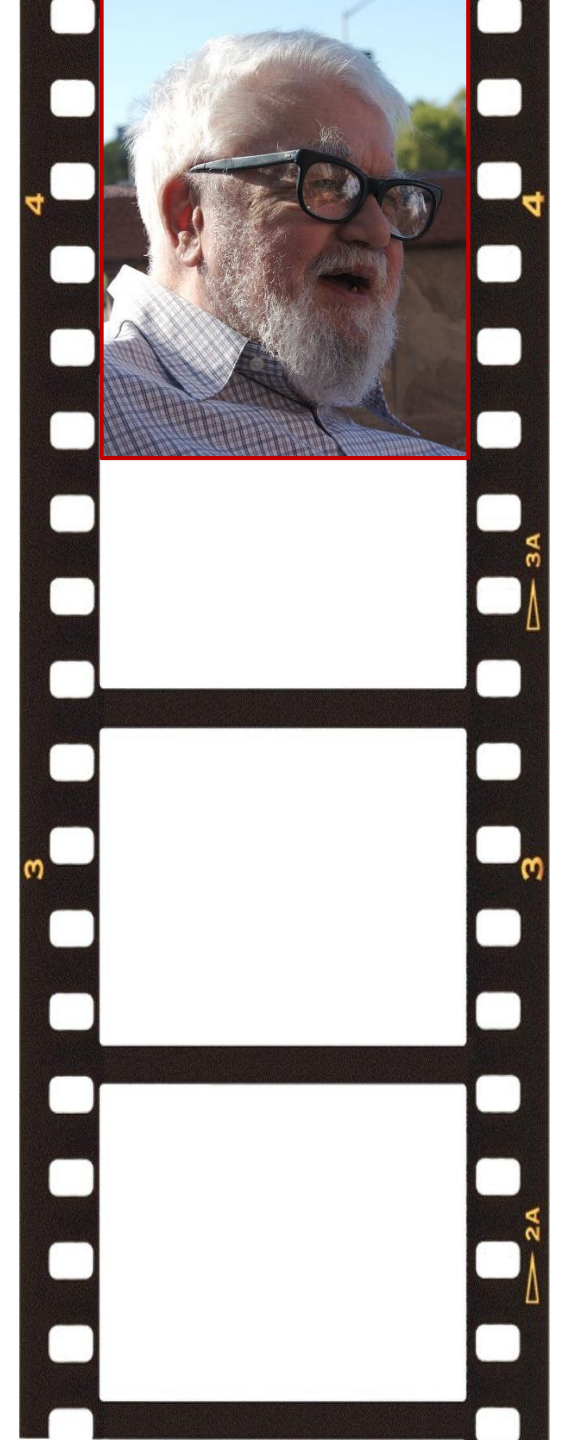
and only now are you reading this top left text... predictable means manipulable

You will read this first

Then you will read this...

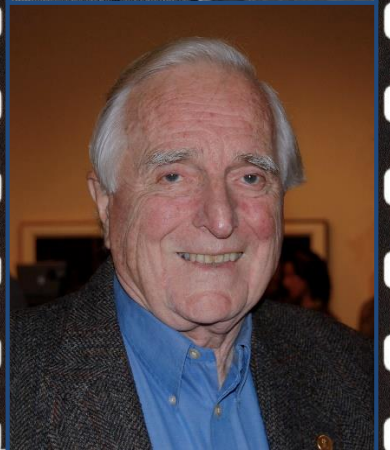
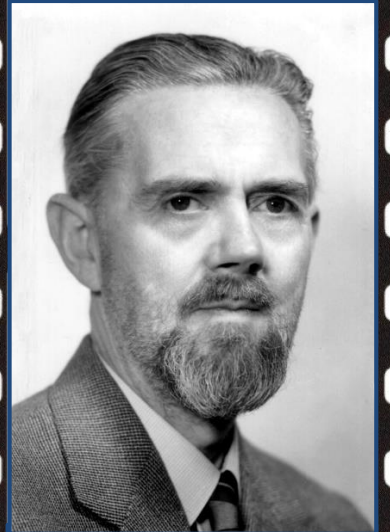
Then this...

AI for **Artificial Intelligence** (McCarthy et al., 1955)



AI for **Artificial Intelligence** (McCarthy et al., 1955)

IA for **Intelligence Amplification** (Ashby, 1956) and
Intelligence Augmentation (Engelbart, 1962)

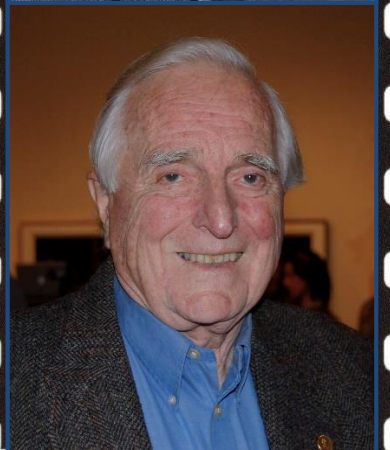
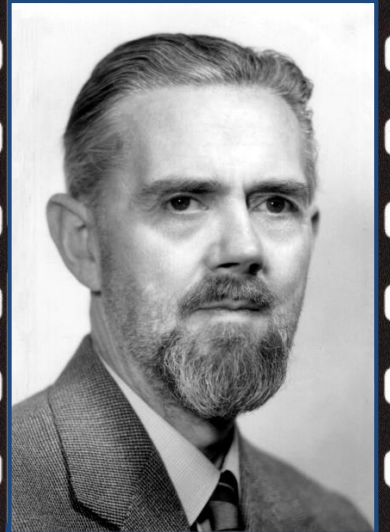


AI for **Artificial Intelligence** (McCarthy et al., 1955)

IA for **Intelligence Amplification** (Ashby, 1956) and
Intelligence Augmentation (Engelbart, 1962)

Web as the *rendez-vous* point for two fields
born in the 50s...

AI & IA



AI and Web societies



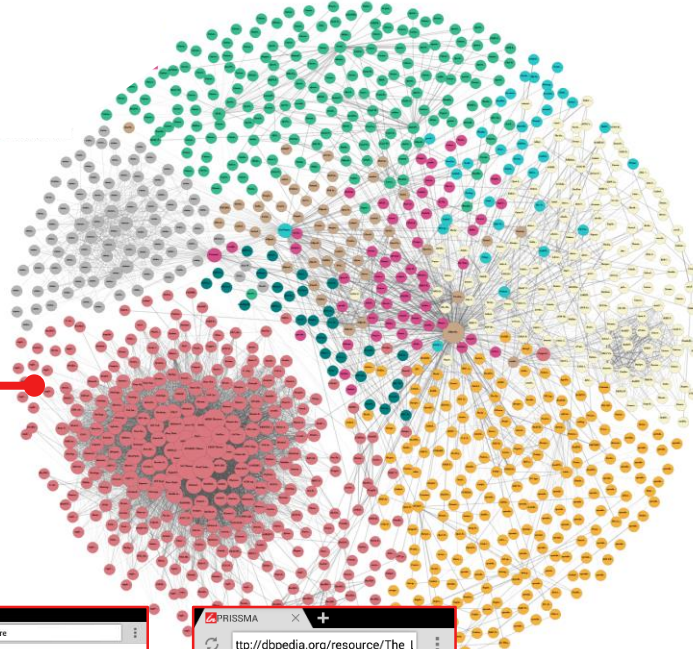
AI and Web augmentation

Web for classical AI tasks

AI for classical Web tasks

MODELING USERS

- individual context
- social structures



PRISSMA

PRISSMA

http://dbpedia.org/resource/The_Louvre

The Louvre

The Musée du Louvre—in English, the Louvre Museum or simply The Louvre—is one of the world's largest museums, and a historic monument. A central landmark of Paris, France, it is located on the right bank of the Seine in the 1st arrondissement (district). Nearly 30,000 objects from prehistory to the 19th century are exhibited over an area of 60,600 square metres (652,300 square feet). With more than 8 million visitors each year, the Louvre is the world's most-visited museum.

City
Paris

Director
Jean-Luc Martinez

Mayor
Marie-Laure de Rochebrune

Established
1793

PRISSMA

http://dbpedia.org/resource/The_L

The Louvre

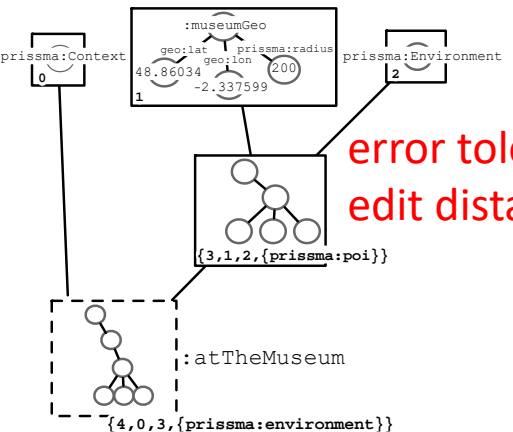
Address
Palais Royal, Musée du Louvre - 75058 Paris - France

Public Transportation
Palais Royal – Musée du Louvre *Louvre-Rivoli

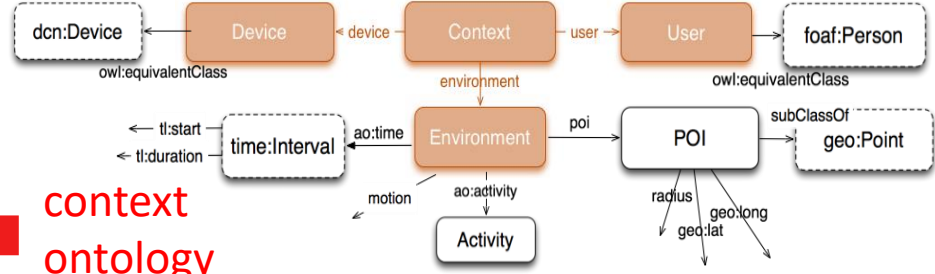
Tel:
+33 (0)1 40 20 53 17

Opening Hours
Monday, Thursday, Saturday, Sunday: from 9 a.m. to 6 p.m. Wednesday, Friday: from 9 a.m. to 9:45 p.m. Closed on Tuesdays

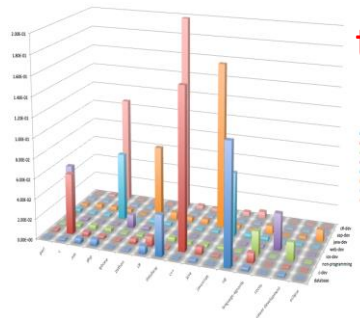
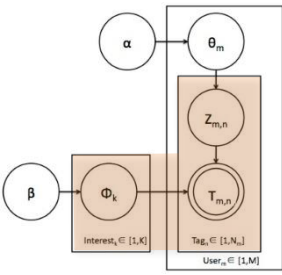
Ticket Price
Tickets for the Permanent Collections: 12€
Tickets for Exhibitions in the Hall Napoléon: 13€
Combined Ticket: 16€



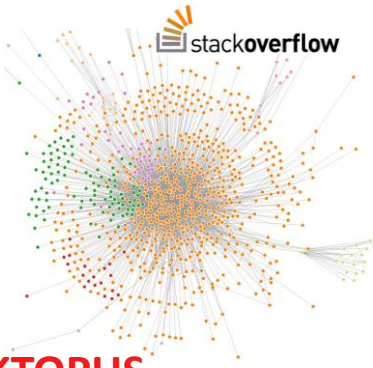
error tolerant graph
edit distance



context
ontology



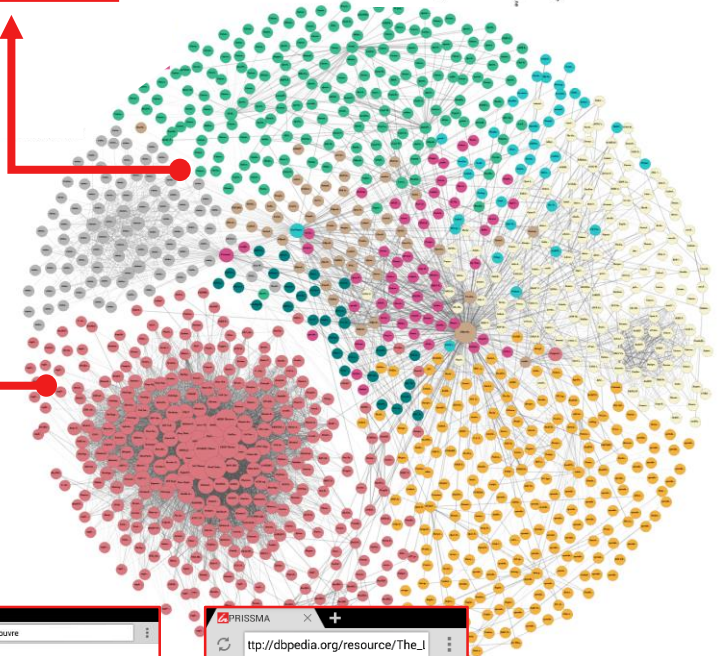
tag, topic, user distribution



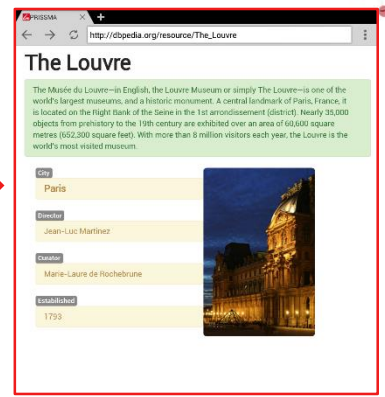
tag and folksonomy restructuring with prefix trees



OCKTOPUS



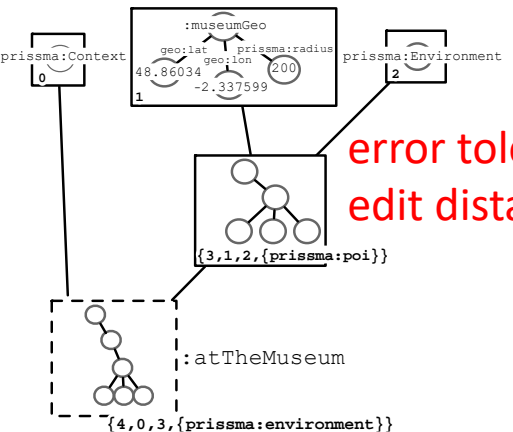
PRISSMA



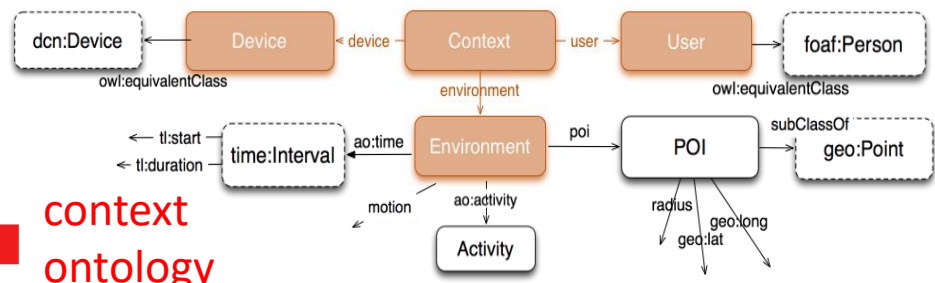
[Meng et al.]

MODELING USERS

- individual context
- social structures

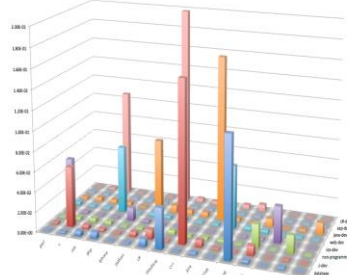
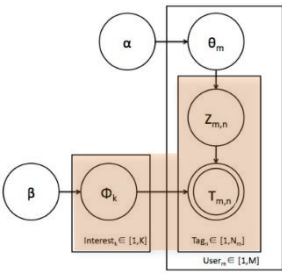


error tolerant graph edit distance

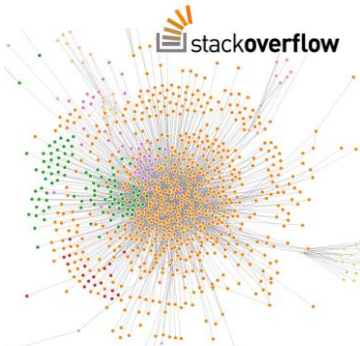


context ontology

[Costabello et al.]



tag, topic, user distribution



tag and folksonomy restructuring with prefix trees

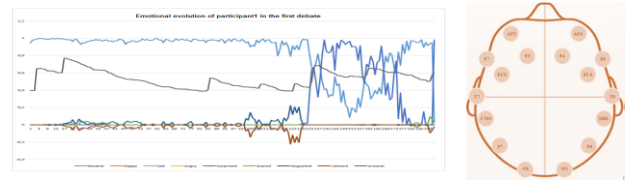


[Villata, Cabrio et al.]

[Meng et al.]

OCKTOPUS

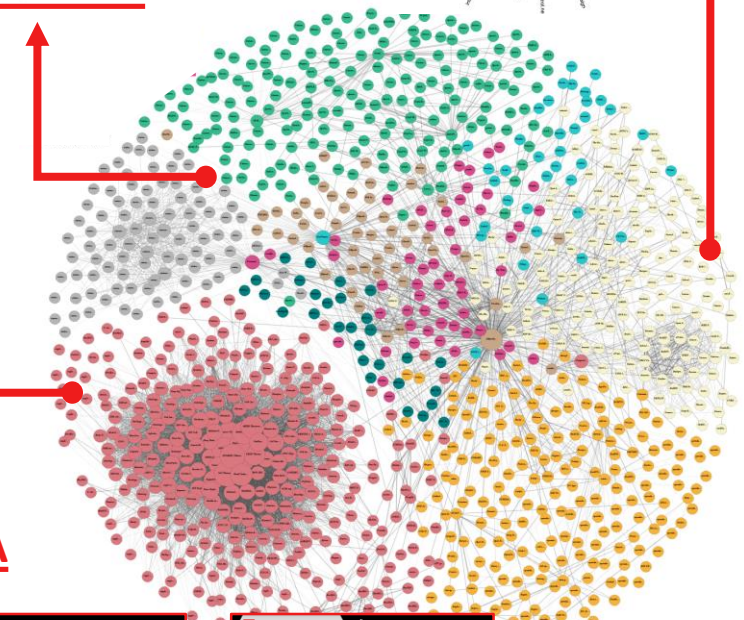
EMOCA&SEEMPAD



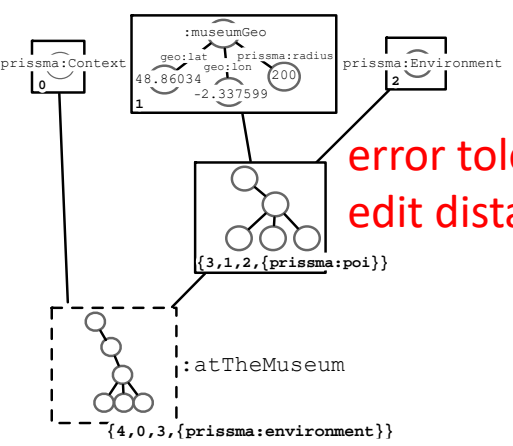
emotion detection & annotation

MODELING USERS

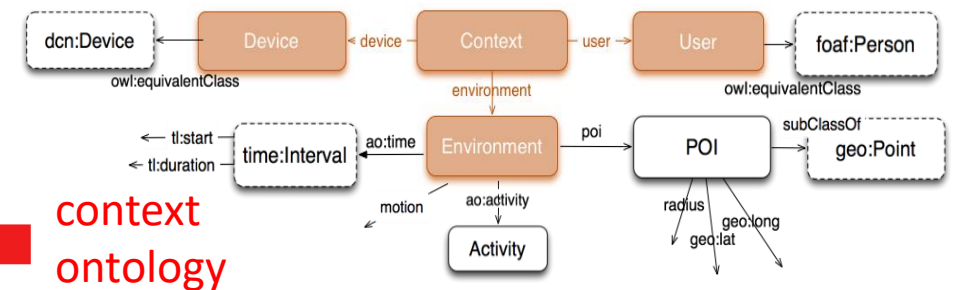
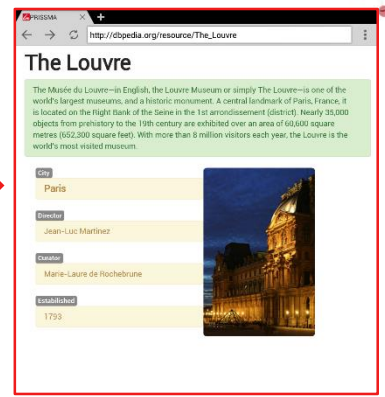
- individual context
- social structures



PRISSMA



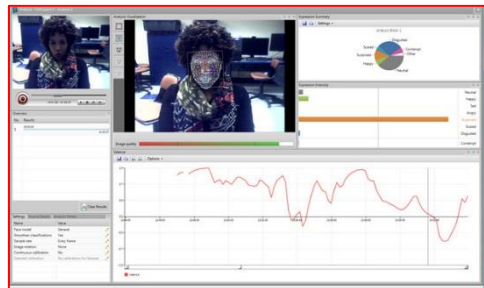
error tolerant graph edit distance



context ontology

[Costabello et al.]

DEBATES & EMOTIONS



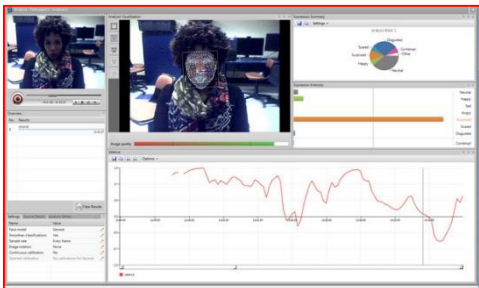
#IRC



Dataset				
Topic	#arg	#pair	#att	#sup
BAN ANIMAL TESTING	49	28	18	10
GO NUCLEAR	40	24	15	9
HOUSEWIVES SHOULD BE PAID	42	18	11	7
RELIGION DOES MORE HARM THAN GOOD	46	23	11	12
ADVERTISING IS HARMFUL	71	16	6	10
BULLIES ARE LEGALLY RESPONSIBLE	71	12	3	9
DISTRIBUTE CONDOMS IN SCHOOLS	68	27	11	16
ENCOURAGE FEWER PEOPLE TO GO TO THE UNIVERSITY	55	14	7	7
FEAR GOVERNMENT POWER OVER INTERNET	41	32	18	14
BAN PARTIAL BIRTH ABORTIONS	41	26	15	11
USE RACIAL PROFILING FOR AIRPORT SECURITY	31	10	1	9
CANNABIS SHOULD BE LEGALIZED	43	33	20	13
TOTAL	598	263	136	127

[Villata, Cabrio et al.]

DEBATES & EMOTIONS

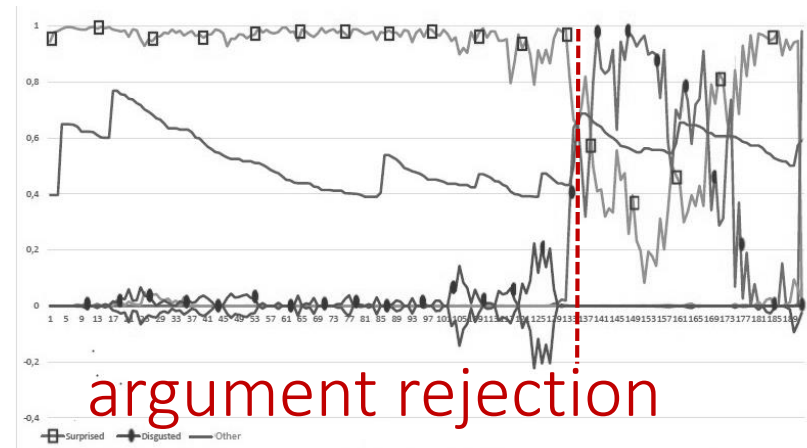


#IRC



Dataset				
Topic	#arg	#pair	#att	#sup
BAN ANIMAL TESTING	49	28	18	10
GO NUCLEAR	40	24	15	9
HOUSEWIVES SHOULD BE PAID	42	18	11	7
RELIGION DOES MORE HARM THAN GOOD	46	23	11	12
ADVERTISING IS HARMFUL	71	16	6	10
BULLIES ARE LEGALLY RESPONSIBLE	71	12	3	9
DISTRIBUTE CONDOMS IN SCHOOLS	68	27	11	16
ENCOURAGE FEWER PEOPLE TO GO TO THE UNIVERSITY	55	14	7	7
FEAR GOVERNMENT POWER OVER INTERNET	41	32	18	14
BAN PARTIAL BIRTH ABORTIONS	41	26	15	11
USE RACIAL PROFILING FOR AIRPORT SECURITY	31	10	1	9
CANNABIS SHOULD BE LEGALIZED	43	33	20	13
TOTAL	598	263	136	127

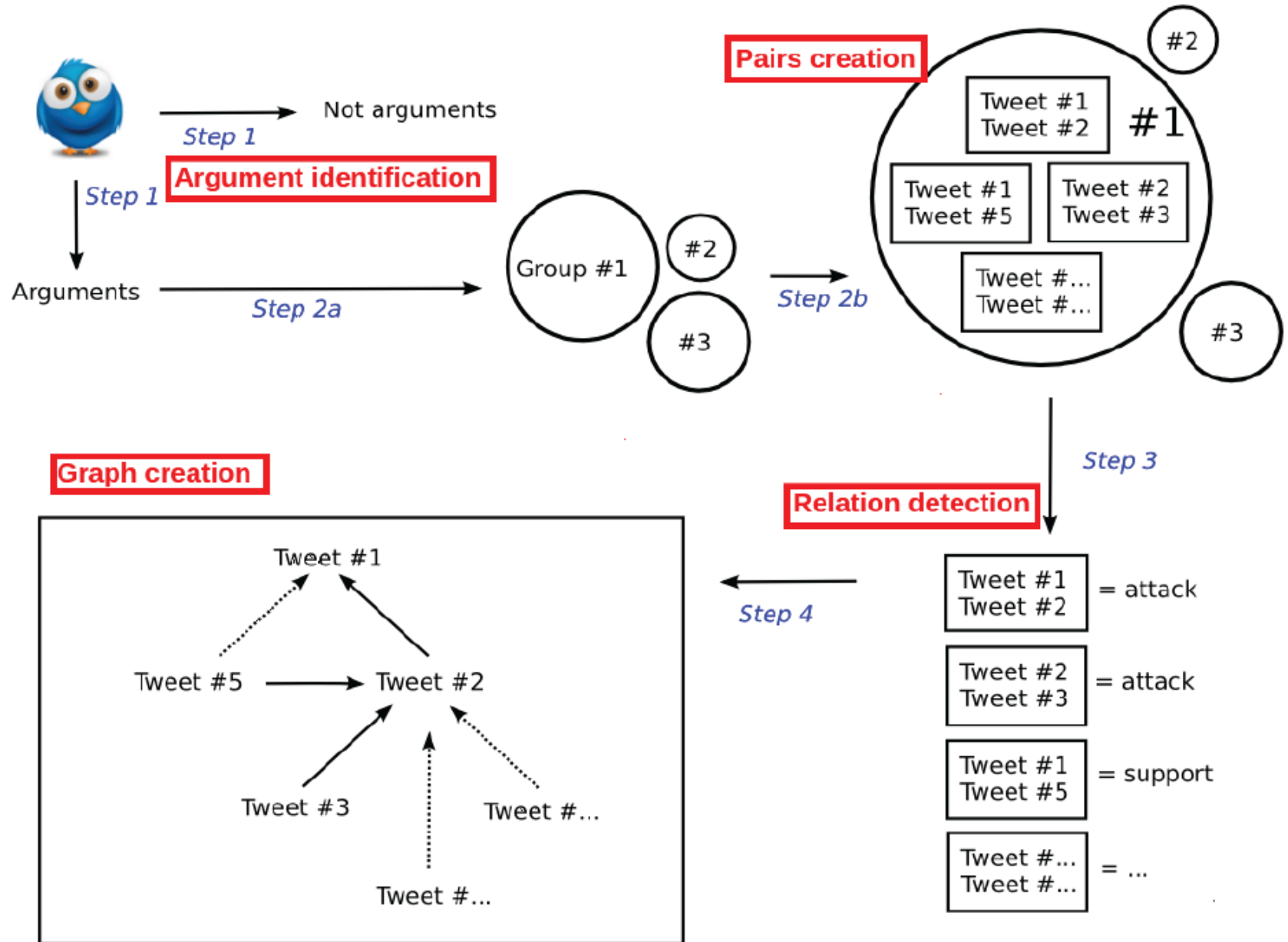
[Villata, Cabrio et al.]



	NB ARG	ATTACK	SUPPORT
Pleasant	0,7067	-0,3383	-0,3800
Unpleasant	-0,7067	0,3383	0,3800
High ENG	-0,6903	-0,3699	-0,1117
LowENG	-0,1705	0,5337	-0,0615
Neutral	0,8887	-0,0895	-0,3739
Disgusted	0,1017	0,8379	0,5227
Scared	0,2606	-0,4132	-0,7107
Angry	-0,7384	-0,5072	-0,0937

attacks-disgust

Argument mining pipeline



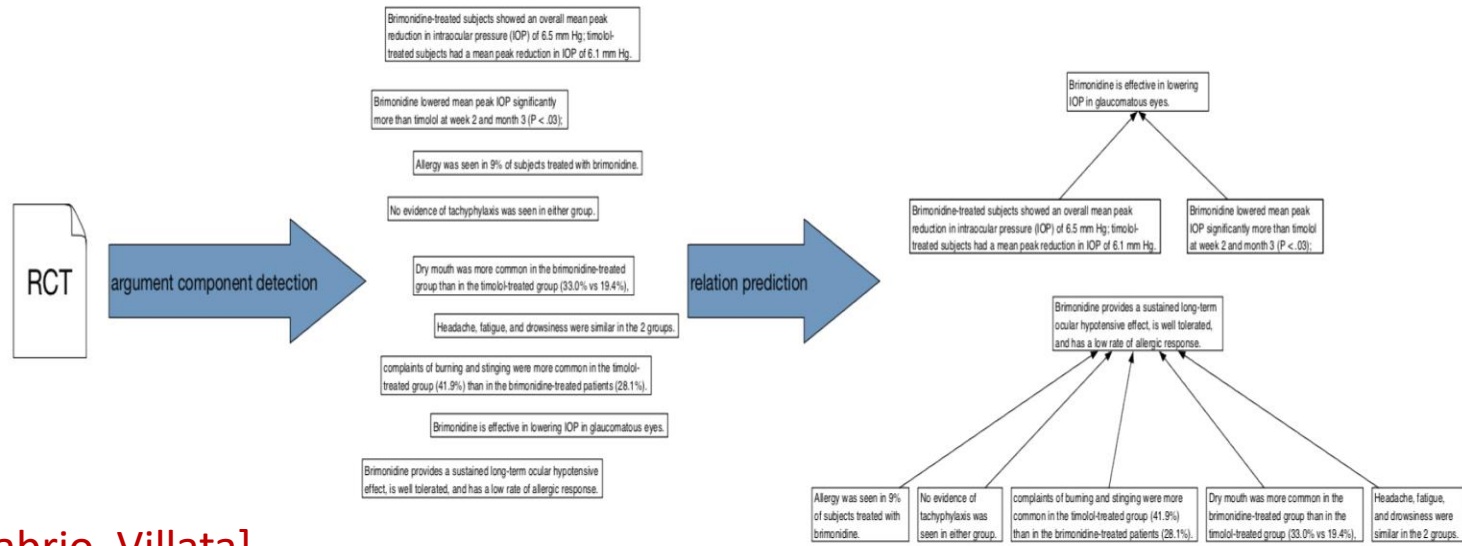
OPINIONS

NLP, ML and arguments to monitor online image

[Villata, Cabrio, et al.]

ARGUMENT MINING ON CLINICAL TRIALS

[Mayer, Cabrio, Villata]



- NLP, ML and arguments
- assist evidence-based medicine
- support doctors and clinicians
- identify doc. for certain disease
- analyze argumentative content and PICO elements

ACTA
Argumentative Clinical Trial Analysis

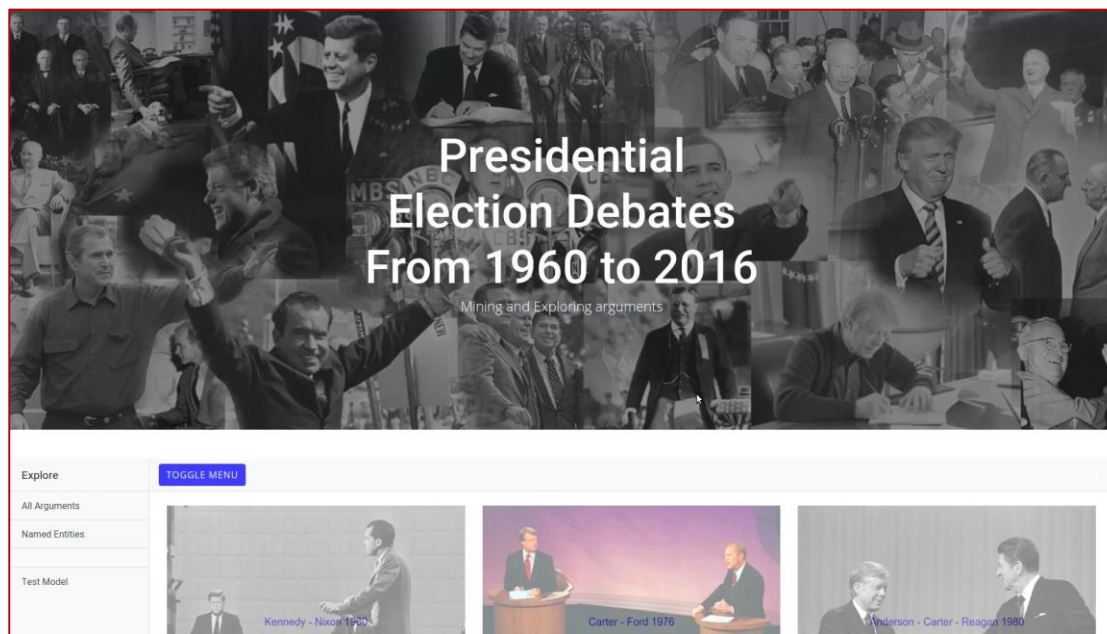
Home About Contacts Services

Search on PubMed:

randomized controlled trial[Publication Type] neoplasm[MeSH Terms] quality of life[MeSH]

Search

PubMed ID	Title
-----------	-------



ARGUMENT MINING ON POLITICAL SPEECHES [Mayer, Cabrio, Villata]

- NLP and Machine Learning.
- Support historians/social science scholars
- Analyze arguments in political speeches
- DISPUTool : 39 political debates, last 50 years of US presidential campaigns (1960-2016)

Options Toggle Menu Home

Explore All Arguments

Named Entities

Test Model

19 Oct 2016 👍 09 Oct 2016

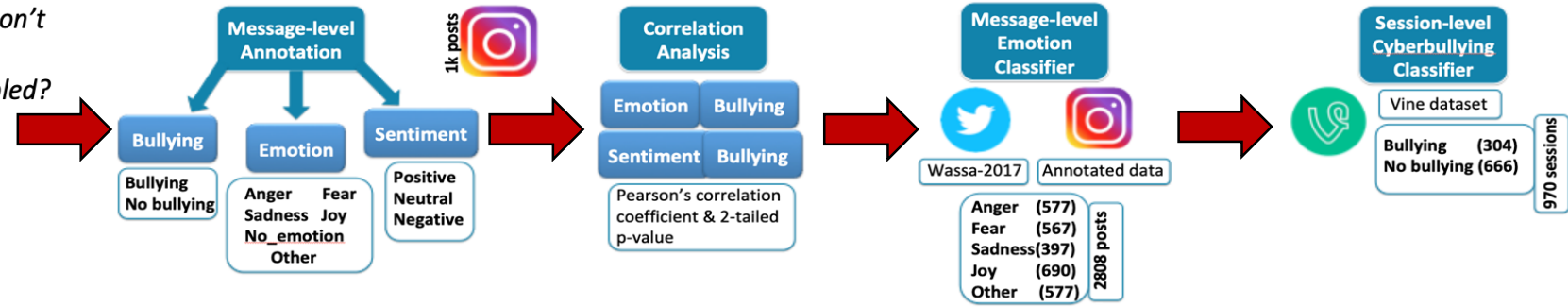
Arguments Table

Show 10 entries Search:

Speech	Name	Component	Year	Date
Good evening from the Thomas and Mack Center ...	Chris Wallace	0	2016	19 Oct 2016
I'm Chris Wallace of Fox News, and I welcome y...	Chris Wallace	0	2016	19 Oct 2016
This debate is sponsored by the Commission on...	Chris Wallace	0	2016	19 Oct 2016
The commission has designed the format. Six ro...	Chris Wallace	0	2016	19 Oct 2016
Both campaigns have agreed to those rules.	Chris Wallace	0	2016	19 Oct 2016
No noise, except right now, as we welcome the...	Chris Wallace	0	2016	19 Oct 2016
Which means that you will, in effect, determi...	Chris Wallace	0	2016	19 Oct 2016
Thank you very much, Chris.	Hillary D. R. Clinton	0	2016	19 Oct 2016
	Hillary D. R. Clinton	0	2016	19 Oct 2016

localhost:5000/chat/19 Oct 2016

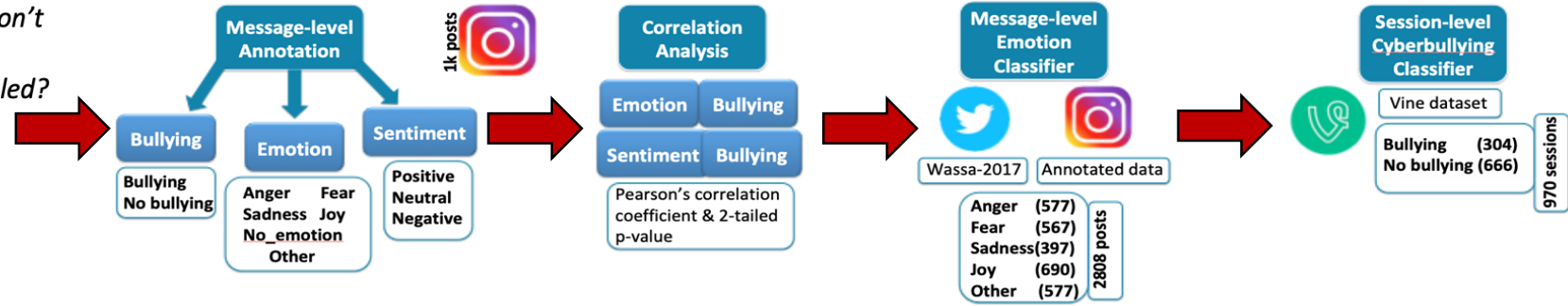
- anger negative bullying 🤔 "Shove off baby ugly @username"
- 🤔 "@username fuck you dump piece of shit. Don't be talking shit to the homie like that ight."
- 🤔 "Are you fucking retarded or mentally disabled? Oh wait I think you're both..."
- joy positive no bullying 😊 "Happy birthday #myidol"
- 😊 "You're lovely, wow"
- 😊 "Marry me___!? Your perff"



CYBERBULLYING [Corazza, Arslan, Cabrio, Villata]

CREEP EU project: detect and prevent

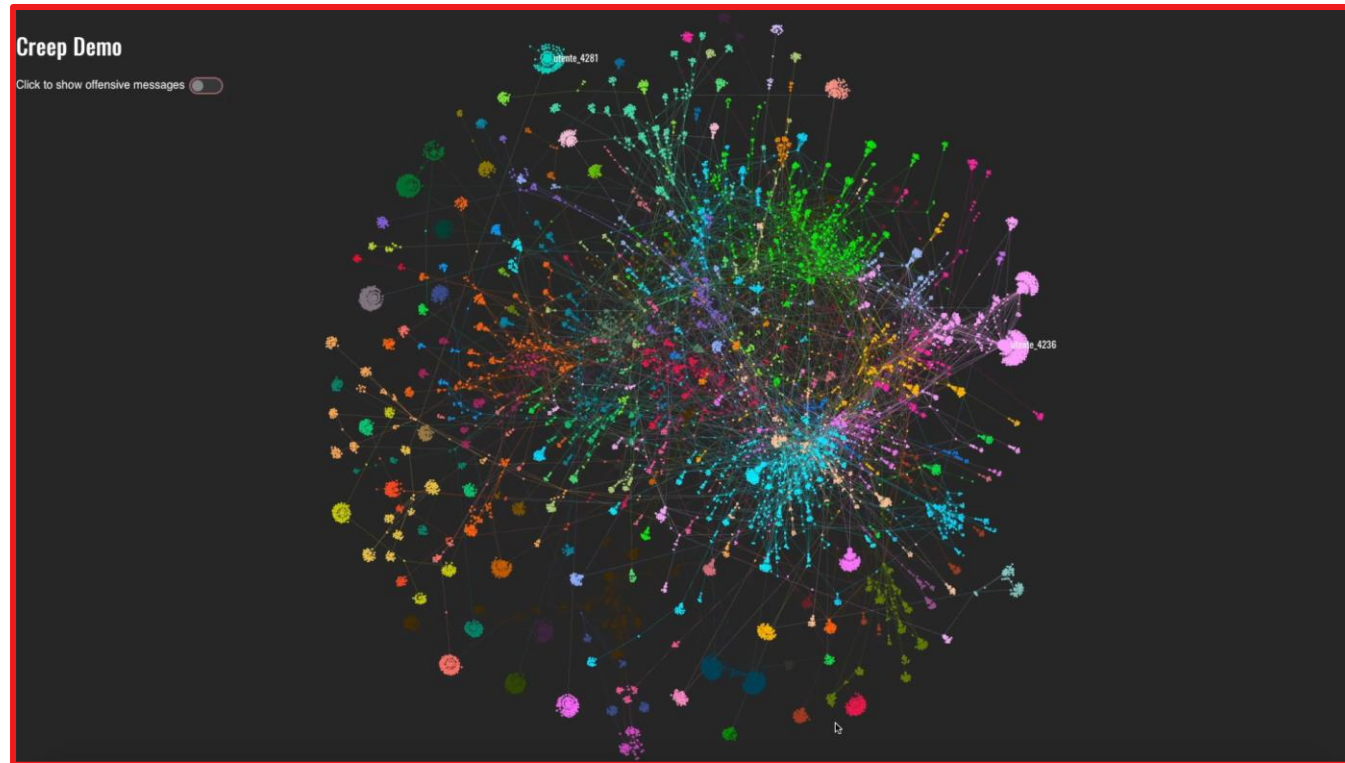
- anger 🤔 "Shove off baby ugly @username"
- negative 🤔 "@username fuck you dump piece of shit. Don't be talking shit to the homie like that ight."
- bullying 🤔 "Are you fucking retarded or mentally disabled? Oh wait I think you're both..."
- joy 😄 "Happy birthday #myidol"
- positive 😄 "You're lovely, wow"
- no bullying 😄 "Marry me___!?! Your perfff"



CYBERBULLYING [Corazza, Arslan, Cabrio, Villata]

CREEP EU project: detect and prevent

Majority Base	Bullying	No Bullying	Avg
word-1gram	0.00	0.81	0.4071
word-2gram	0.62	0.83	0.7277
char-3gram	0.36	0.83	0.5946
char-4gram	0.62	0.84	0.7333
char-5gram	0.62	0.85	0.7376
EmqLex	0.61	0.85	0.7290
predemotion	0.56	0.66	0.6113
SenticNet	0.61	0.77	0.6875
	0.62	0.81	0.7154
✓	0.59	0.84	0.7142
✓ ✓	0.64	0.86	0.7497
✓ ✓ ✓	0.63	0.86	0.7419
✓ ✓ ✓ ✓	0.62	0.86	0.7374
✓ ✓ ✓ ✓ ✓	0.63	0.85	0.7407
✓ ✓ ✓ ✓ ✓ ✓	0.64	0.85	0.7466
✓ ✓ ✓ ✓ ✓ ✓ ✓	0.65	0.86	0.7520
✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓	0.64	0.85	0.7442
✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓	0.64	0.84	0.7412



but...





facebook: 2 400 millions

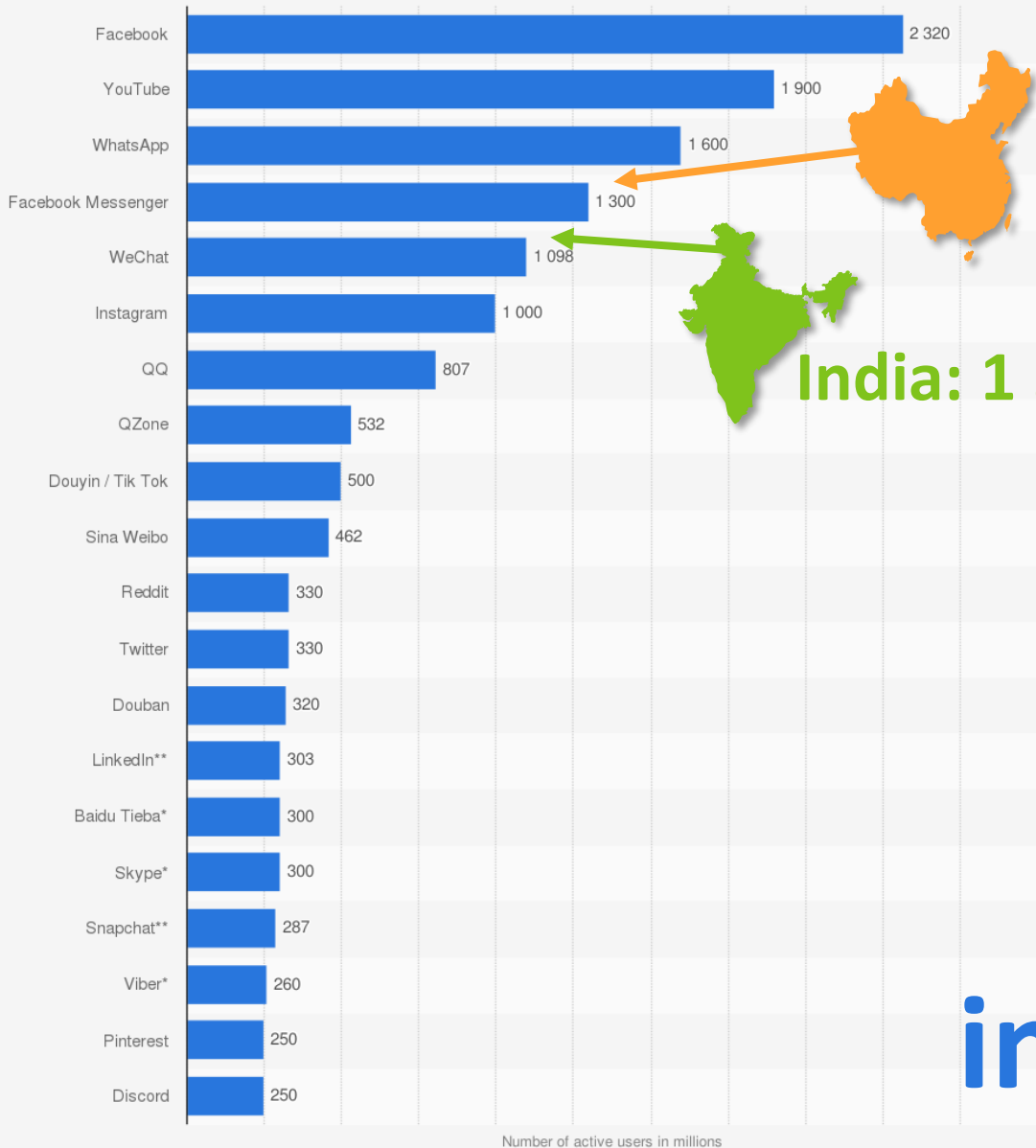


China: 1 400 millions



India: 1 300 millions

Most popular social networks worldwide as of April 2019, ranked by number of active users (in millions)



China: 1 400 millions

India: 1 300 millions

inverse view of the world

Web Robots



e.g. Wikipedia bots

2 187 bot approved for use on the English Wikipedia to help maintain 45 223 137 pages

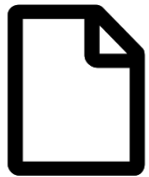


Examples

Some examples of bots are:

- [User:AAAlertBot](#) – delivering [article alerts](#) to WikiProjects about ongoing discussions.
- [User:AnomieBOT](#) – large variety of tasks, most well known for adding dates to amboxes.
- [User:BracketBot](#) – notifies users of mismatched brackets in recently edited articles.
- [User:ClueBot NG](#) – reverts [vandalism](#).
- [User:Cydebot](#) – generally carries out tasks associated with deletion.
- [User:DatBot](#) – patrols the edit filters and resizes non-free images.
- [User:DumbBOT](#) – often removes protection templates from recently unprotected pages.
- [User>ListeriaBot](#) – Experimental bot by [Magnus Manske](#). It generates and updates lists on Wikipedia.
- [User:Lowercase sigmabot III](#) – archives talk pages.
- [User:ProcseeBot](#) – automatically blocks proxies due to both the local policy against [open proxies](#).
- [User:SineBot](#) – signs comments left on talk pages.
- [User:TheMagikBOT](#) – often adds protection templates to articles without them.
- [User:WP 1.0 bot](#) – works with the [Version 1.0 Editorial Team](#).
- [User:Yobot](#) – syntax fixes and tagging.

web...



page



social

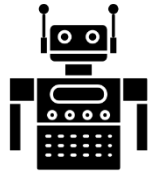


apps

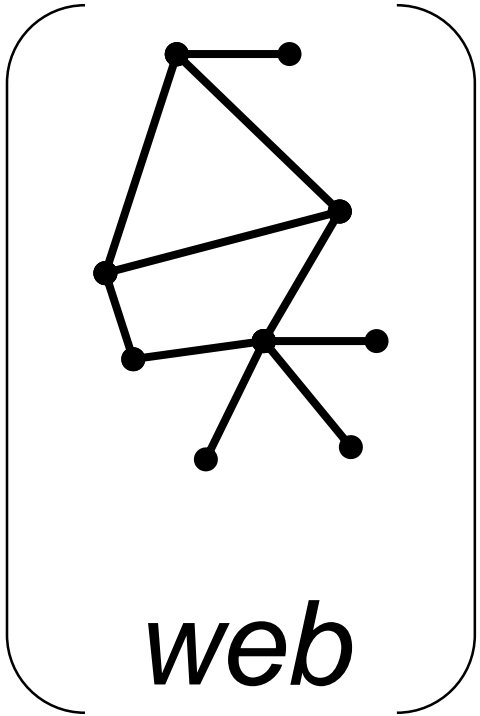


data

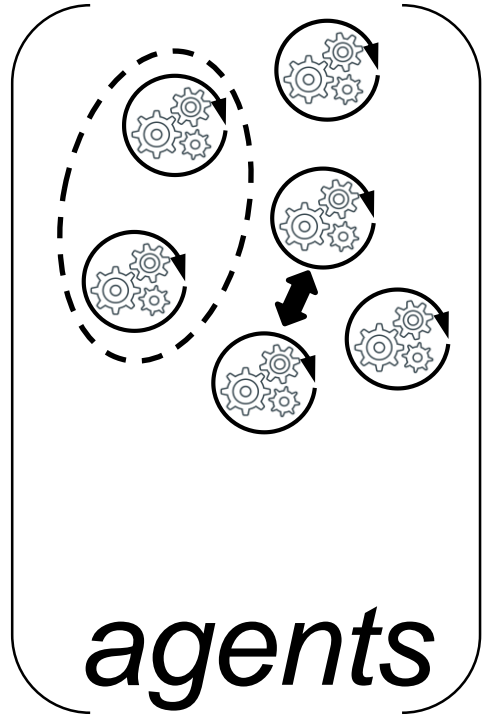
...



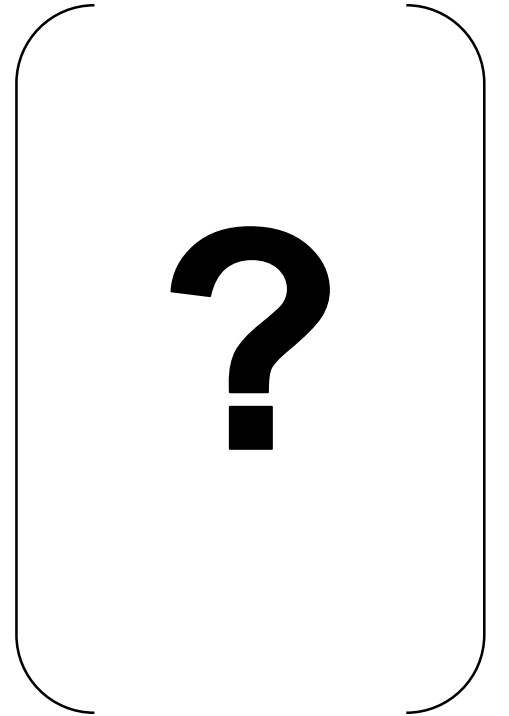
a.i.



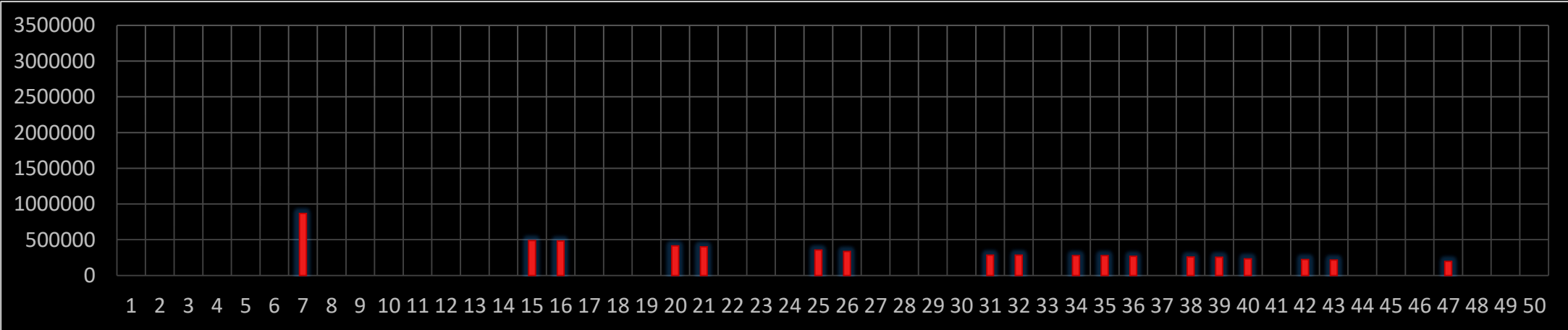
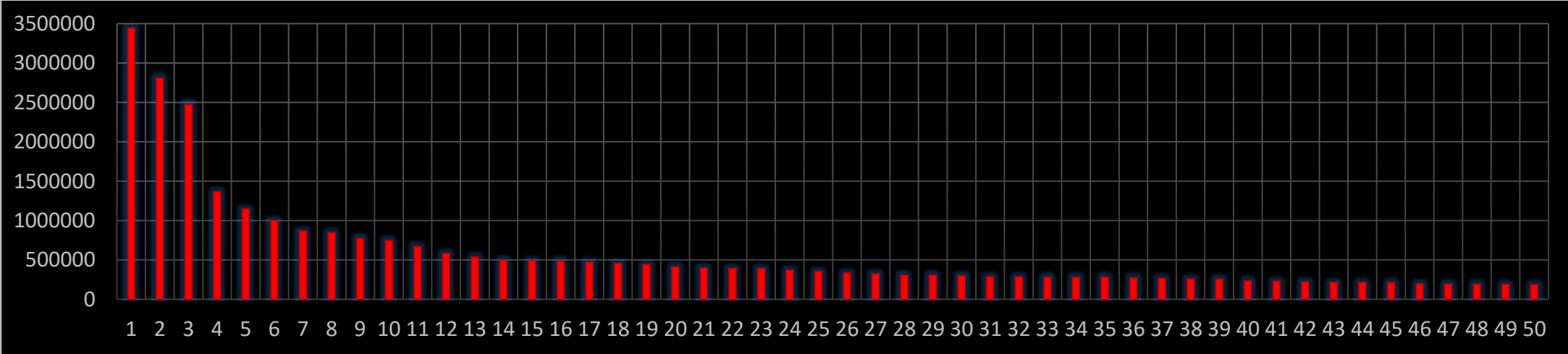
+



=



EDITS





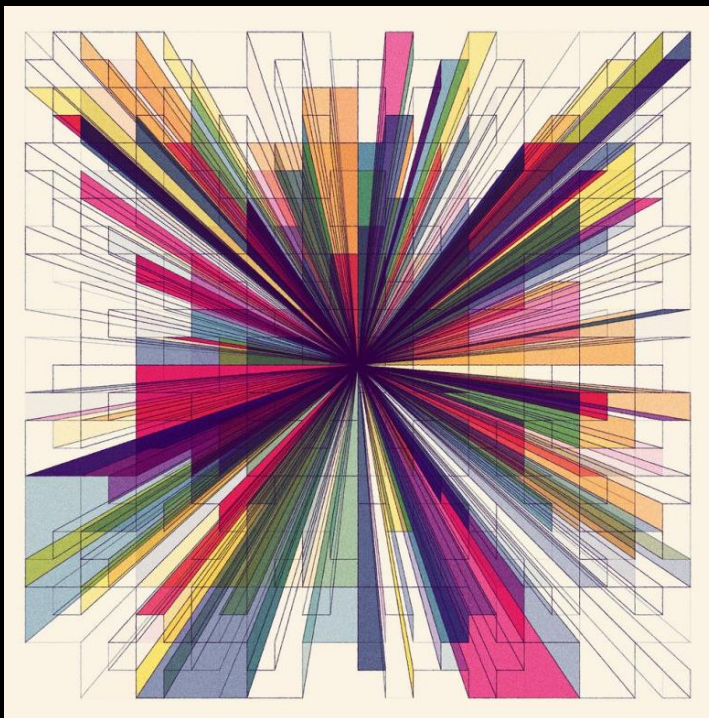
massive interaction design
ergonomics in hybrid web communities

- 1 ° K / -458 ° F / -272 ° C ✓
- 420 ° K / 300 ° F / 150 ° C ✓
- Vacuum / High pressure ✓
- Ionizing radiation ✓
- Himalayas / Deep sea ✓

Extremophile

tardigrades Web AIs



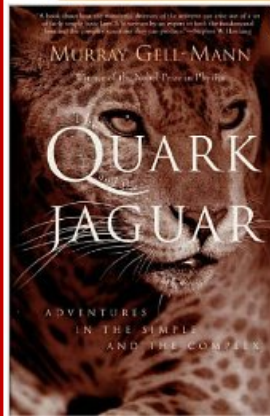


perspectives

Web 1.0, ...

amazon.com

Click to **LOOK INSIDE!**



The Quark and the Jaguar: Adventures in the Simple and the Complex [Paperback]

[Murray Gell-Mann](#) (Author)

★★★★☆ (30 customer reviews)

Price: **\$13.59**

In Stock.

Ships from and sold by **Amazon.com**.

Formats

	Amazon Price	New from	Used from
Hardcover	--	\$13.67	\$0.55
Paperback	\$13.59	\$10.01	\$7.84

[Show 4 more formats](#)

Product Details

Paperback: 392 pages

Publisher: St. Martin's Griffin (September 15, 1995)

Language: English

ISBN-10: 0805072535

ISBN-13: 978-0805072532

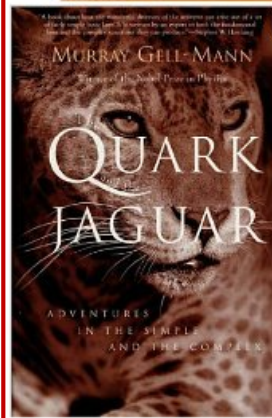
Product Dimensions: 9.1 x 6.1 x 1.2 inches

Shipping Weight: 1.2 pounds ([View shipping rates and policies](#))

Web 1.0, 2.0...

amazon.com

Click to **LOOK INSIDE!**



The Quark and the Jaguar: Adventures in the Simple and the Complex [Paperback]

[Murray Gell-Mann](#) (Author)

★★★★☆ (30 customer reviews)

Price: **\$13.59**

In Stock.

Ships from and sold by **Amazon.com**.

Formats

	Amazon Price	New from	Used from
<input type="checkbox"/> Hardcover	--	\$13.67	\$0.55
<input checked="" type="checkbox"/> Paperback	\$13.59	\$10.01	\$7.84

[Show 4 more formats](#)

Product Details

Paperback: 392 pages

Publisher: St. Martin's Griffin (September 15, 1995)

Language: English

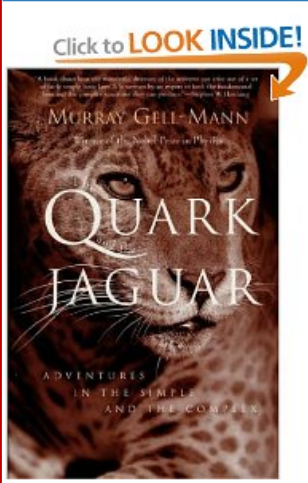
ISBN-10: 0805072535

ISBN-13: 978-0805072532

Product Dimensions: 9.1 x 6.1 x 1.2 inches

Shipping Weight: 1.2 pounds ([View shipping rates and policies](#))

Web 1.0, 2.0, 3.0 ...



Click to **LOOK INSIDE!** **The Quark and the Jaguar: Adventures in the Simple and the Complex** [Paperback]

[Murray Gell-Mann](#) (Author)
★★★★☆ (30 customer reviews)

Price: **\$13.59**

In Stock.
Ships from and sold by **Amazon.com**.

Formats	Amazon Price	New from	Used from
<input type="checkbox"/> Hardcover	--	\$13.67	\$0.55
<input checked="" type="checkbox"/> Paperback	\$13.59	\$10.01	\$7.84

[Show 4 more formats](#)

Product Details

- Paperback:** 392 pages
- Publisher:** St. Martin's Griffin (September 15, 1995)
- Language:** English
- ISBN-10:** 0805072535
- ISBN-13:** 978-0805072532
- Product Dimensions:** 9.1 x 6.1 x 1.2 inches
- Shipping Weight:** 1.2 pounds ([View shipping rates and policies](#))
- Average Customer Review:** ★★★★★ (30 customer reviews)
- Amazon Bestsellers Rank:** #373,365 in Books

Customer Reviews



Average Customer Review
★★★★☆ (30 customer reviews)

The **World-Wide Web**

The World-Wide Web (W3) was developed to be a pool of human knowledge, which would allow collaborators in remote sites to share their ideas and all aspects of a common project. Physicists and engineers at CERN, the European Particle Physics Laboratory in Geneva, Switzerland, collaborate with many other institutes to build the software and hardware for high-energy physics research. The idea of

- Evolution of objects from being principally human-readable documents to contain more machine-oriented semantic information, allowing more sophisticated processing;
- Conventions on the Internet for charging and commercial use to allow direct access to for-profit services.

Conclusion

It is intended that after reading this article you will have an idea of what W3 is, where it fits in with other systems in the field, and where it is going. There is much more to be said, especially about providing information, but this is described

(“anchor”) causes the client program to retrieve another object from some other computer, a “server.” The retrieved object is normally also in a hypertext format, so the process of navigation continues (see Figure 1).

When viewing some documents, the reader can request a search, by typing in plain text (or complex commands) to send to the server, rather than following a link. In either case, the client sends a request off to the

publishers or system managers of one sort or another. However, the incredible diversity of information available gives great credit to the creativity and ingenuity of information providers, and points to a very exciting future.

Appendix. **Getting Started**

If you have a vt100 terminal, you can try out a full-screen interface by telnet to `ukanaix.cc.ukans.edu` and logging in as `www`. With any terminal, you can telnet to `info.cern.ch` for the simplest interface. These browsers are also available in source and in some cases binary form. Details of status and coordinates of about 20 differ-

ple) the directory structure of an existing file store. This allows existing data to be put “on the Web” without further human effort.

- There is a very extendable system for introducing new formats for multimedia data.
- There are many W3 client programs. As hypertext information is transmitted on the network in logical (mark-up) form, each client can interpret this in a way natural for the

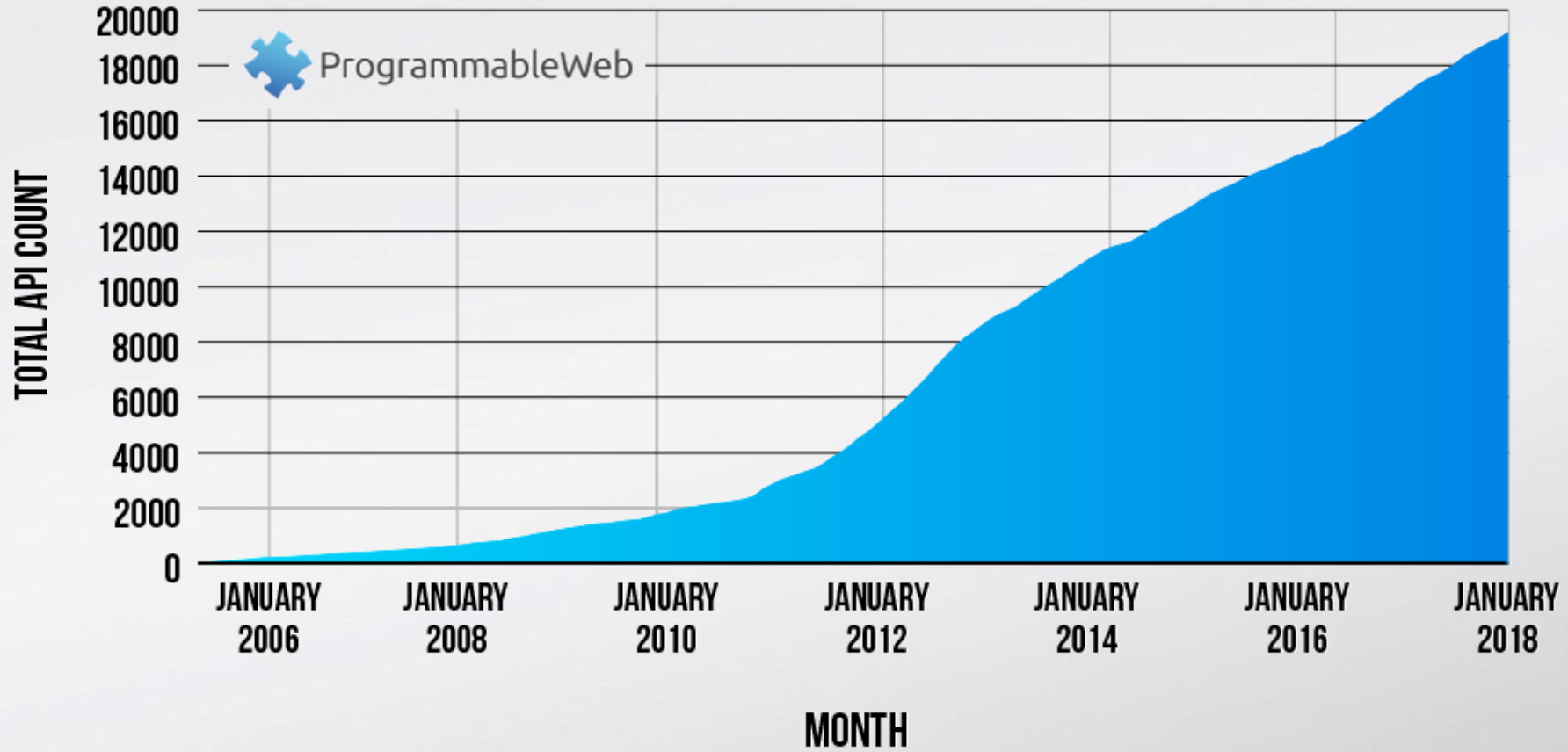
Paris, France, as a CERN technical student. Among other tasks in the CERN W3 team, he currently organizes the cataloging of new W3 material in the “virtual library.” email: `secret@info.cern.ch`

Authors' Present Address: CERN, 1211 Geneva 23, Switzerland.

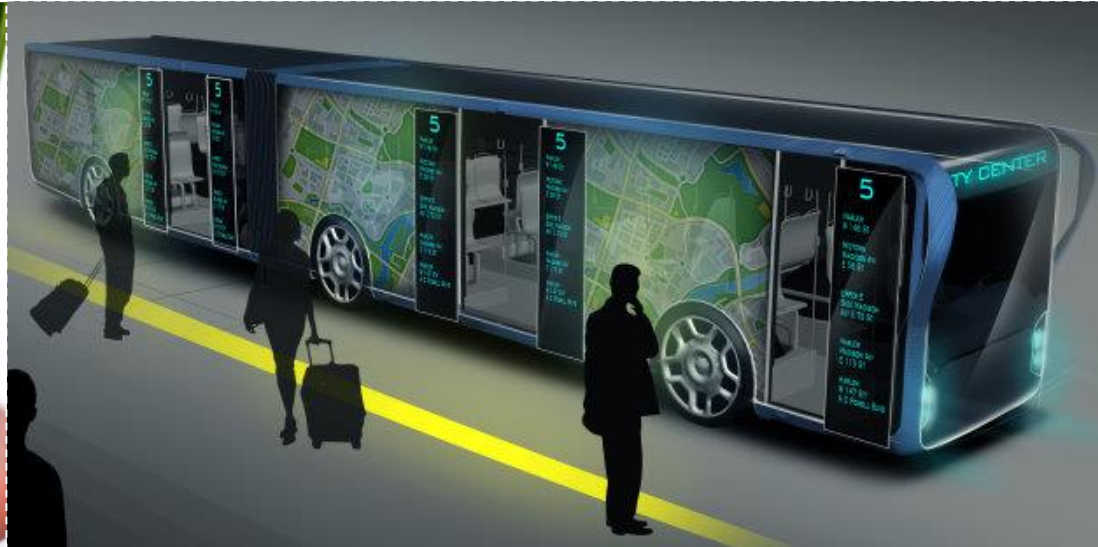
Permission to copy without fee all or part of this material is granted provided that the copies are not made or distributed for direct commercial advantage, the ACM copyright notice and the title of the publication and its date appear, and notice is given that copying is by permission of the Association for Computing Machinery. To copy otherwise, or to republish, requires a fee and/or specific permission.

© ACM 0002-0782/94/0800 \$3.50

GROWTH IN WEB APIS SINCE 2005



Toward a Web of Things

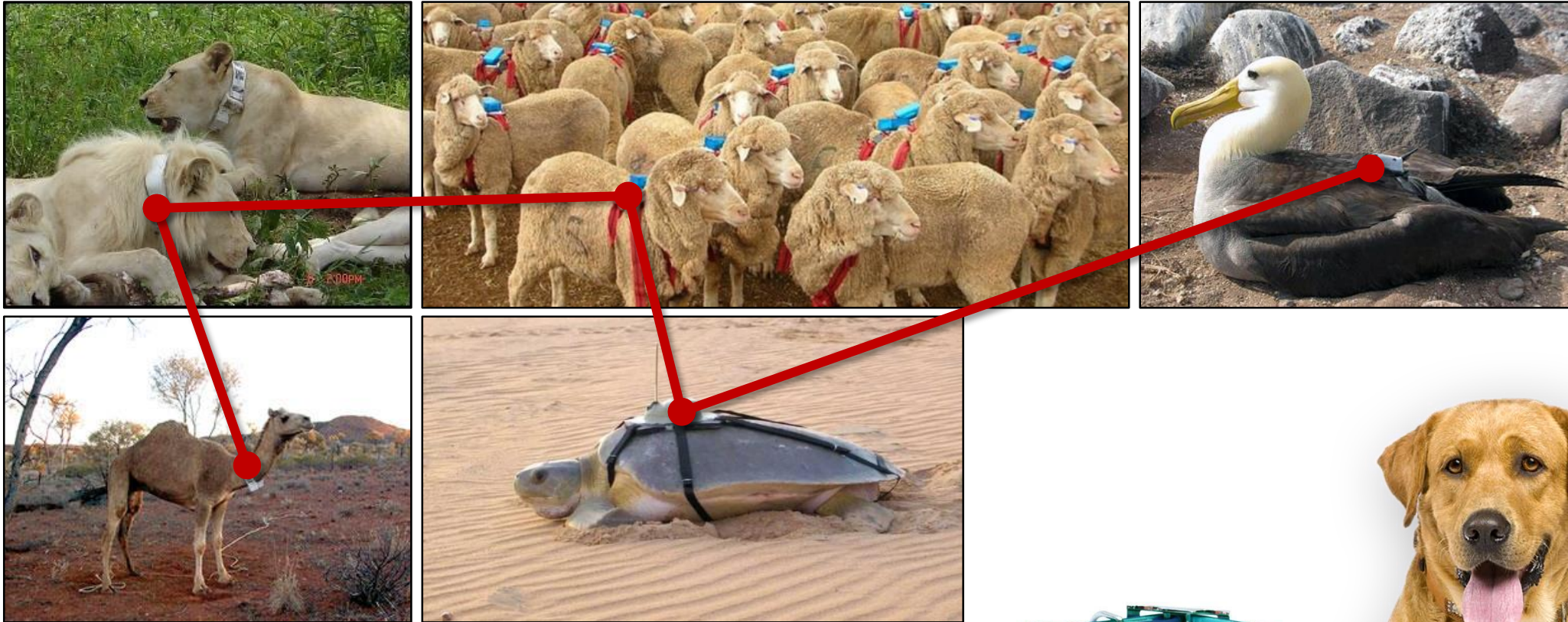




the web as a global blackboard
for artificial intelligence

Connected Animals, Animal-computer interaction (ACI)

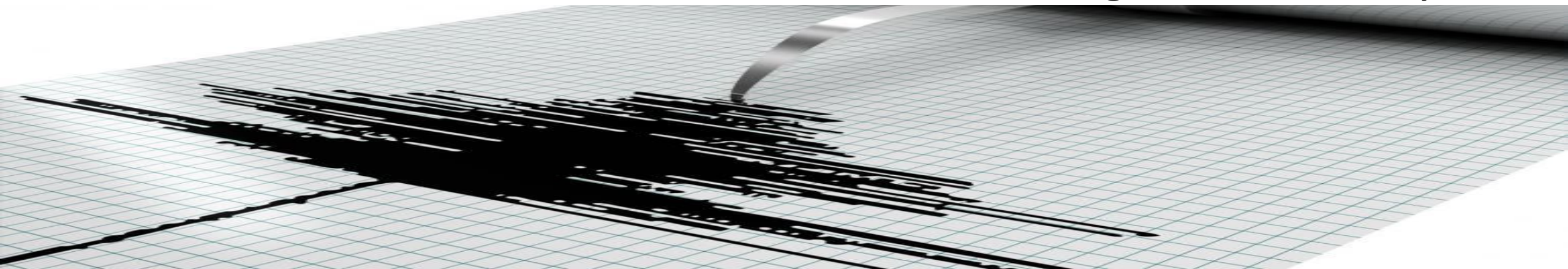
Herdsourcing: monitoring collective animal behavior



Consider a Web of intelligence linking

- Web of Animals: connected **animals** to predict earthquakes
- Web of Things, Web of **Sensors**
- Web of **Als**: reasoning systems, learning systems, etc.
- Web of People: **experts** from all over the world
(i.e. some expert always awake // crowdsourcing & Q&A expert routing)

... to form a collective intelligent decision system.

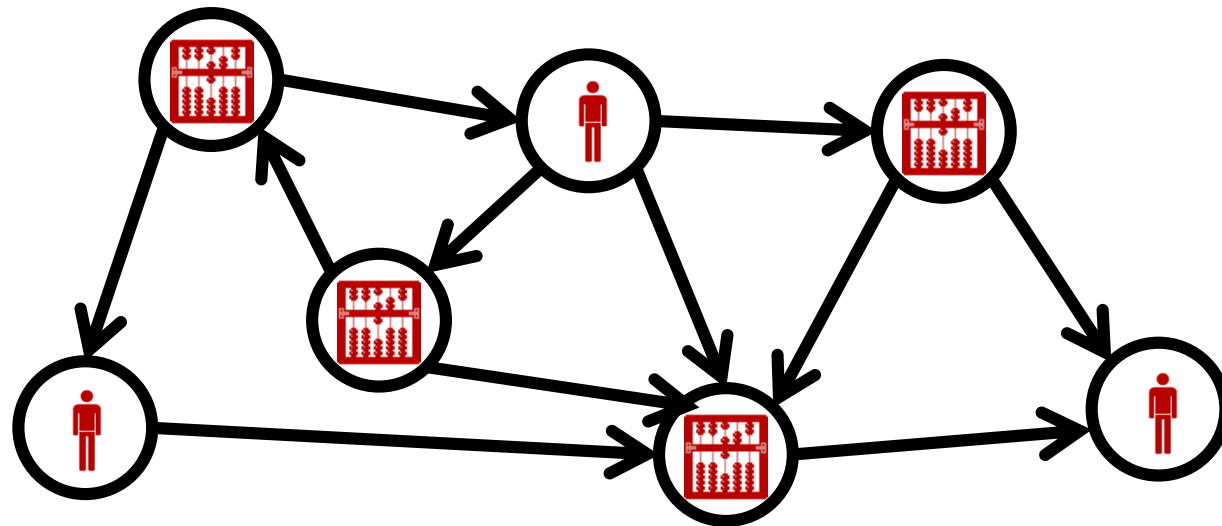


the web as a global blackboard
for ~~artificial~~ intelligence
all kinds of

IMAG_NE
—

IMAG_NE

a Web linking all forms of Intelligence



a Web Science research agenda must account for the fact that the long term potential of the Web is to augment and link everything.



"On the Internet, nobody knows you're a dog."

Peter Steiner, *The New Yorker*, 1993



"Remember when, on the Internet, nobody knew who you were?"

Kaamran Hafeez, *The New Yorker*, 2015



*I pass CAPTCHAs,
Nobody knows I'm a bot*

2019

but...



Energy and Policy Considerations for Deep Learning in NLP

Emma Strubell Ananya Ganesh Andrew McCallum

College of Information and Computer Sciences

University of Massachusetts Amherst

{strubell, aganesh, mccallum}@cs.umass.edu

Abstract

Recent progress in hardware and methodology for training neural networks has ushered in a new generation of large networks trained on abundant data. These models have obtained notable gains in accuracy across many NLP tasks. However, these accuracy improvements depend on the availability of exceptionally large computational resources that necessitate similarly substantial energy consumption. As a result these models are costly to train and develop, both financially, due to the cost of hardware and electricity or cloud compute time, and environmentally, due to the carbon footprint required to fuel modern tensor processing hardware. In this paper we bring this issue to the attention of NLP researchers by quantifying the approximate financial and environmental costs of training a variety of recently successful neural network models for NLP. Based on these findings, we propose actionable recommendations to reduce costs and improve equity in NLP research and practice.

Consumption	CO₂e (lbs)
Air travel, 1 passenger, NY↔SF	1984
Human life, avg, 1 year	11,023
American life, avg, 1 year	36,156
Car, avg incl. fuel, 1 lifetime	126,000
Training one model (GPU)	
NLP pipeline (parsing, SRL)	39
w/ tuning & experimentation	78,468
Transformer (big)	192
w/ neural architecture search	626,155

Table 1: Estimated CO₂ emissions from training common NLP models, compared to familiar consumption.¹

training common NLP models nearly five times the lifetime emissions of the average American car including the manufacturing of the car itself



downscaling AI and Web



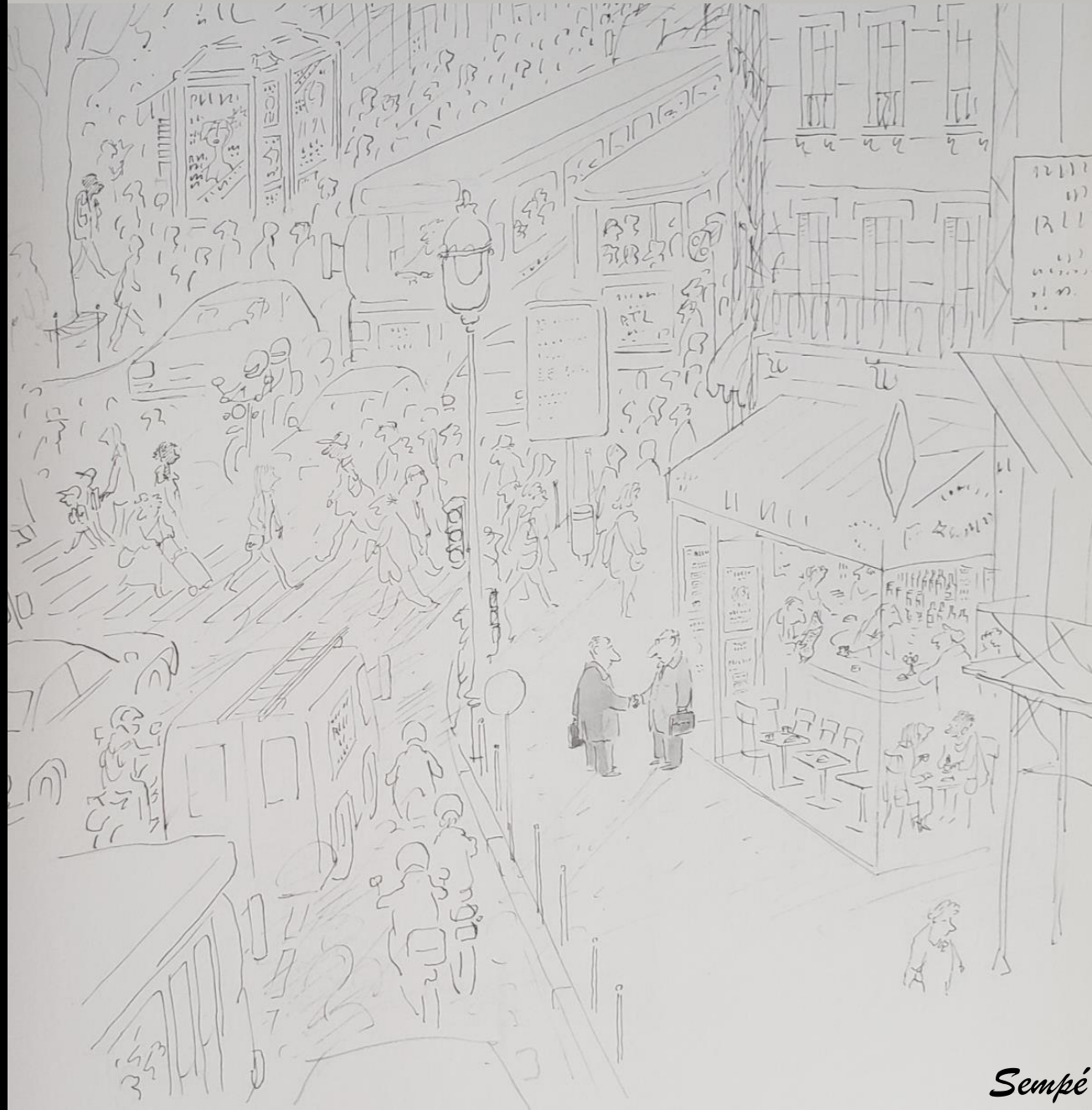
Business Insider, US National Archives, George Grantham Bain Collection

we need historians, sociologists,...

world wide web
www

mmm
massively multidisciplinary method

- Look, we should stay away for a while; you and I are so over-informed that we come out of every conversation exhausted.



Sempé

Make the Web AI-friendly

