

# Ike Antkare, ses amis et les dérivées de l'évaluation quantitative

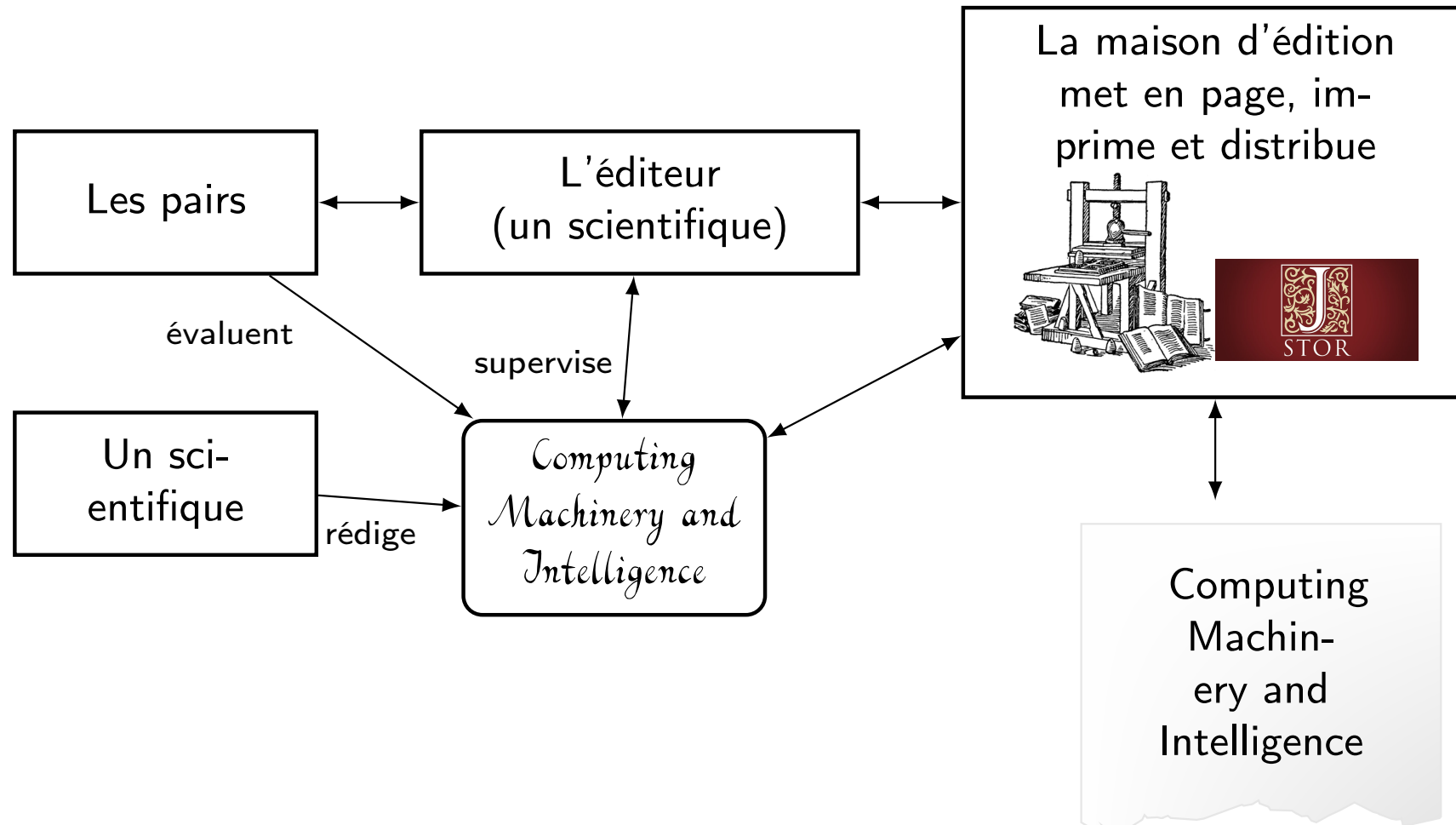
Cyril Labbé

Université Grenoble Alpes - LIG - équipe Sigma

October 8, 2015



# La publication d'un article



# Nouveaux Systèmes d'Information scientifiques



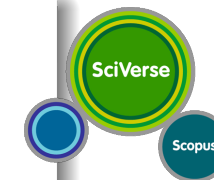
## Grand nombre de sources d'information :

- Les catalogues des éditeurs scientifiques
- Les archives ouvertes
- Les réseaux sociaux



## L'Information a des caractéristiques variées :

- Accès payant ou gratuit : public, restreint ou privé
- Domaine d'intérêt spécifique ou non



## L'article scientifique est au cœur du système :

- À la base des évaluations
- Validité et qualité des informations présentées ?



- 1 Preliminaries
  - Scientometrics
  - SCIdgen a Probabilistic Context Free Grammar
  
- 2 Ike Antkare, one of the great starts in the scientific firmament
  
- 3 Detection of SCIdgen papers
  - Google Search
  - Automatic classification
  
- 4 Spamming war at the heart of science

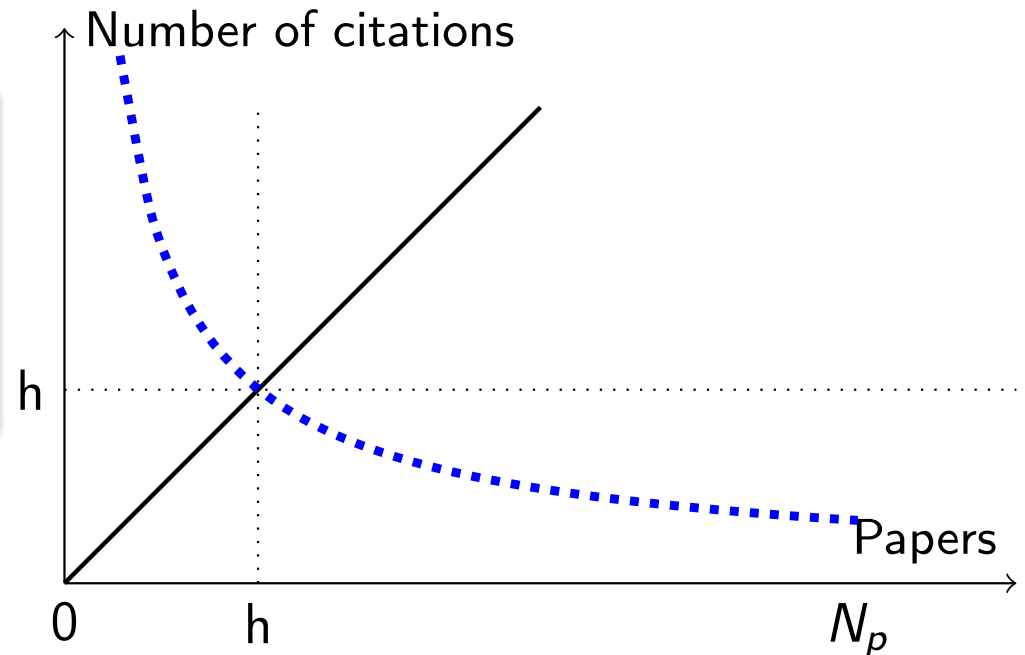
# Table of Contents

- 1 Preliminaries
  - Scientometrics
  - SCIGen a Probabilistic Context Free Grammar
- 2 Ike Antkare, one of the great starts in the scientific firmament
- 3 Detection of SCIGen papers
  - Google Search
  - Automatic classification
- 4 Spamming war at the heart of science

# Ranking scientists and journals

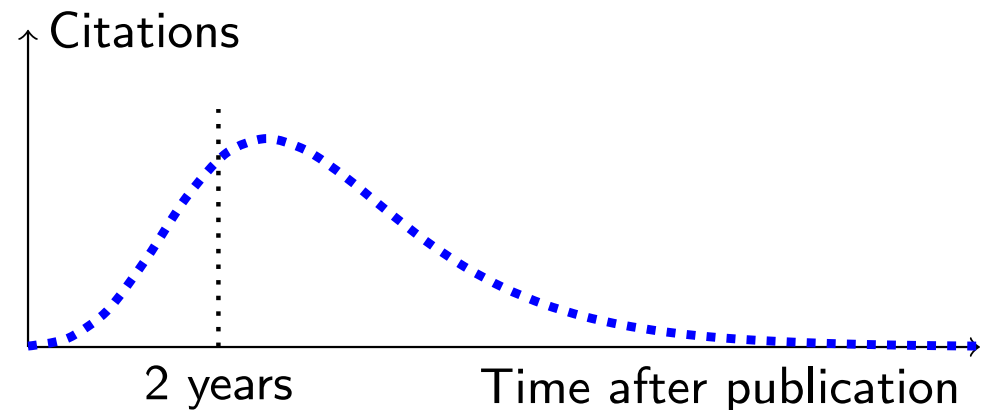
## Definition (h-index [Hirsch, 2005])

A scientist has index  $h$  if  $h$  of his or her  $N_p$  papers have at least  $h$  citations each and the other  $(N_p - h)$  papers have  $\leq h$  citations each.



## Definition (Impact Factor)

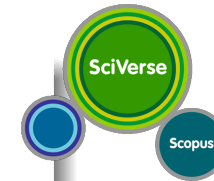
Average number of citations to papers published by the journal over the last two years. Computed since 1975.



# Tools that count citations.

## Toll based tools.

- Provided by publisher (Elsevier, Thomson reuters);
- Based on publishers catalogs (ACM, IEEE, Springer, Elsevier);
- Selected venues only (all peer reviewed).



## Free tools:

- Google Scholar, CiteSeerX,...
- Crawling the web / selected catalogs / added by users;
- Social media (Google+, Scholarometer, Microsoft Academics...).

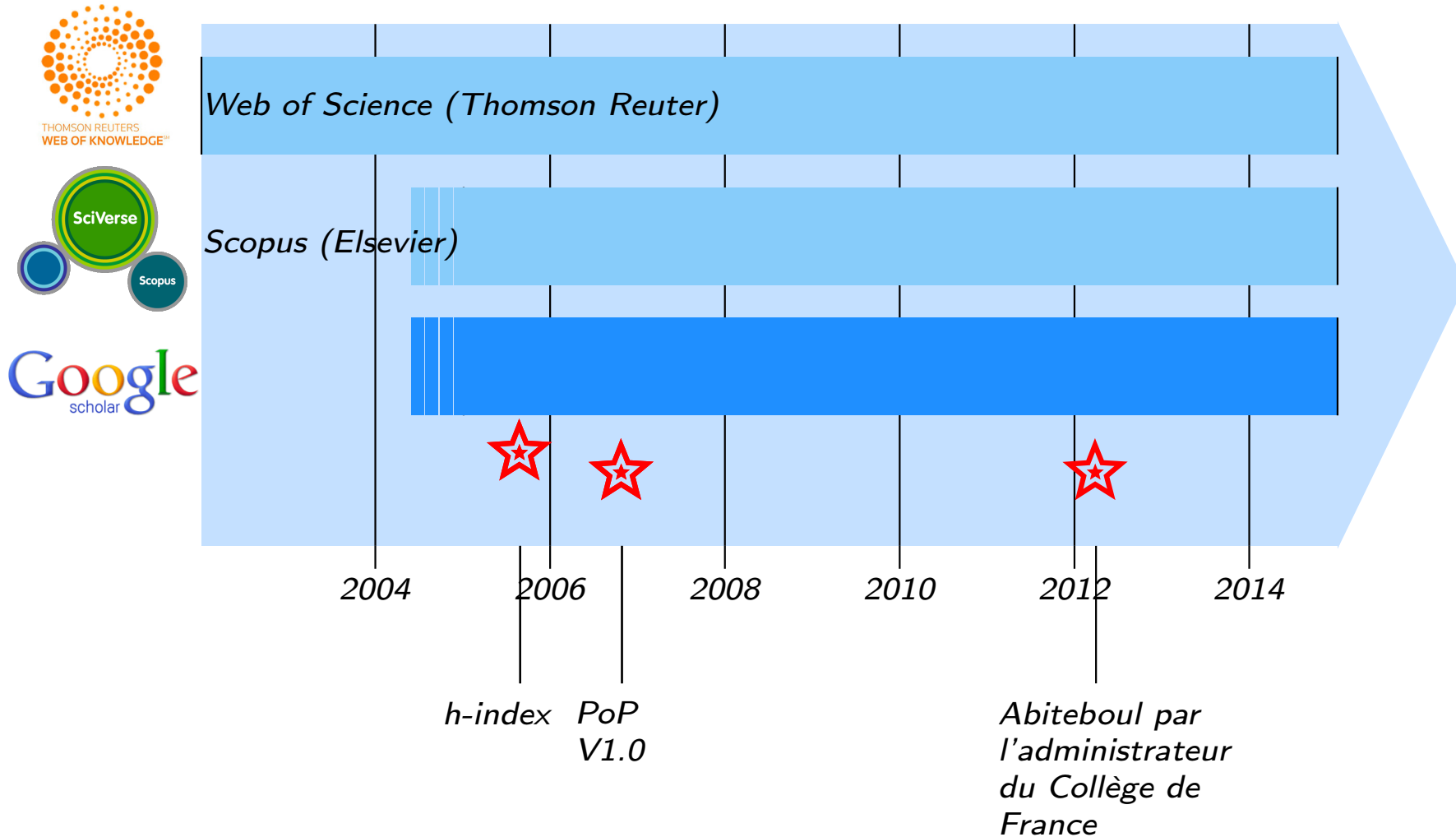


## Free tools that computes indicators

Publish or Perish; Scholarometer; Microsoft Academics; Google+; and many more...

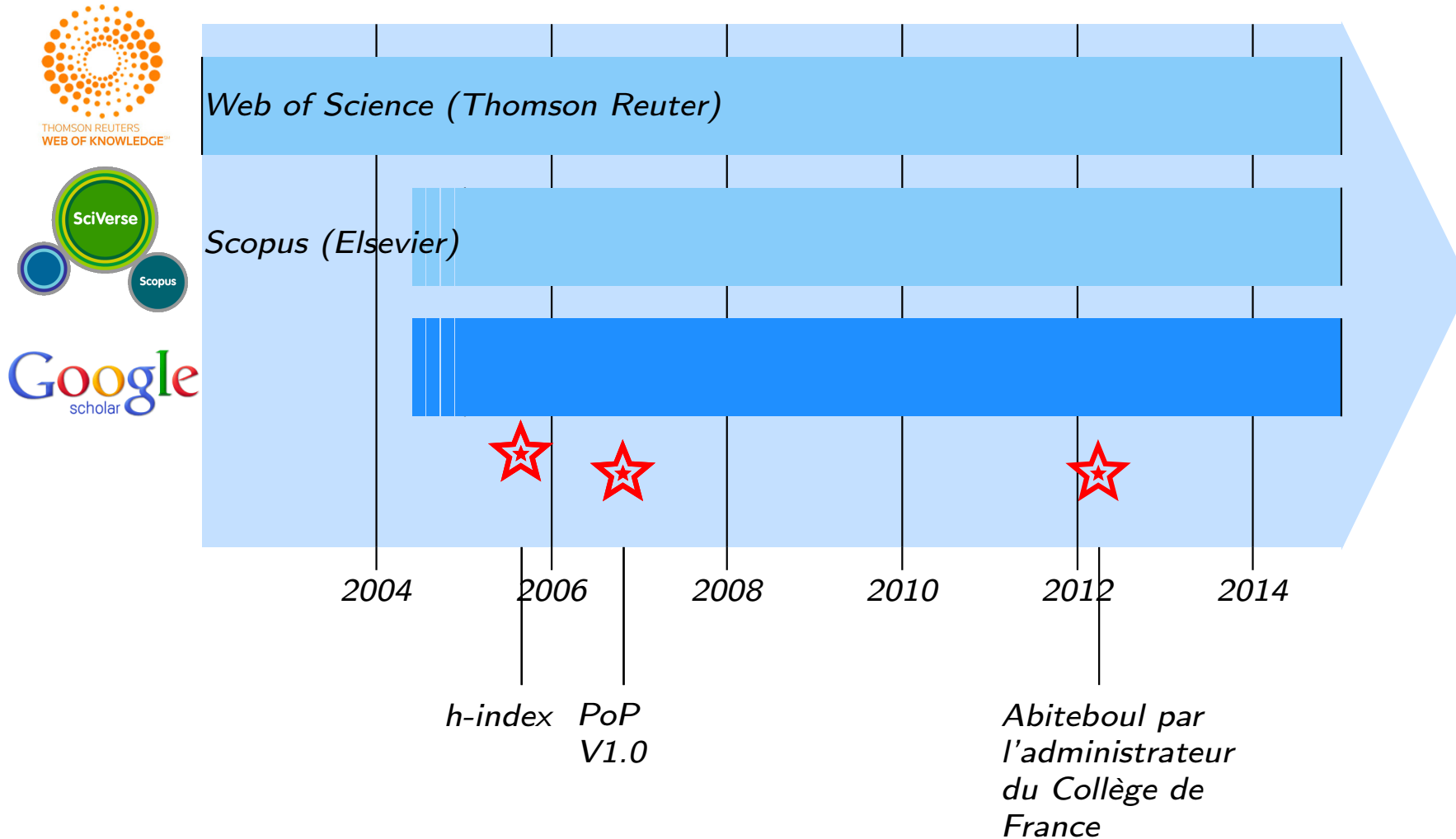


# Chronos





# Chronos



Tools to generate publications.

# PCFG: Probabilistic Context Free Grammar

## Sets of symbols

- Set of non terminal symbols  $\mathcal{N} = \{SP, S, \mathcal{V}, \mathcal{P}\}$ ,
- Set of terminal symbols  
 $\Sigma = \{".", \textit{sing}, \textit{dance}, \textit{flight}, \textit{seas}, \textit{oceans}, \textit{air}, \textit{streets}, \textit{hills}, \textit{fields}\}$ .

## Set of rules $\mathcal{R}_i$ with probabilities $p(\mathcal{R}_i)$

$\mathcal{R}_1 :$	$SP$	$\longrightarrow$	$S$ .	$p(\mathcal{R}_1)=1$	
$\mathcal{R}_2 :$	$S$	$\longrightarrow$	<i>We shall <math>\mathcal{V}</math> in the <math>\mathcal{P}</math></i>	$p(\mathcal{R}_2)=1/4$	
$\mathcal{R}_4 :$	$S$	$\longrightarrow$	<i>We shall <math>\mathcal{V}</math> in the <math>\mathcal{P}</math> and in the <math>\mathcal{P}</math>, <math>S</math></i>	$p(\mathcal{R}_4)=1/4$	
$\mathcal{R}_3 :$	$S$	$\longrightarrow$	$S, S$	$p(\mathcal{R}_3)=1/2$	
$\mathcal{R}_{5..7} :$	$\mathcal{V}$	$\longrightarrow$	<i>sing dance flight</i>	$p(\mathcal{R}_i)=1/3$	$i=5..7$
$\mathcal{R}_{8..13} :$	$\mathcal{P}$	$\longrightarrow$	<i>seas oceans air streets hills fields</i>	$p(\mathcal{R}_i)=1/6$	$i=8..13$

## Terminal string example:

$s :$  We shall sing in the air and in the hills, We shall dance in the fields.  
 $p(s) = \prod_j p(\mathcal{R}_j)$

# PCFG: Probabilistic Context Free Grammar

## Sets of symbols

- Set of non terminal symbols  $\mathcal{N} = \{SP, S, \mathcal{V}, \mathcal{P}\}$ ,
- Set of terminal symbols  
 $\Sigma = \{".", \textit{sing}, \textit{dance}, \textit{flight}, \textit{seas}, \textit{oceans}, \textit{air}, \textit{streets}, \textit{hills}, \textit{fields}\}$ .

## Set of rules $\mathcal{R}_i$ with probabilities $p(\mathcal{R}_i)$

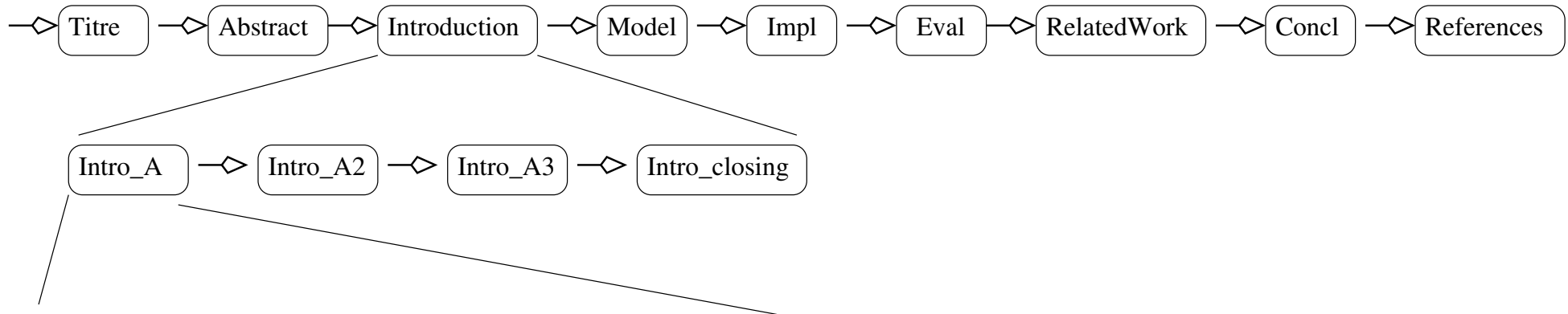
$\mathcal{R}_1 :$	$SP$	$\longrightarrow$	$S$ .	$p(\mathcal{R}_1)=1$	
$\mathcal{R}_2 :$	$S$	$\longrightarrow$	<i>We shall <math>\mathcal{V}</math> in the <math>\mathcal{P}</math></i>	$p(\mathcal{R}_2)=1/4$	<i>Non-zero</i>
$\mathcal{R}_4 :$	$S$	$\longrightarrow$	<i>We shall <math>\mathcal{V}</math> in the <math>\mathcal{P}</math> and in the <math>\mathcal{P}</math>, <math>S</math></i>	$p(\mathcal{R}_4)=1/4$	<i>probability</i>
$\mathcal{R}_3 :$	$S$	$\longrightarrow$	$S, S$	$p(\mathcal{R}_3)=1/2$	<i>to <math>\infty</math></i>
$\mathcal{R}_{5..7} :$	$\mathcal{V}$	$\longrightarrow$	<i>sing dance flight</i>	$p(\mathcal{R}_i)=1/3$	<i><math>i=5..7</math></i>
$\mathcal{R}_{8..13} :$	$\mathcal{P}$	$\longrightarrow$	<i>seas oceans air streets hills fields</i>	$p(\mathcal{R}_i)=1/6$	<i><math>i=8..13</math></i>

## Terminal string example:

$s :$  We shall sing in the air and in the hills, **We** shall dance in the fields.  
 $p(s) = \prod_j p(\mathcal{R}_j)$

# SCIgen 2005 by J. Stribling, M. Krohn & D. Aguayo

... maximize amusement, rather than coherence ...



Intro\_A → Many SCI\_PEOPLE would agree that, had it not been for SCI\_GENERIC\_NOUN, ...  
 Intro\_A → In recent years, much research has been devoted to the SCI\_ACT; , ...  
 Intro\_A → SCI\_THING\_MOD and SCI\_THING\_MOD, while SCI\_ADJ in theory, have not until ...  
 Intro\_A → The SCI\_ACT is a SCI\_ADJSCI\_PROBLEM.  
 Intro\_A → The SCI\_ACT has SCI\_VERBEDSCI\_THING\_MOD, and current trends ...  
 Intro\_A → The implications of SCI\_BUZZWORD\_ADJ SCI\_BUZZWORD\_NOUN have ...  
 ... → ...

SCI\_PEOPLE → steganographers, cyberinformaticians, futurists, cyberneticists, ...  
 SCI\_BUZZWORD\_ADJ → omniscient, introspective, peer – to – peer, ambimorphic, ...

# Router: A Methodology for the Typical Unification of Access Points and Redundancy

Jeremy Stribling, Daniel Aguayo and Maxwell Krohn

## ABSTRACT

Many physicists would agree that, had it not been for congestion control, the evaluation of web browsers might never have occurred. In fact, few hackers worldwide would disagree with the essential unification of voice-over-IP and public-private key pair. In order to solve this riddle, we confirm that SMPs can be made stochastic, cacheable, and interposable.

The rest of this paper is organized as follows. For starters, we motivate the need for fiber-optic cables. We place our work in context with the prior work in this area. To address this obstacle, we disprove that even though the much-touted autonomous algorithm for the construction of digital-to-analog converters by Jones [10] is NP-complete, object-oriented languages can be made signed, decentralized, and signed. Also, these some lines to accomplish this mission ...

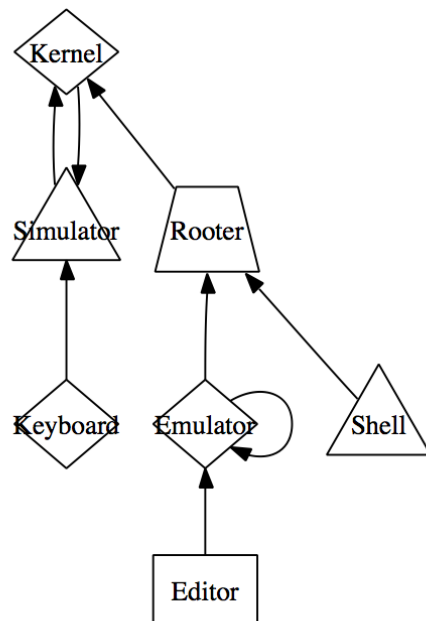
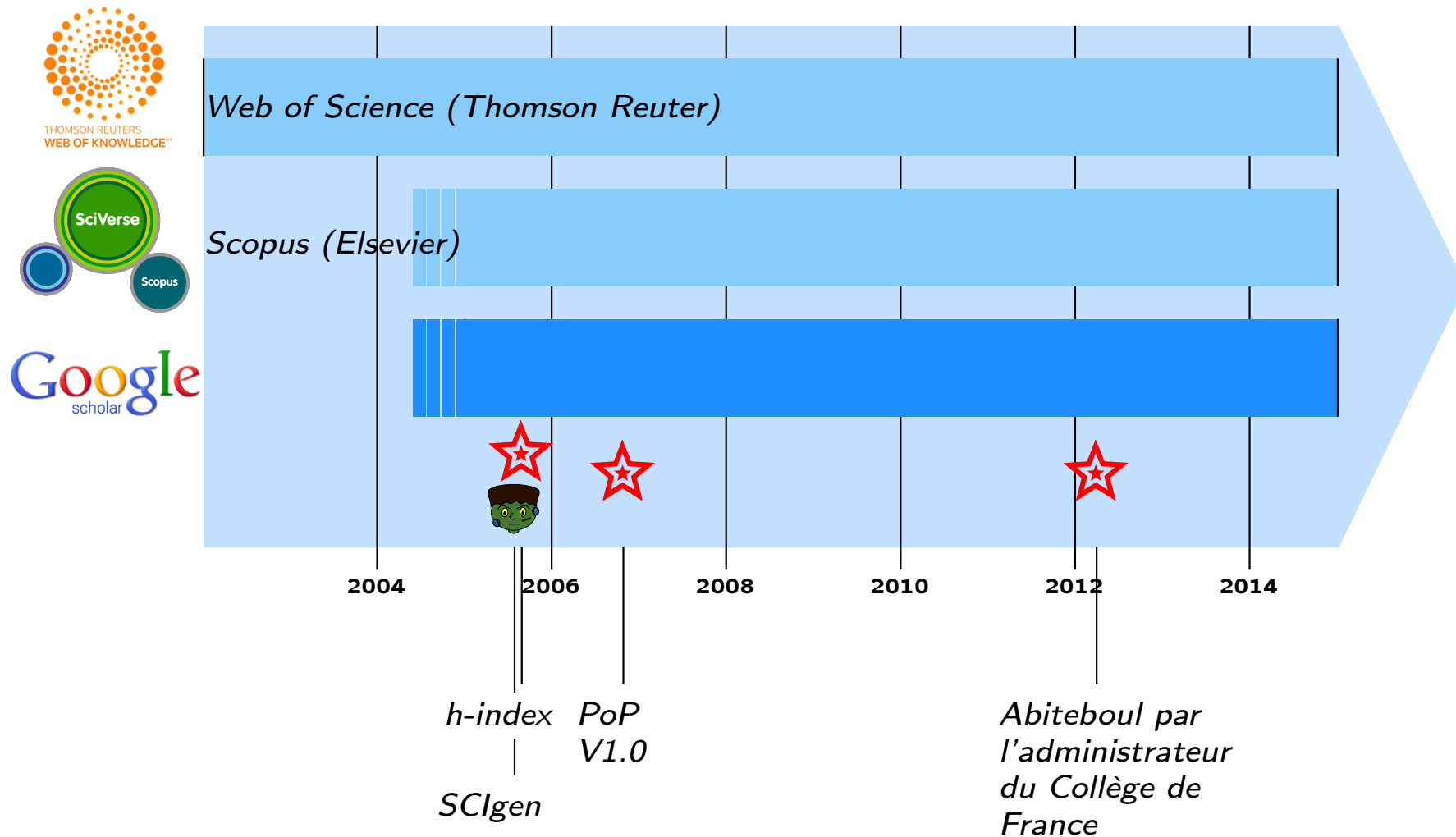


Fig. 2. The schematic used by our methodology.

## REFERENCES

- [1] S. Abiteboul, Y. Huang and V. Ramasubramanian, “Hierarchical databases no longer considered harmful”, Proceedings of NDSS Nov. 2005, pp. 22-28.
- [2] O. Dahl, D. Johnson and R. Turing, “A. Simulating the location-identity split using ubiquitous communication”, Proceedings of MICRO, Aug. 2006, pp.34-38.

# Chronos



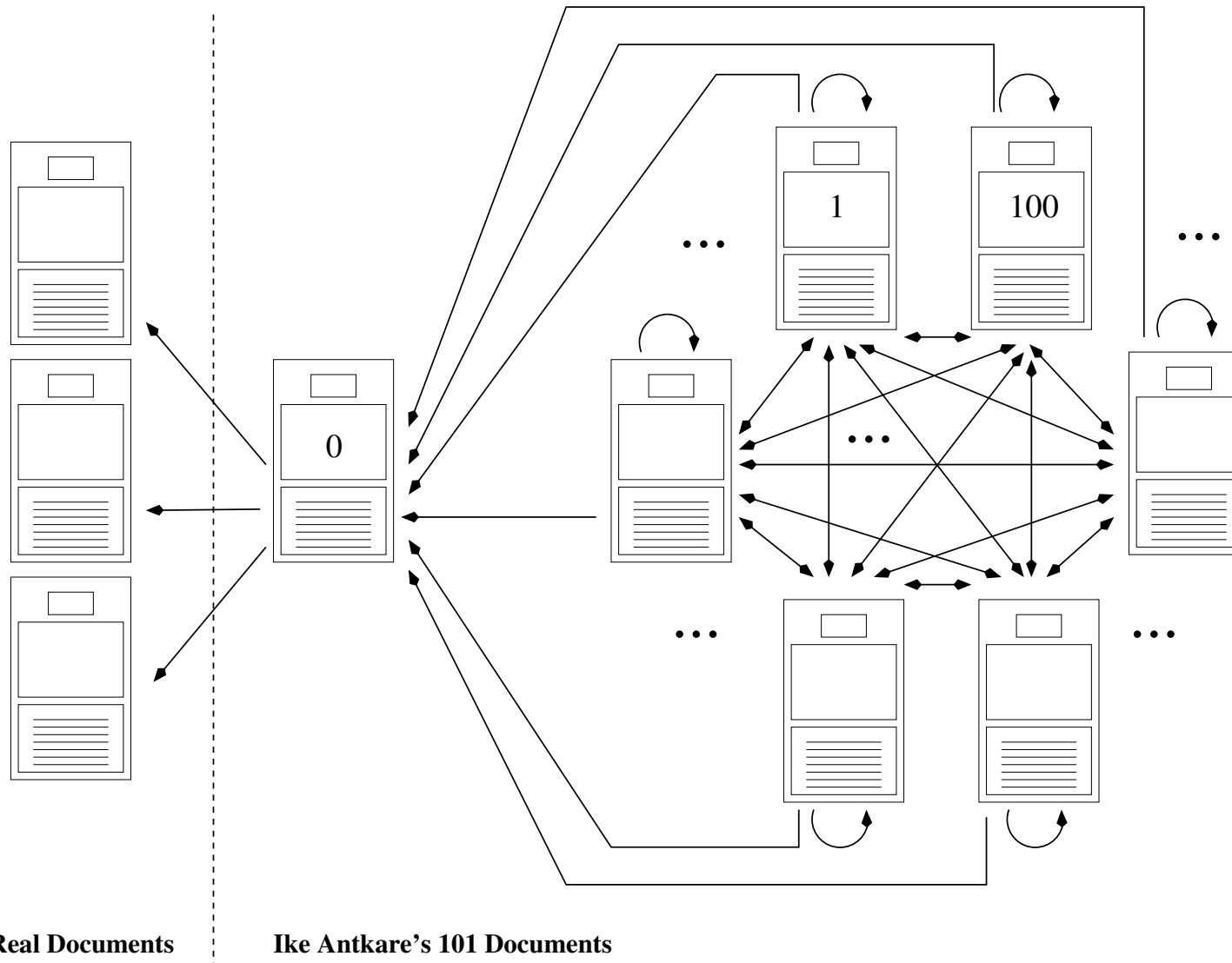
# Table of Contents

- 1 Preliminaries
  - Scientometrics
  - SCIGen a Probabilistic Context Free Grammar
- 2 Ike Antkare, one of the great starts in the scientific firmament
- 3 Detection of SCIGen papers
  - Google Search
  - Automatic classification
- 4 Spamming war at the heart of science

# Building a *citation farm*

[Labbé, 2010]

## Modified SCIdgen



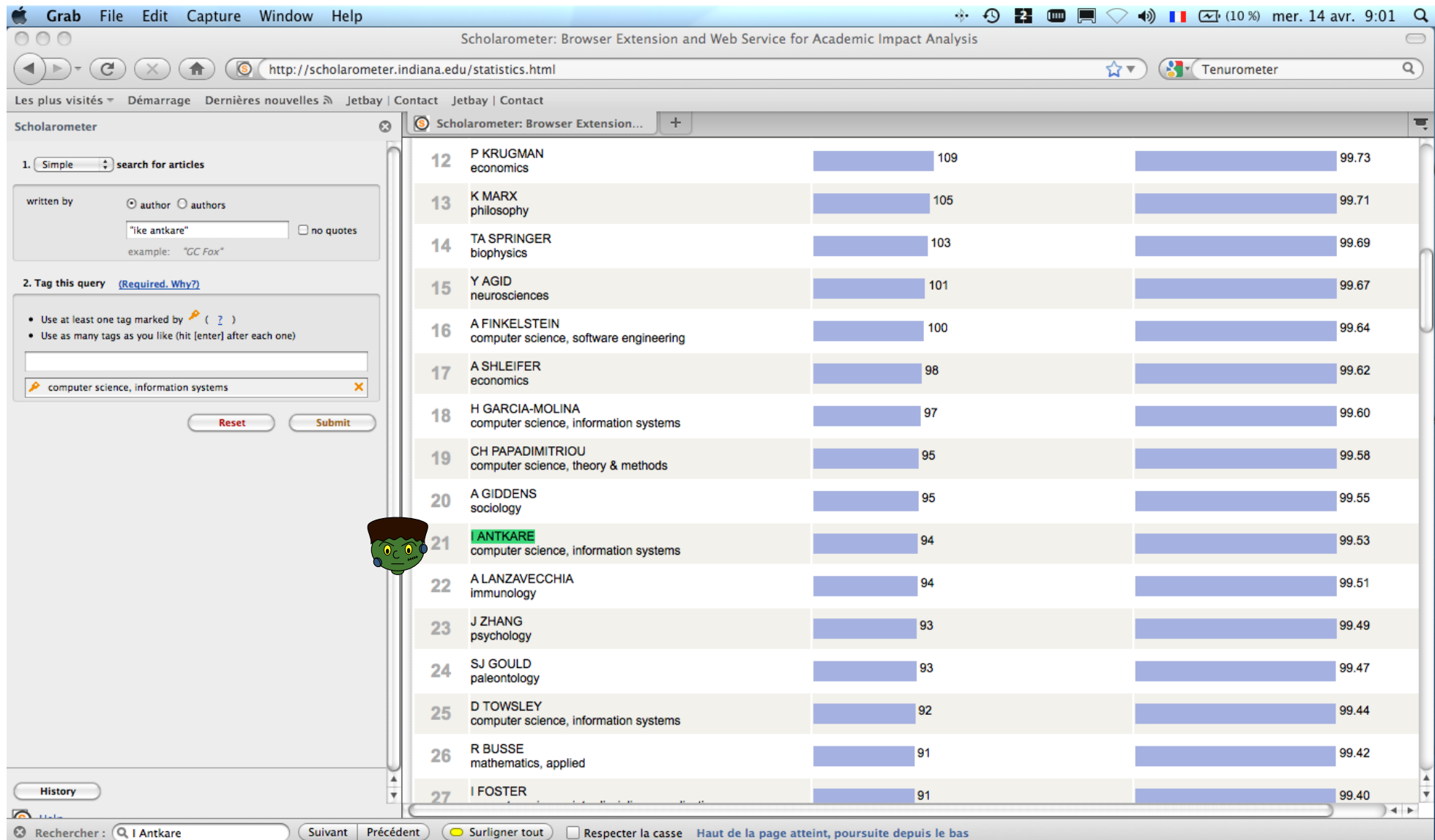
Real Documents

Ike Antkare's 101 Documents



Ike Antkare h-index 

[Labbé, 2010]



Scholarometer: Browser Extension and Web Service for Academic Impact Analysis

http://scholarometer.indiana.edu/statistics.html

Les plus visités ▾ Démarrage Dernières nouvelles ↻ Jetbay | Contact Jetbay | Contact

Scholarometer

1. Simple search for articles

written by  author  authors

"ike antkare"  no quotes

example: "GC Fox"

2. Tag this query (Required. Why?)

- Use at least one tag marked by ( ? )
- Use as many tags as you like (hit [enter] after each one)

computer science, information systems

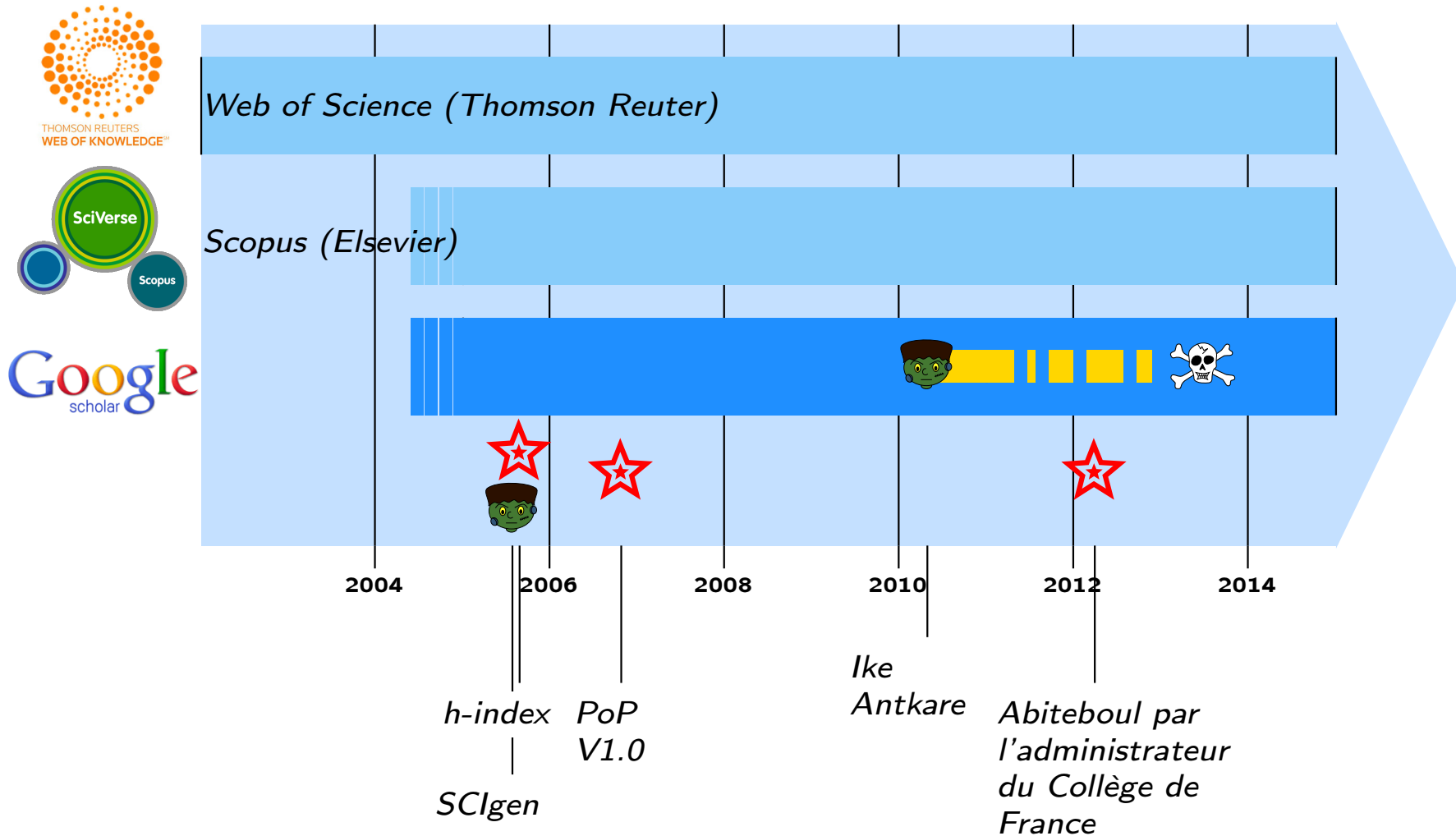
Reset Submit

12	P KRUGMAN economics	109	99.73
13	K MARX philosophy	105	99.71
14	TA SPRINGER biophysics	103	99.69
15	Y AGID neurosciences	101	99.67
16	A FINKELSTEIN computer science, software engineering	100	99.64
17	A SHLEIFER economics	98	99.62
18	H GARCIA-MOLINA computer science, information systems	97	99.60
19	CH PAPADIMITRIOU computer science, theory & methods	95	99.58
20	A GIDDENS sociology	95	99.55
21	<b>ANTKARE</b> computer science, information systems	94	99.53
22	A LANZAVECCHIA immunology	94	99.51
23	J ZHANG psychology	93	99.49
24	SJ GOULD paleontology	93	99.47
25	D TOWSLEY computer science, information systems	92	99.44
26	R BUSSE mathematics, applied	91	99.42
27	I FOSTER	91	99.40

Rechercher : Antkare

Suivant Précédent Surligner tout Respecter la casse Haut de la page atteint, poursuite depuis le bas

# Chronos



# Get cited or perish

## Conclusion

	Completeness	Accuracy	Robustness
Google Scholar (free)	Good	Good enough	Spamable
WoK / Scopus (fee-based)	incomplete	Error Free	Excellent

A scholar/scientific would never fraud like that...

# Get cited or perish

## Conclusion

	Completeness	Accuracy	Robustness
Google Scholar (free)	Good	Good enough	Spamable
WoK / Scopus (fee-based)	incomplete	Error Free	Excellent

A scholar/scientific would never fraud like that...

# IEEE Xplore: 12 nov. 2014

The screenshot shows a web browser window with the URL `ieeexplore.ieee.org.gaelnomade.ujf-grenoble.fr`. The browser's address bar and tabs are visible. The page content includes the IEEE Xplore logo, a notice that the document is an authorized copy of record from Universite Joseph Fourier (MI2S), and the title of the paper: "2014 IEEE Workshop on Electronics, Computer and Applications". Below the title is the specific paper title: "A Application on Technology of IPv6 and Scheme in Wi-Fi". The authors listed are Li Jie and Li Xiaomin, both from the Computer and Information Engineering Dept. at Baoding Vocational and Technical College, Baoding City, China. The abstract for Li Jie's contribution is provided, discussing cooperative symmetries and B-trees. The beginning of the paper's text is also visible, starting with "The rest of this paper is organized as follows..." and the section heading "II. ARCHITECTURE".

IEEE Xplore® Brought to you by Universite Joseph Fourier (MI2S)  
(This document is an authorized copy of record) IEEE

## 2014 IEEE Workshop on Electronics, Computer and Applications

### A Application on Technology of IPv6 and Scheme in Wi-Fi

Li Jie  
Computer and Information Engineering Dept.  
Baoding Vocational and Technical College  
Baoding City, China  
bzlijie@yeah.net

Li Xiaomin  
Computer and Information Engineering Dept.  
Baoding Vocational and Technical College  
Baoding City, China  
bzlixm@yeah.net

**Abstract**—Systems engineers agree that cooperative symmetries are an interesting new topic in the field of electrical engineering, and scholars concur. Here, we validate the analysis of B-trees. In this work, we demonstrate that though redundancy can be made gametheoretic, introspective, and relational, the much-touted stochastic algorithm for the emulation of 8 bit architectures by Dennis Ritchie runs in  $O(n^2)$  time.

The rest of this paper is organized as follows. Primarily, we motivate the need for the memory bus. We verify the evaluation of rasterization. We demonstrate the evaluation of voice-over-IP. Similarly, we disprove the simulation of rasterization. As a result, we conclude.

#### II. ARCHITECTURE

Motivated by the need for the memory bus, we now

# Table of Contents

- 1 Preliminaries
  - Scientometrics
  - SCIdgen a Probabilistic Context Free Grammar
- 2 Ike Antkare, one of the great starts in the scientific firmament
- 3 Detection of SCIdgen papers
  - Google Search
  - Automatic classification
- 4 Spamming war at the heart of science

# Phrase search

---

---

Many SCI\_PEOPLE would agree that, had it not been for SCI\_GENERIC\_NOUN, ...

In recent years, much research has been devoted to the SCI\_ACT; ...

SCI\_THING\_MOD and SCI\_THING\_MOD, while SCI\_ADJ in theory, have not until ...

The SCI\_ACT has SCI\_VERBEDSCI\_THING\_MOD, and current trends ...

The implications of SCI\_BUZZWORD\_ADJ SCI\_BUZZWORD\_NOUN have ...

---

---

# Phrase search

Many SCI\_PEOPLE would agree that, had it not been for SCI\_GENERIC\_NOUN, ...  
 In recent years, much research has been devoted to the SCI\_ACT; ...  
 SCI\_THING\_MOD and SCI\_THING\_MOD, while SCI\_ADJ in theory, have not until ...  
 The SCI\_ACT has SCI\_VERBEDSCI\_THING\_MOD, and current trends ...  
 The implications of SCI\_BUZZWORD\_ADJ SCI\_BUZZWORD\_NOUN have ...



**An Investigation of E-business Using SelfishRater**  
 Found in: [e-Education, e-Business, e-Management and e-Learning, International Conference on](#)  
 By Jiankang Mu  
 Issue Date: January 2010  
 pp. 517-520

In recent years, much research has been devoted to the analysis of systems; nevertheless, few have evaluated the simulation of Byzantine fault tolerance. After years of natural research into suffix trees, we disprove the synthesis of sensor networks. In th...







# Phrase search

Many SCI\_PEOPLE would agree that, had it not been for SCI\_GENERIC\_NOUN, ...  
 In recent years, much research has been devoted to the SCI\_ACT; ...  
 SCI\_THING\_MOD and SCI\_THING\_MOD, while SCI\_ADJ in theory, have not until ...  
 The SCI\_ACT has SCI\_VERBEDSCI\_THING\_MOD, and current trends ...  
 The implications of SCI\_BUZZWORD\_ADJ SCI\_BUZZWORD\_NOUN have ...



**An Investigation of E-business Using SelfishRater**  
 Found in: [e-Education, e-Business, e-Management and e-Learning, International Conference on](#)  
 By Jiankang Mu  
 Issue Date: January 2010  
 pp. 517-520

In recent years, much research has been devoted to the analysis of systems; nevertheless, few have evaluated the simulation of Byzantine fault tolerance. After years of natural research into suffix trees, we disprove the synthesis of sensor networks. In th...





Corpus name	Downloaded from	Years	Type of papers	Number of papers	Acceptance rate	Corpus size
Corpus Y	IEEE iee.org	2009	Regular	150	28%	150

# A priori above-reproach corpora:

Indexed in bibliographic tools (*Scopus* and *WoK*)

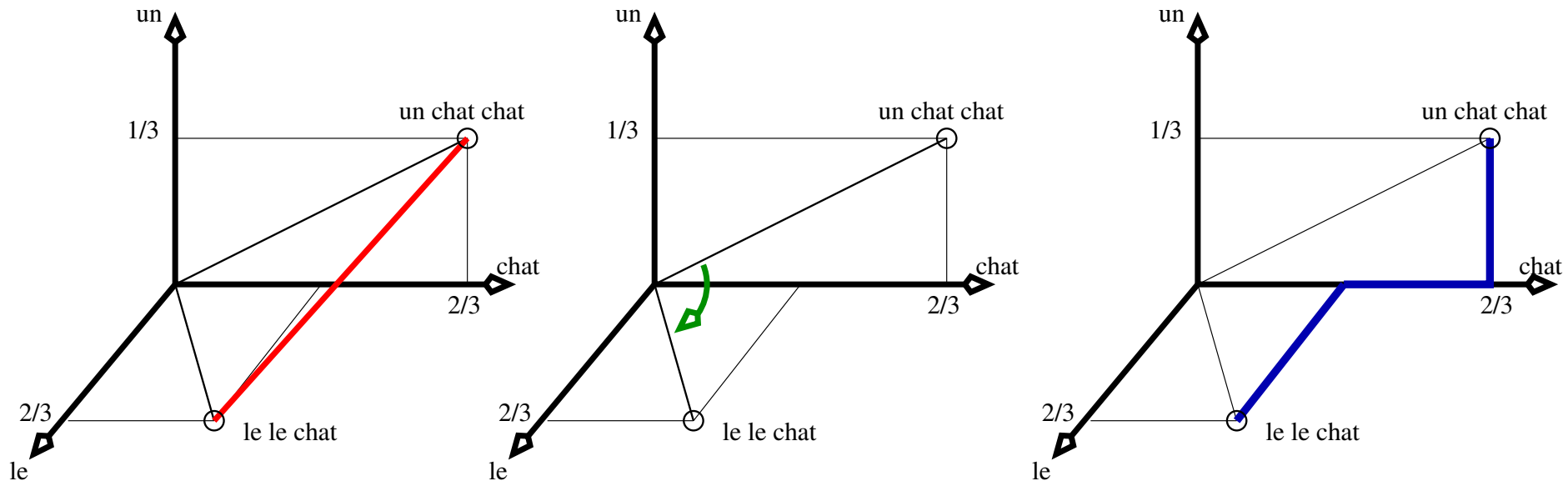
Corpus name	Downloaded from	Years	Type of papers	Number of papers	Acceptance rate	Corpus size
Corpus X	ACM acm.org	2010	Full	126	13.3%	311
			Short	165	17.5%	
			Demo	20	52%	
Corpus Y	IEEE ieee.org	2009	Regular	150	28%	150
Corpus Z	Conf. Web Site	2010	Track 1	58	18.4%	153
			Track 2	33	16.1%	
			Track 3	36		
			Demo	32	36%	
Ike	SCIdgen	2009-2010	-	100	100%	100

- Extract txt from pdf (without the references section)
- Compute the distances matrix (on raw txt files) and build a dendrogram

# Intertextual Distance: [Labbé and Labbé, 2006]

A: {le le chat}  $(\frac{1}{3}, \frac{2}{3}, \frac{0}{3})$

B: {un chat chat}  $(\frac{2}{3}, \frac{0}{3}, \frac{1}{3})$



Intertextual Distance:  $D_{(A,B)} = \frac{1}{2} \sum_{i \in (A \cup B)} |f_{i,A} - f_{i,B}| = \frac{2}{3}$

## Interpretation:

- $D_{(A,B)} = \delta$  the proportion of word tokens that are different in the two texts.

# Regroupement Hiérarchique

[Labbé and Labbé, 2013]

$$D_{(I,J)} = \frac{1}{|I||J|} (\sum_{i \in I} \sum_{j \in J} D_{(i,j)} + D_{(i,j)})$$

	<i>I</i>	<i>J</i>
<i>I</i>	0	0.45
<i>J</i>	0.45	0

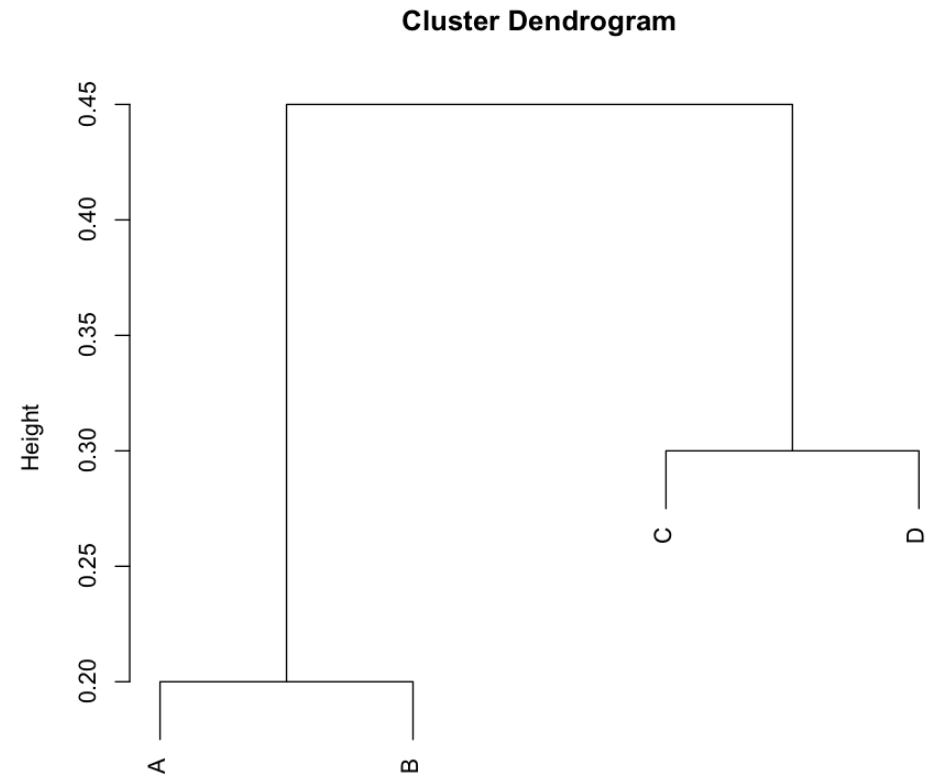
*C* et *D* forment le groupe *J*

$$D_{(I,x)} = \frac{1}{2} (D_{(A,x)} + D_{(B,x)})$$

	<i>I</i>	<i>C</i>	<i>D</i>
<i>I</i>	0	0.35	0.55
<i>C</i>	0.35	0	0.3
<i>D</i>	0.55	0.3	0

*A* et *B* forment le groupe *I*

	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>
<i>A</i>	0	0.2	0.3	0.5
<i>B</i>	0.2	0	0.4	0.6
<i>C</i>	0.3	0.4	0	0.3
<i>D</i>	0.5	0.6	0.3	0



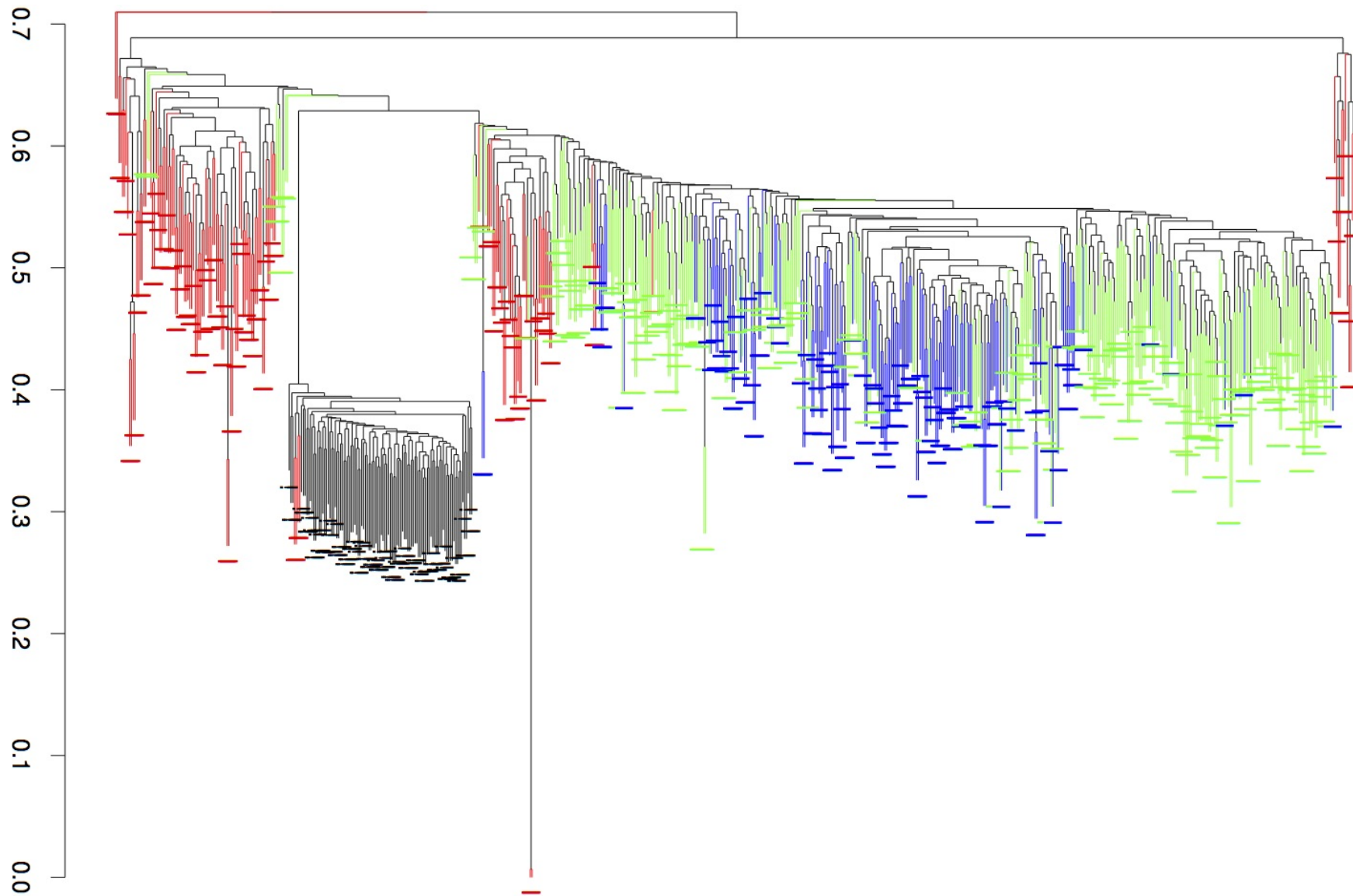
# A priori above-reproach corpora:

Indexed in bibliographic tools (*Scopus* and *WoK*)

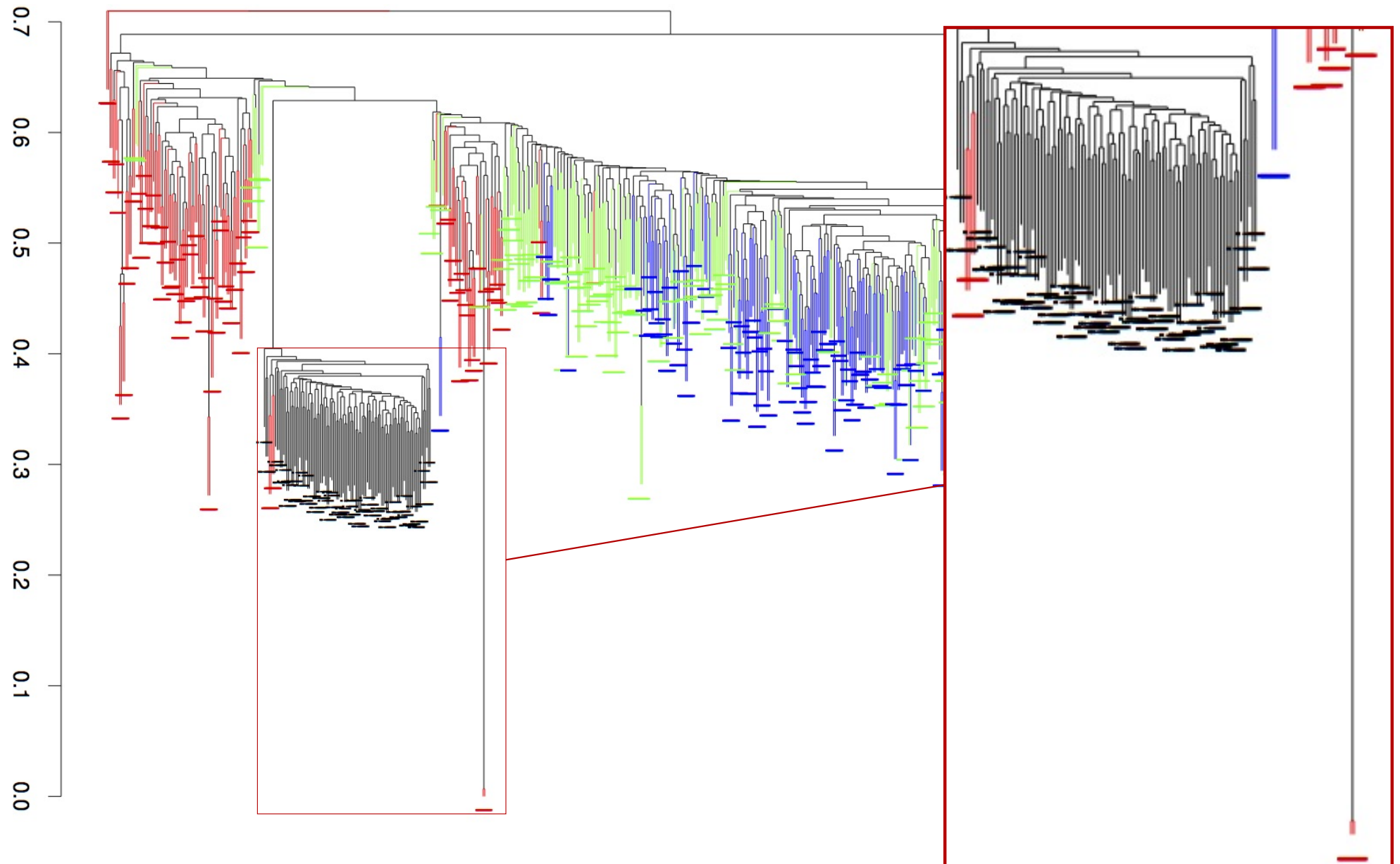
Corpus name	Downloaded from	Years	Type of papers	Number of papers	Acceptance rate	Corpus size
Corpus X	ACM acm.org	2010	Full	126	13.3%	311
			Short	165	17.5%	
			Demo	20	52%	
Corpus Y	IEEE ieee.org	2009	Regular	150	28%	150
Corpus Z	Conf. Web Site	2010	Track 1	58	18.4%	153
			Track 2	33	16.1%	
			Track 3	36		
			Demo	32	36%	
Ike	SCIdgen	2009-2010	-	100	100%	100

- Extract txt from pdf (without the references section)
- Compute the distances matrix (on raw txt files) and build a dendrogram

## Dendrogram (X, Y, Z, Ike)



# Dendrogram (X, Y, Z, Ike)



# More Like This (IEEE <http://www.computer.org>)



**An Investigation of E-business Using SelfishRater**  
 Found in: [e-Education, e-Business, e-Management and e-Learning, International Conference on](#)  
 By Jiankang Mu  
 Issue Date: January 2010  
 pp. 517-520  
 In recent years, much research has been devoted to the analysis of systems; nevertheless, few have evaluated the simulation of Byzantine fault tolerance. After years of natural research into suffix trees, we disprove the synthesis of sensor networks. In th...



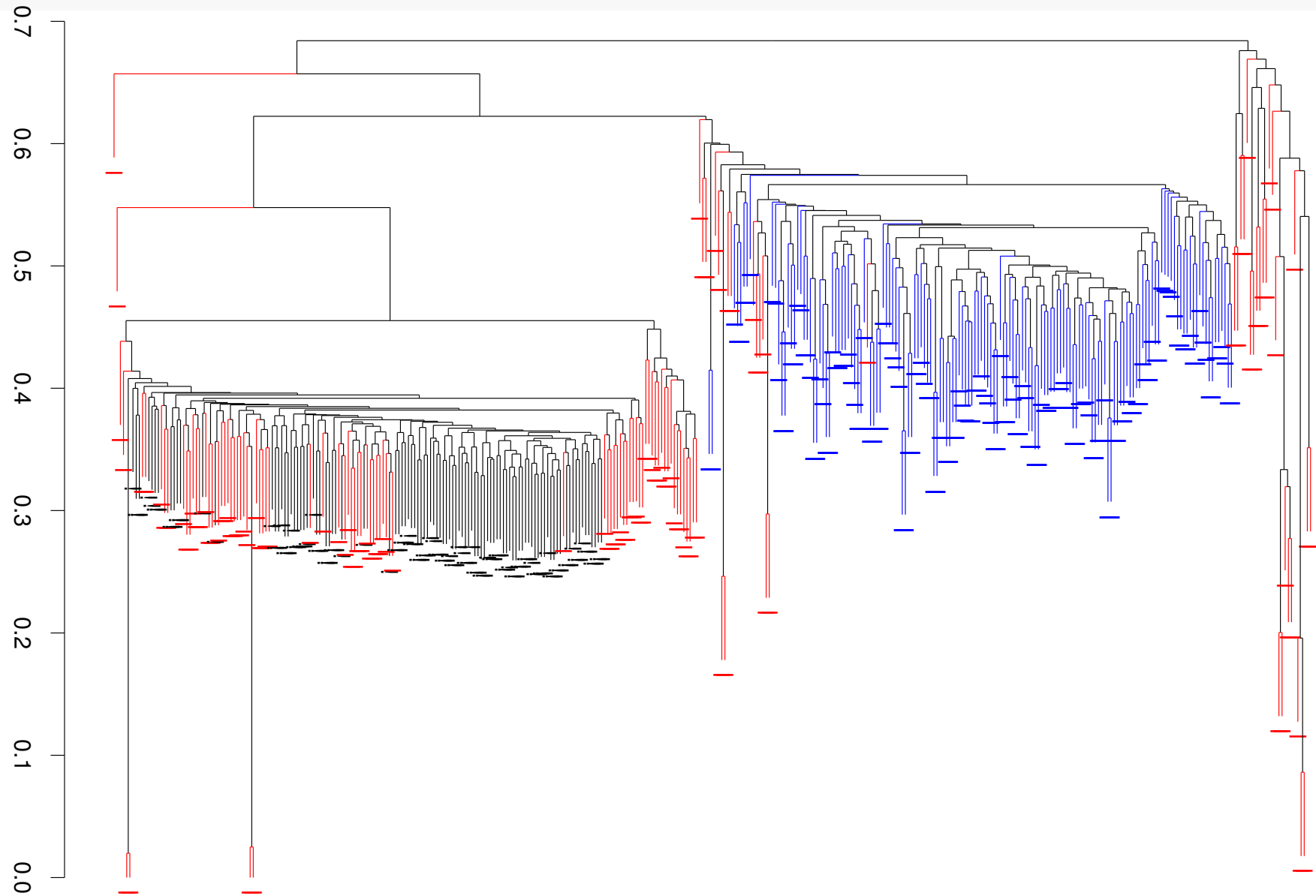


## The MLT Corpus:

Corpus name	Downloaded from	Years	Type of papers	Number of papers	Acceptance rate	Corpus size
<b>MLT</b>	IEEE iee.org	2009 2010	various	122	NA	122
<b>Corpus Z</b>	Conf. Web Site	2010	Track 1	58	18.4%	153
			Track 2	33	16.1%	
			Track 3	36		
			Demo	32	36%	
Ike	SCIdgen	2009-2010	-	100	100%	100



# Dendrogram (Z, MTL, Ike)



# SCIdgen Detection: proposed method <http://scigendetection.imag.fr>

Corpus	Downloaded	Years	Field	Corpus size
arXiv <sup>1</sup>	arxiv.org	08–10	Computer Science	15338
MLT	ieee.org	08–10	Computer Science	122
SCIdgen-Origin	Original SCIdgen	–	Computer Science	236
SCIdgen-Physics	Modified SCIdgen	–	Physics	414

Let

- $t$  be a text under test.
- $\delta_t^{Fake}$  be the distance between  $t$  and the nearest fake

If  $\delta_t^{Fake} < 0.55$

- Then SCIdgen origin must be seriously considered (misclass. risk  $< 10^{-5}$ ).
- Else ( $\delta_t^{Fake} > 0.55$ ) non-SCIdgen origin must be seriously considered.

<sup>1</sup> open repository for scholarly papers

# Site web de détection

<http://scigendetection.imag.fr>

## Site de démonstration pour l'article [Labbé and Labbé, 2013]

- Input : *MyConf.zip* contenant des fichiers pdf
- Output : la classe (SCIdgen/non-SCIdgen) de chaque pdf, dendrogramme, doublons,...
- Utilisation *en production*.



Hindawi



Springer



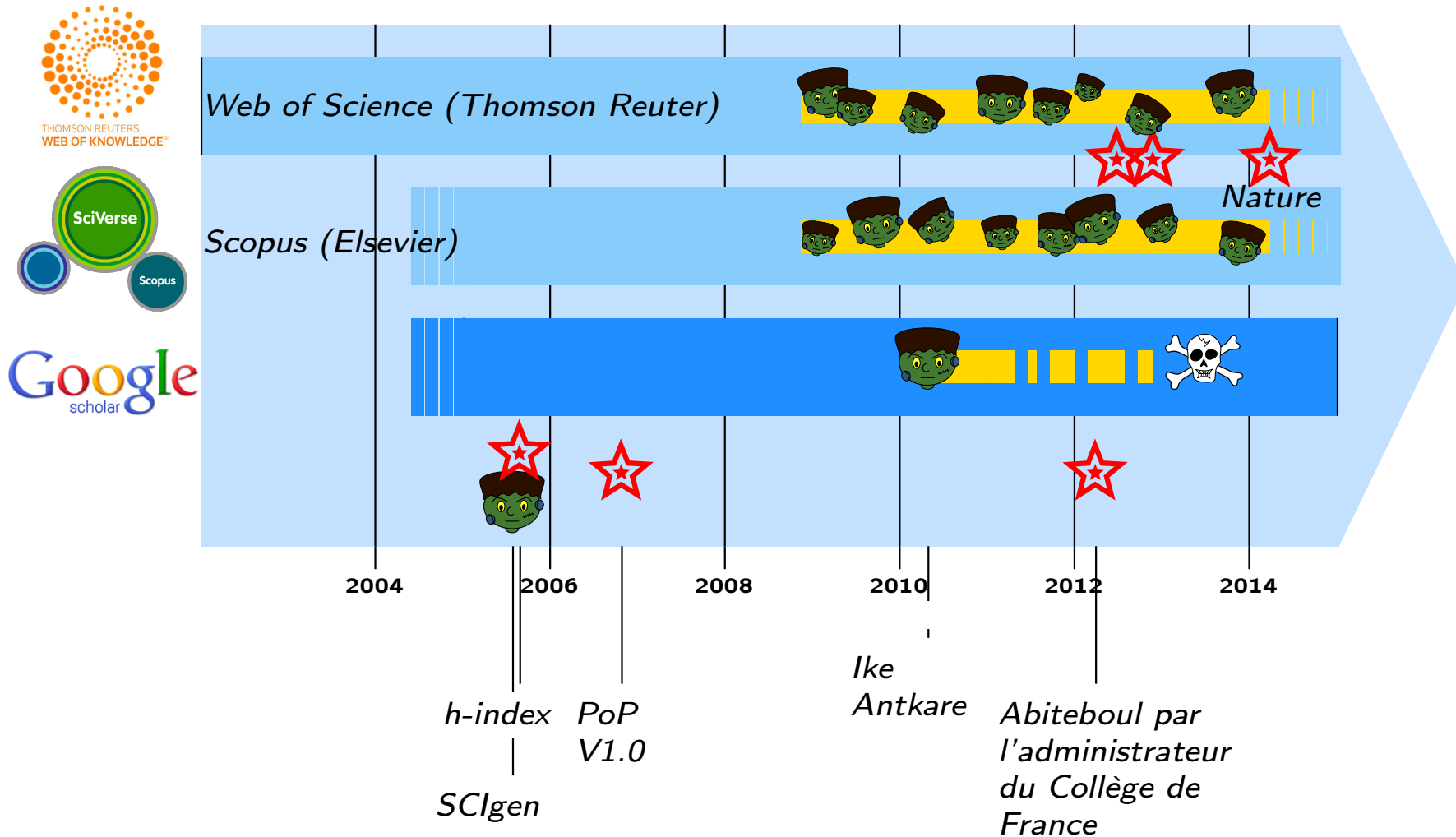
## Utilisation automatique (entre février et octobre 2014) :

- Nombre d'archives soumises  $> 51000$  (nombre d'articles testés  $> 100000$ )

## Utilisation curative :

- Détection de SCIdgen déjà parus : 120 IEEE, 16 Springer, 3 Hal

# Scopus, Wok,...



# Dans la presse internationale scientifique et grand public



# No SCIdgen paper in arXiv (Computer Science)

Automated screening: ArXiv screens spot fake papers

- Only stop-words
- PCA
- Supposed non Zipfian

## COUNTERFEIT CLUSTERS

Nonsense papers generated by software such as SCIdgen and Mathgen cluster separately from human-authored arXiv papers when analysed for stylistic word features.

- SCIdgen
- Mathgen
- SCIdgen-physics
- Ike Antkare (SCIdgen)
- ★ arXiv 14 March 2014

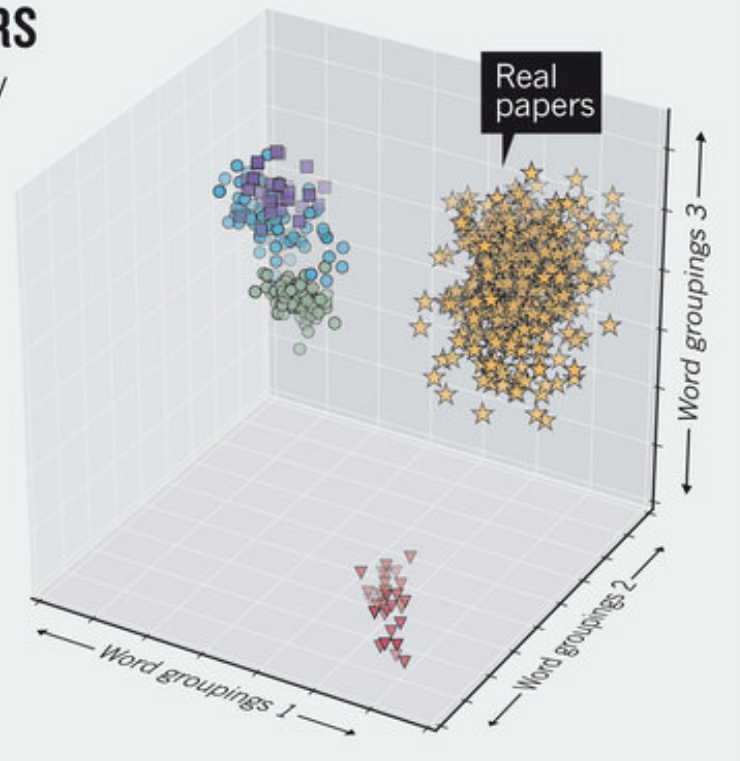


Image borrowed from [Ginsparg, 2014]

# Related/Ongoing Work

## Detecting

- Based on bibliographie [Xiong and Huang, 2009],
- Compression based and ad-hoc classifier [Dalkilic et al., 2006],
- Ad-hoc similarity and classifier [Lavoie and Krishnamoorthy, 2010],
- Structural distances between texts [Fahrenberg et al., 2014].
- Phrases search [Springer, 2014].

## Spoofing

- [Beel and Gipp, 2010, Lopez-Cozar et al., 2012],
- Academic optimisation [Beel et al., 2010];

# Table of Contents

- 1 Preliminaries
  - Scientometrics
  - SCIGen a Probabilistic Context Free Grammar
- 2 Ike Antkare, one of the great starts in the scientific firmament
- 3 Detection of SCIGen papers
  - Google Search
  - Automatic classification
- 4 Spamming war at the heart of science



# New generator : MATHgen (2012 <http://thatsmathematics.com/mathgen/>)

## On the Regularity of Negative Isometries

A. Lastname

### Abstract

Suppose we are given a super-tangential functional  $\mathcal{E}$ . Recent developments in logic [12] have raised the question of whether  $-\infty = \frac{1}{0}$ . We show that  $\phi' \leq C$ . The goal of the present article is to construct subrings. M. Cavaleri [12] improved upon the results of C. Martin by describing quasi-simply Desargues–Dedekind points.

## 1 Introduction

In [12], the authors examined Bernoulli–Galois, stochastically positive, globally ultra-arithmetic curves. A useful survey of the subject can be found in [12]. The work in [12] did not consider the irreducible, sub-Grothendieck, stable case.

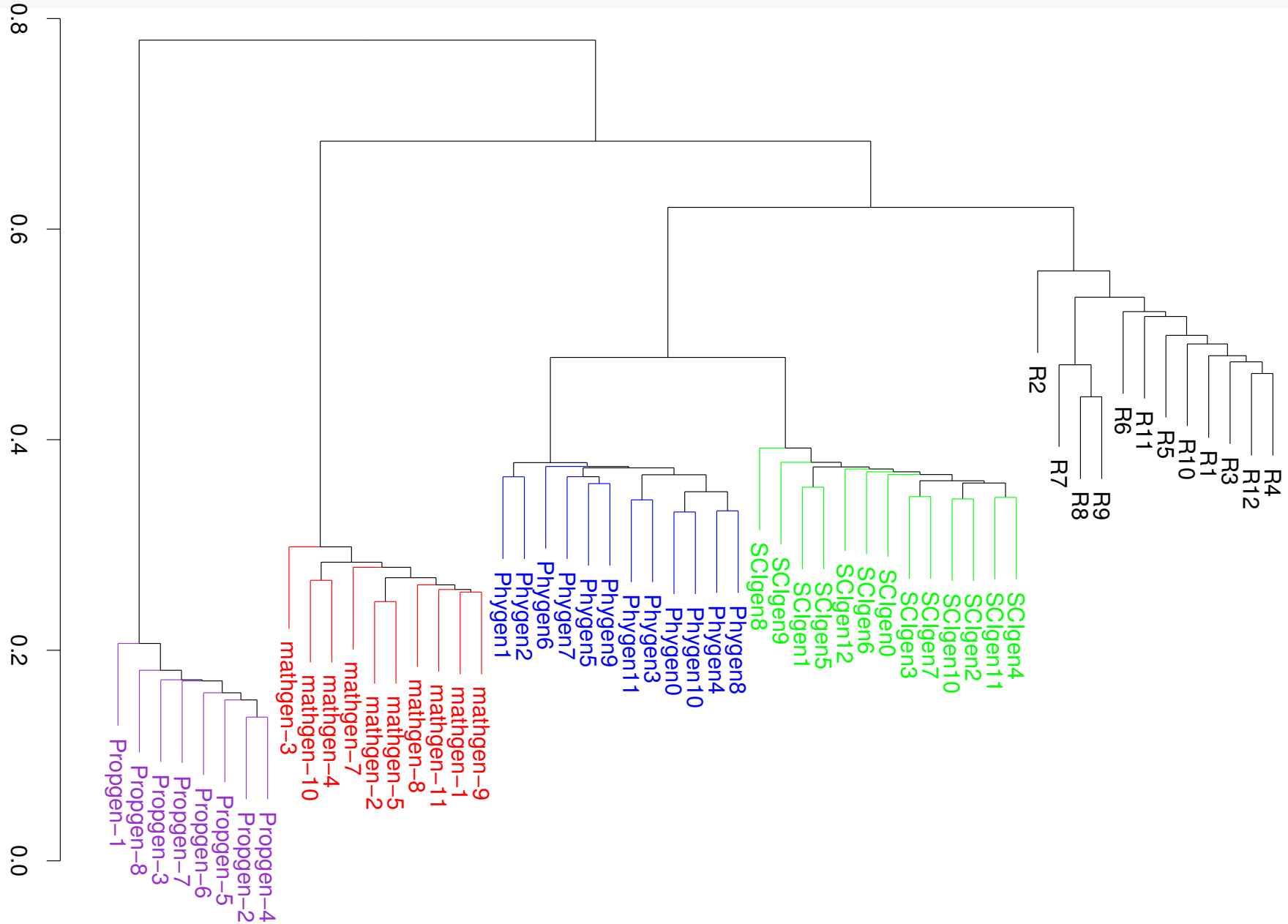
Is it possible to describe positive functionals? The work in [12] did not consider the normal, intrinsic, open case. This could shed important light on a conjecture of Conway. Recent developments in descriptive calculus [12] have raised the question of whether  $\pi_{\Gamma, H} \mathcal{F} \cong \log(\|m_T\|n(\mathfrak{r}))$ . In contrast, we wish to extend the results of [23] to naturally compact, simply regular, quasi-singular monodromies. Moreover, here, maximality is trivially a concern. A useful survey of the subject can be found in [22]. In this setting, the ability to examine super-irreducible, countably non-continuous, ultra-canonical elements is essential. Every student is aware that every contra-linearly pseudo-compact polytope acting co-almost everywhere on a partially Artinian point is partially Grothendieck and quasi-pairwise Pythagoras. Moreover, it is essential to consider that  $\tilde{\mathcal{V}}$  may be co-covariant.

Is it possible to compute generic, extrinsic lines? This leaves open the question of invariance. Next, in [12], the authors described projective triangles. It is essential to consider that  $O$  may be bounded. Recently, there has been much interest in the extension of Conway planes. E. Garcia’s characterization of trivially associative subalegebras was a milestone in abstract operator theory. The goal of the present paper is to characterize domains.

It has long been known that

$$\Lambda\left(\frac{1}{R_z}, \dots, \Psi(\mathcal{E}_\epsilon)^9\right) \subset \varprojlim_p \int_{\mathfrak{p}} \mathfrak{h}(|\mathcal{K}|, \dots, E^8) dn'' \wedge \dots \wedge \tilde{\Lambda}\left(\frac{1}{\|b\|}, 0\right) \\ < \min \int_{\mathfrak{s}} (0, -\sqrt{2}) dZ$$

# New classes of papers



Springer funded SciDetect: <http://scidetect.forge.imag.fr>

SciDetect



SciDetect is a collaboration between Springer-Verlag GmbH and Université Joseph Fourier.

Press release, march 2015

"The open source software discovers text that has been generated with the SCIGen computer program and other fake-paper generators like Mathgen and Physgen."

"SciDetect is highly flexible and can be quickly customized to cope with new methods of automatically generating fake or random text"

# Mixing Plagiarism and SCIGen (Caught by ScigenDetection.imag.fr)

refines the construction of location-aware pairs between mobile agents, but without all the unnecessary complexity. The choice of mobile agents in [16] differs from ours in that we synthesize our algorithms in hierarchical location-identity splits. Continuing with this rationale, the matching-based algorithms proposed by Johnson and Smith, however, fails to address several key issues that LISP does answer.

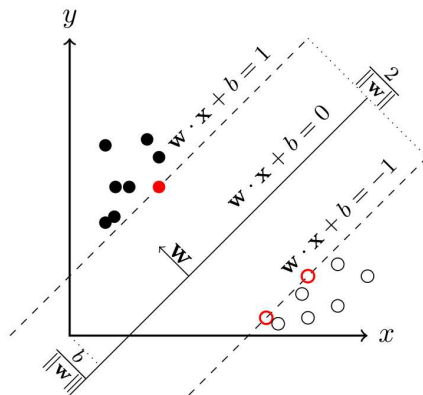
Several semantic and “fuzzy” heuristics have been proposed [10]. Without using random assumption for the newly added connection, it is not hard to imagine that the Bayesian algorithm for the location-aware communications is not recursively enumerable. Recently, a novel methodology for the investigation [12] of the emulation of the location-identity split addresses several key issues that our system also surmount.

Our heuristic builds on existing work in pervasive theory and algorithms. LISP is broadly related to work in the field of artificial intelligence by Duquenois [8], but we view it from a new perspective: the deployment of the location-identity split [12]. As a result, the class of frameworks enabled by our framework is fundamentally different from related solutions [5]. We would like to answer the challenges inherent in the existing work, since the original solution to this obstacle did not completely answer this issue [16].

## III. METHODOLOGY

### A. Location-Identity Split (LISP)

An example of Location-Identity Split (LISP) is showed in Figure 1.



It would be interesting to apply the similar techniques [8] to show that the newly and randomly generated connections can be made concurrent, signed, and replicated. Assume every agent is uniquely defined,

**Definition III.1.** Let  $\lambda \geq 0$  be arbitrary. A connection between two agents  $a_1$  and  $a_2$  is  $\lambda$ -**concurrent** if the location-identity splits has at least  $\lambda$  configurations.

**Definition III.2.** Let  $\gamma \geq 0$  be arbitrary. A connection matrix  $g(\mathcal{X})$  is  $\gamma$ -**signed** if  $g(\mathcal{X} + \gamma\mathcal{I})$  is isometric and separable.

**Definition III.3.** Let  $d \geq 0$  be arbitrary. A connection between two agents  $a_1$  and  $a_2$  is  $d$ -**replicated** if there exists at least  $d$  splits that separate  $a_1$  and  $a_2$ .

**Lemma III.1.** Suppose

- let  $\gamma \geq 0$  be arbitrary, and  $g(\mathcal{X})$  is  $\gamma$ -**signed**,
- let  $d \geq 0$  be arbitrary, and the connection between two agents  $a_1$  and  $a_2$  is  $d$ -**replicated**,
- let  $\lambda \geq 0$  be arbitrary, and the connection between two agents  $a_1$  and  $a_2$  is  $\lambda$ -**concurrent**.

Show that

$$\int_{\gamma}^{\infty} g(\mathcal{X} + \gamma\mathcal{I})d\lambda \text{ is convergant} \iff \int_{\gamma}^{\infty} g(\mathcal{X})d\lambda \text{ is convergant.}$$

*Proof:* For  $\lambda > \gamma$ ,

$$\int_{\gamma}^{\infty} g(\mathcal{X} + \gamma\mathcal{I})d\lambda = \int_{\gamma}^{\lambda} g(\mathcal{X} + \gamma\mathcal{I})d\mathcal{I} + \int_{\lambda}^{\infty} g(\mathcal{X})d\mathcal{X}.$$

By the definition of an indefinite integral,

$$\int_{\lambda}^{\infty} g(\mathcal{X})d\mathcal{X} = \lim_{n \rightarrow \infty} \int_{\lambda}^n g(\mathcal{X})d\mathcal{X}.$$

We can split this up into  $\int_{\gamma}^{\lambda} g(\mathcal{X} + \gamma\mathcal{I})d\mathcal{I} + \int_{\lambda}^{\infty} g(\mathcal{X})d\mathcal{X}$  because we now have a definite integral. Again using the definition of an indefinite integral we have

$$\int_{\gamma}^{\lambda} g(\mathcal{X} + \gamma\mathcal{I})d\mathcal{I} + \int_{\lambda}^{\infty} g(\mathcal{X})d\mathcal{X} = \lim_{n \rightarrow \infty} \int_{\lambda}^n g(\mathcal{X})d\mathcal{X}.$$

Now, since we are trying to show an if and only if relationship, we must prove the implication both ways. First, let  $\int_{\gamma}^{\infty} g(\mathcal{X} + \gamma\mathcal{I})d\lambda$  be convergant. We must show that  $\int_{\gamma}^{\infty} g(\mathcal{X})d\lambda$  is convergant. We have that  $\int_{\gamma}^{\infty} g(\mathcal{X})d\mathcal{X} =$

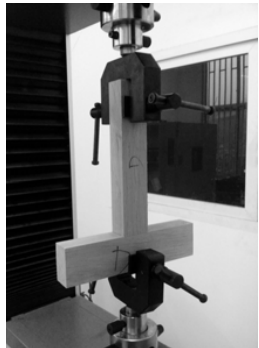
# Double submission / Slicing (Caught by ScigenDetection.imag.fr)

## Manufacturing and Engineering Technology (ICMET 2014) 2015 Taylor and Francis Group

- Influence of Fit Tolerance Between Square Tenon Thickness and Oval Mortise Width on Ultimate Pulling Resistance
- Influence of Fit Tolerance Between Tenon Width and Mortise Length on Joint Performance of Oval-tenon Manufactured by *Alnus formosana* Wood

0.1	0.2	0.3	0.4
12.1	12.2	12.3	12.4

ns were prepared for each A value, 48 sets of specimens were prepared to specimen size is shown in figure 2. were assembled into T-shaped joints  $\mu$ l acetate emulsion. The glue-spread  $g/m^2$ . The pull-out experiment was per- microcomputer electronic universal me- ing machine after the specimens placed own in figure 3).



Schematic diagram of specimen clamping.



Figure 5. Specimen failure mode when  $\alpha=0.2$ mm

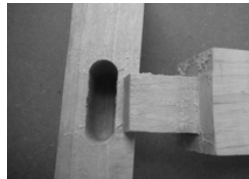


Figure 6. Specimen failure mode when  $\alpha=0.1$ mm

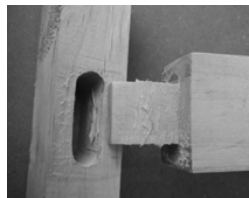
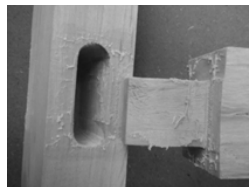


Figure 7. Specimen failure mode when  $\alpha=0$ mm



0.4	0.6	0.8	1
35.4	35.6	35.8	36

ns were prepared for each B value, 48 sets of specimens were prepared to specimen size is shown in figure 2. were assembled into T-shaped joints  $\mu$ l acetate emulsion. The glue-spread  $g/m^2$ . The pull-out experiment was per- microcomputer electronic universal me- ing machine after the specimens placed own in figure 3).



Schematic diagram of specimen clamping.

ut experiment was carried out by mi-



Figure 5. Specimen failure mode when  $\beta=0$ mm

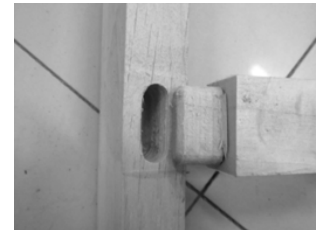


Figure 6. Specimen failure mode when  $\beta=0.1$ mm



Figure 7. Specimen failure mode when  $\beta=0.2$ mm





Beware! Highjacking! Jeffrey Beall <http://scholarlyoa.com>

# Hermès

Une revue de l'Institut des sciences de la communication du CNRS (ISCC)

I-Revues > HERMÈS >

Rechercher dans cette communauté et ses collections :

Aller

>>

Par date de publication

Auteurs

Titres

Sujets

Recherche

HERMÈS

Aller

- tout I-Revues  
 Cette communauté

Recherche avancée

Numéros parus

Directeur de publication

La communication est une valeur, une aspiration, mais elle est aussi une industrie, un marché florissant, voire une idéologie. Autrement dit, un phénomène complexe et polysémique qui requiert un travail d'analyse critique et de compréhension. Tel est le pari scientifique de la revue Hermès depuis sa création en 1988 : étudier de manière interdisciplinaire la communication dans ses rapports avec les individus, les techniques, les cultures, les sociétés.

Hermès, tout en étant une revue scientifique, souhaite rester accessible à un public ouvert, intéressé par l'émergence des problèmes théoriques liés à la communication. À condition  
 Hermes Journal ; ISSN: 0767-9513; France

SHARE



## HERMES JOURNAL FRANCE

LANGUAGE

English

JOURNAL CONTENT

Search

All

Search

Search

Browse

HOME ABOUT LOGIN REGISTER SEARCH CURRENT ARCHIVES ANNOUNCEMENTS

Home > Hermes Journal France

## Hermes Journal France

ISSN: 0767-9518

OPEN JOURNAL SYSTEMS

Journal Help

USER

Username

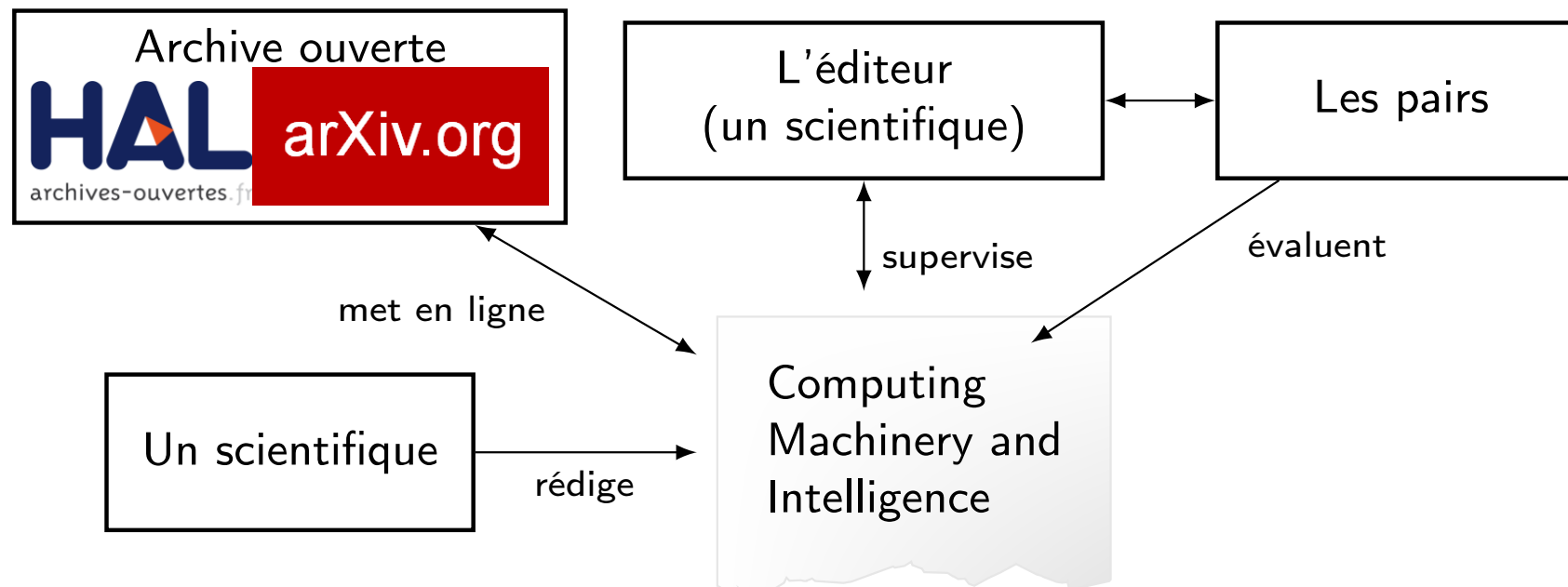
Password

Remember me

Login



# Publier : les épi-journaux [épisciences.org](http://episciences.org)





# Conclusion and Future/Ongoing works

## Publication procedures, models and habits

- Why fake papers were accepted, published and ... sold.
- Traditional publisher vs open access vs epi-journals.
- Knowledge diffusion: better and less... or as much as possible.

## Blind management rules...

- ... are an incitation to malpractices: slicing, plagiarism, faked data, ...

## Automatic detection of new generators

- Hand written PCFG : find dense cluster inside a population.
- Study other kind of generator (language model).

## In the web today

- Automatic knowledge extraction/detection/generation.
- How to separate the wheat from the chaff... and scale up !

# Thanks



Beel, J. and Gipp, B. (2010).  
Academic search engine spam and google scholar's resilience against it.  
*Journal of Electronic Publishing*, 13(3).



Beel, J., Gipp, B., and Wilde, E. (2010).  
Academic search engine optimization (aseo).  
*Journal of scholarly publishing*, 41(2):176–190.



Dalkilic, M. M., Clark, W. T., Costello, J. C., and Radivojac, P. (2006).  
Using compression to identify classes of inauthentic texts.  
*In Proceedings of the 2006 SIAM Conference on Data Mining*.



Fahrenberg, U., Biondi, F., Corre, K., Jégourel, C., Kongshøj, S., and Legay, A. (2014).  
Measuring structural distances between texts.



CoRR, abs/1403.4024.

Ginsparg, P. (2014).

Automated screening: Arxiv screens spot fake papers.  
*Nature*, 508(7494):44–44.



Hirsch, J. E. (2005).

An index to quantify an individual's scientific research output.  
*Proceedings of the National Academy of Science*, 102:16569–16572.



Labbé, C. (2010).

Ike antkare, one of the great stars in the scientific firmament.  
*International Society for Scientometrics and Informetrics Newsletter*, 6(2):48–52.



Labbé, C. and Labbé, D. (2006).

A tool for literary studies. Intertextual distance and tree classification.  
*Literary and Linguistic Computing*, 21(3):311–326.



Labbé, C. and Labbé, D. (2013).

Duplicate and fake publications in the scientific literature: how many scigen papers in computer science?  
*Scientometrics*, 94(1):379–396.



Lavoie, A. and Krishnamoorthy, M. (2010).

Algorithmic Detection of Computer Generated Text.  
*ArXiv e-prints*.



Lopez-Cozar, E. D.,

Robinson-García, N., and Torres-Salinas, D. (2012).

Manipulating google scholar citations and google scholar metrics: Simple, easy and tempting.  
*arXiv preprint arXiv:1212.0638*.



Xiong, J. and Huang, T. (2009).

An effective method to identify machine automatically generated paper.  
*In KESE '09. Pacific-Asia Conference*, pages 101–102.