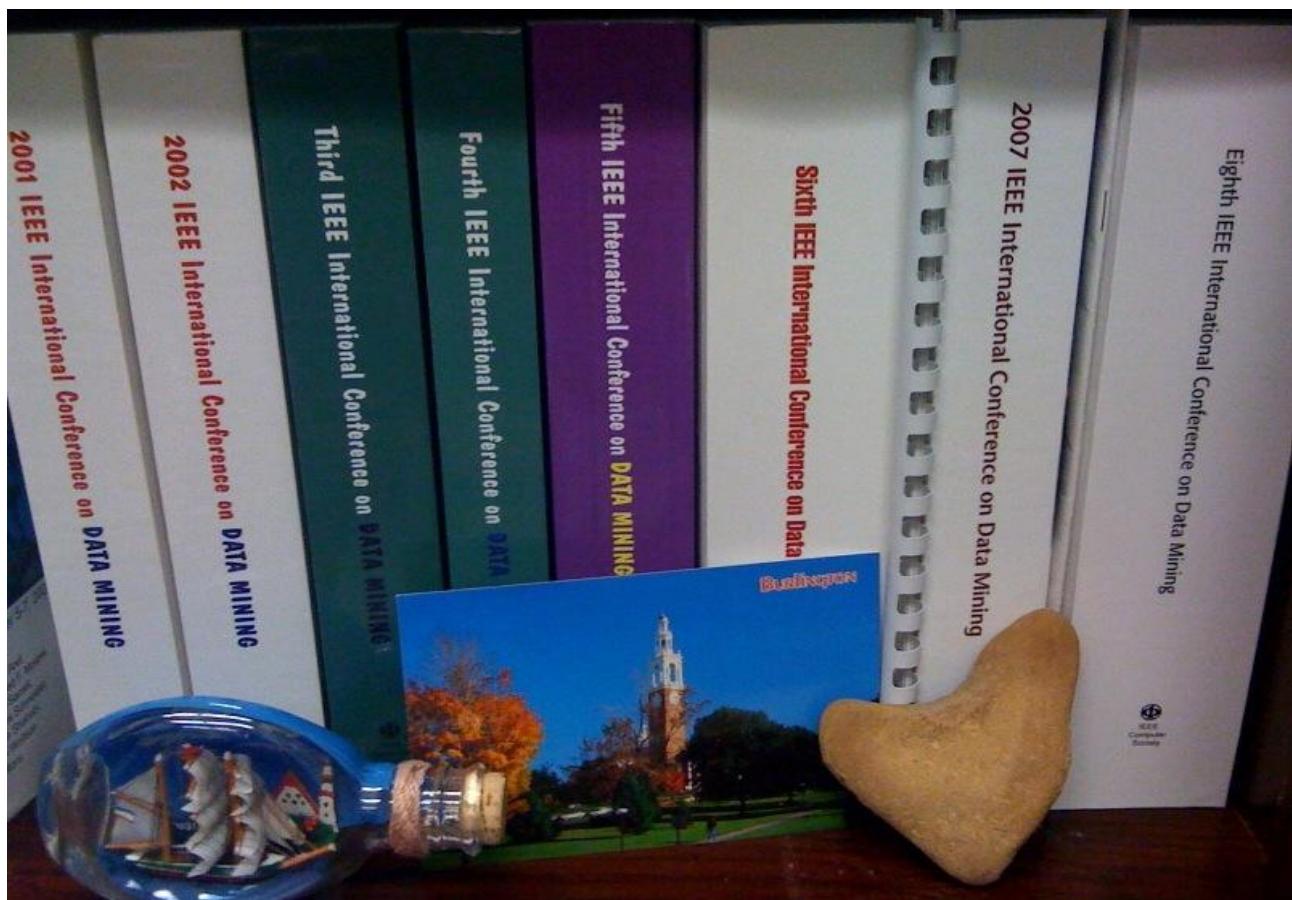


Welcome to 2015 ICDM Community Meeting





2015 ICDM Community Meeting

- 1. ICDM '15 Student Travel Awards**
- 2. ICDM as “the World’s Premier Research Conference in Data Mining”**
By Xindong Wu
- 3. Call for ICDM '17 and '18 proposals**
- 4. ICDM '15 Highlights**
By ICDM '15 PC Chairs
- 5. ICDM '16: Attractions and Call for Papers**
By ICDM '16 General Chairs & PC Chairs
- 6. Data Mining on ICDM '15 Paper Submissions**
By Shusaku Tsumoto/Shoji Hirano

ICDM-2015 Student Travel Awards

Sponsors

- Prof. Ram Kumar Memorial Foundation
- National Science Foundation (NSF)
- Alberta Innovates Center for Machine Learning (AICML)
- Rutgers Business School
- Total award amount: \$47,450



Applicant Statistics

- 119 applications
- Selection criteria:
 - Student (excluding post doctorates)
 - First-author full paper, short paper, workshop paper or PhD forum paper
- NSF supported: 44
- Ramkumar Foundation and AICML: 45



ICDM Student Travel Awardees

Anuj Karpatne

Aubrey Gress

Biao Chang

Bo Li

Bokai Cao

Boshu Ru

Chang Xu

Charalampos

Mavroforakis

Chen Chen

Cheng Long

Chenguang Wang

Chien-Wei Chang

Constantine Alexander

Vitt Dawei Zhou

Dezhong Yao

Dongjin Song

Emre Eftelioglu

Fabio Pulvirenti

Fangzhao Wu

Fenglong Ma

Forough Arabshahi

Handong Zhao

Hanmo Wang

Harsh Dani

Hemanth

Venkateswara

Hongfu Liu

Houping Xiao

Huang Xu

Huayu Li

Hugo Gualdron

Ioakeim Perros

Iuliia Chepurna

Ivan Brugere

Jaekoo Lee

Jan N. van Rijn

Jiawei Zhang

Jin Huang

Jingjing Wang

Jingyuan Yang

Juhua Hu

Junming Liu

Justin Sampson

Mihir Shekhar

ICDM Student Travel Awardees

Keigo Kimura

Lennart Downar

Li Zhang

Liang Zhao

Linyun Yu

Liyuan Liu

Manju Bhardwaj

Miguel Bernard

Minh Nguyen

Mostafa Bayomi

Mustafa Coskun

Olivia Simpson

Partha Mukherjee

Patrick Donnelly

Qiaoying Huang

Raksha Kumaraswamy

Rong Zhu

Rui Liu

Rui Meng

Ryan Scrivens

Samer Al-Khateeb

Suhas Ranganath

Tao Zeng

Tao-Yang Fu

Yilin Wang

Yong Zheng

Yu Liu

Yu Wang

Tianran Hu

Tingting Liang

Tri Doan

Wei Xiong

Wenchao Yu

Wenjun Wang

Xiang Li

Xin Li

Xinhu Zheng

Yang Yang

Yanjie Fu

Yao Zhang

Yifan Chen

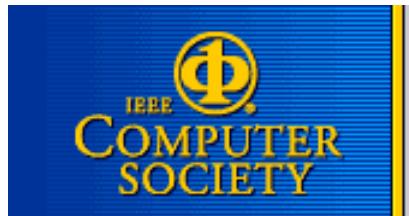
Yijun Zhao

Zhao Kang

Zhe Jiang

Zhenyu Chen

Zitao Liu



2015 ICDM Community Meeting

1. **ICDM '15 Student Travel Awards**
2. **ICDM as “the World’s Premier Research Conference in Data Mining”**
By Xindong Wu
3. **Call for ICDM '17 and '18 proposals**
4. **ICDM '15 Highlights**
By ICDM '15 PC Chairs
5. **ICDM '16: Attractions and Call for Papers**
By ICDM '15 General Chairs & PC Chairs
6. **Data Mining on ICDM '15 Paper Submissions**
By Shusaku Tsumoto/Shoji Hirano



Important ICDM Traditions

1. Truly international - “*The World’s Premier Research Conference in Data Mining*”
2. Exciting social programs as well as outstanding technical programs
3. **World-renowned invited speakers!**
4. Service innovation - Designed and organized by a collection of people.



Recent Changes in ICDM (1)

- Starting at ICDM '06,
 - Tutorials and workshops: complimentary participation
 - Workshops (*with formal proceedings!*) generally on the first day
 - Tutorials in parallel with paper sessions
 - Awards:
 - Data Mining Research Contribution and Service Awards: Highest data mining awards (each with a check of \$2500)
 - An increased number of Student Travel Awards
- Starting at ICDM '07,
 - Double blind reviewing
- Starting at ICDM '08,
 - Best student paper award
 - Demo papers (in workshop proceedings)
- Starting at ICDM '09,
 - No hardcopy, but CD copies/memory sticks only



Recent Changes in ICDM (2)

- Starting at ICDM '10,
 - 10-year highest impact paper award (with a check of \$1K)
- Starting at ICDM '11,
 - A PhD Forum
 - ***Triple blind reviewing***
- Starting at ICDM '14,
 - Reproducibility.



Call for ICDM '17 & '18 Proposals

- Generally in December
- ICDM '17: America
- ICDM '18: Asia
- Proposals to xwu@uvm.edu



ICDM 2015

IEEE International Conference on Data Mining

IEEE ICDM 2015: PC Chairs' Report

- A Data Mining Study about ICDM Submissions

Zhi-Hua Zhou, Nanjing University

Charu Aggarwal, IBM Research





Paper Statistics

- Submitted: 807
 - 49 different countries
 - 440 (55%) student first authors
- Accepted: 146 (18.1%)
 - 68 regular papers (20 minute presentation)
 - 78 short papers (10 minute presentation)
- 24 Technical Sessions



Review Process

- 35 Area Chairs and 307 PC members
- Triple blind reviewing process
- **Phase 1:** 3 reviewers assigned to each paper (based on bids)
- **Phase 2:** Each paper assigned to an area chair who guided the review discussion and made the recommendation.
- **Phase 3:** Program Chairs made the final decision based on all input (seeking additional input when necessary)

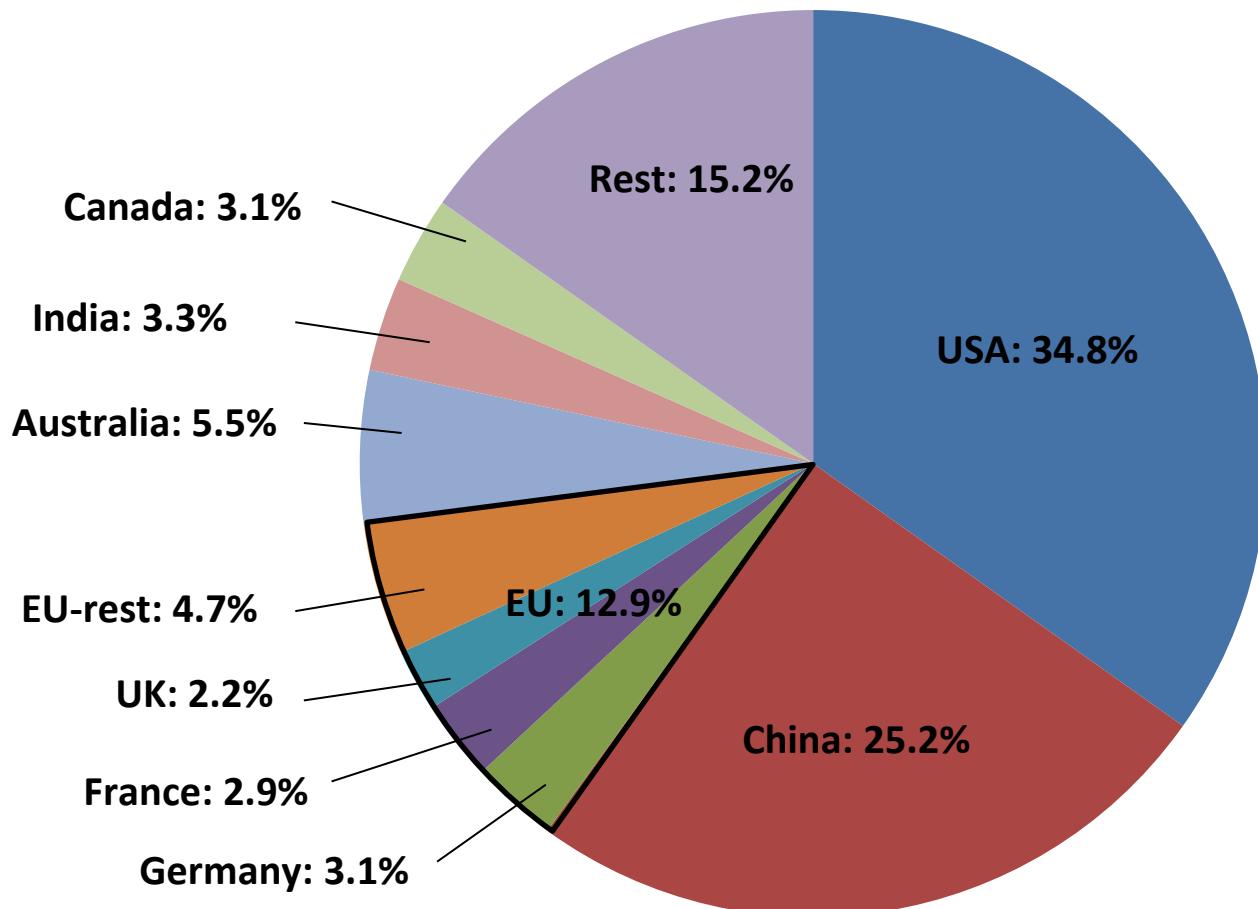


Review Process

- ICDM 2015 is one month earlier, thus we have a much tighter review schedule
- We have tried our best to save time at every section
- PC chairs did paper assignments in 48 hours
- PC members finished reviews in 4 weeks
- ACs made recommendation in 5 days
- Finally, author notifications were sent out 3 weeks earlier (Aug.07 vs. Aug.28): people can have more time for visa application and traveling preparation

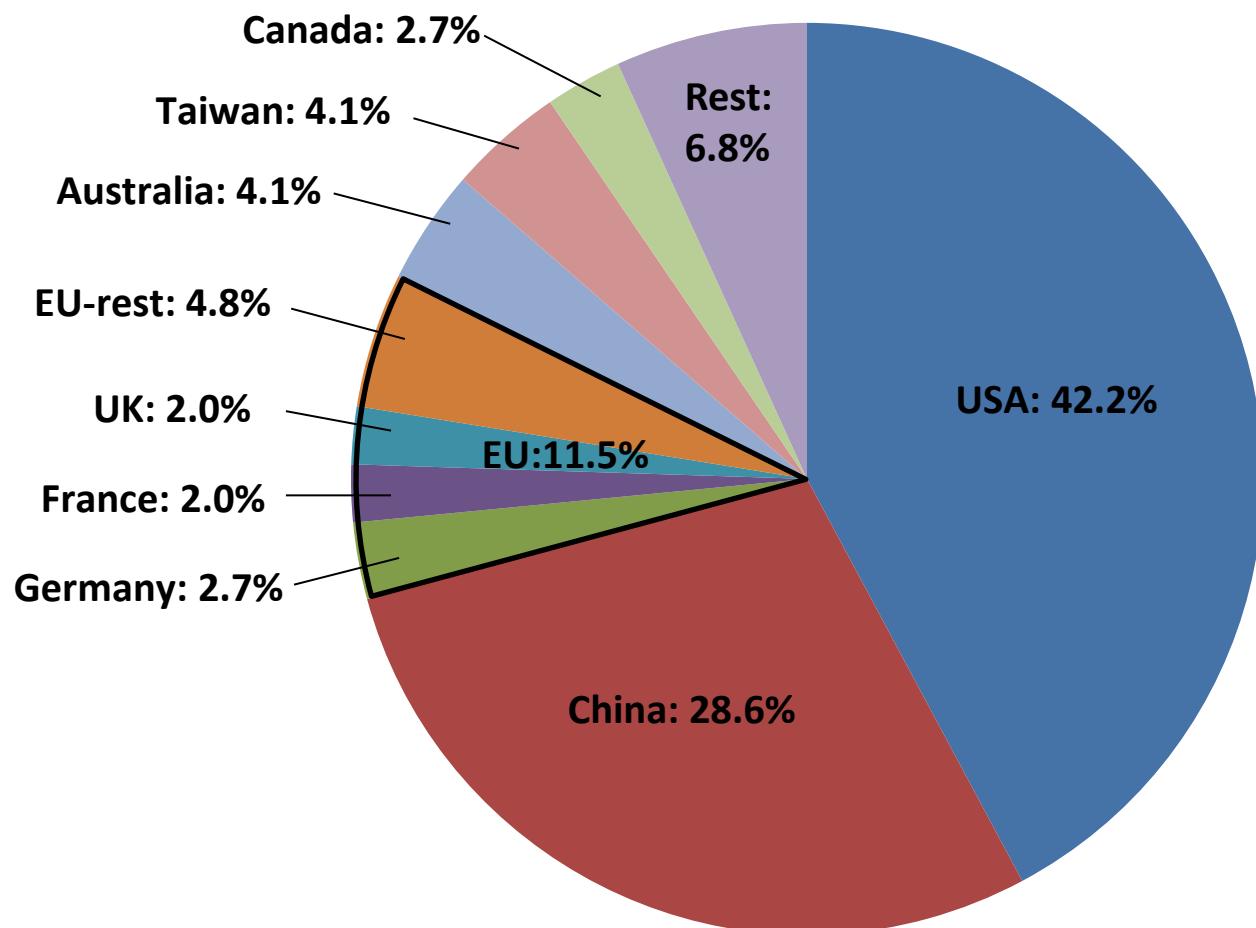


Submitted/All





Accepted/All Accepted





Replicability is Important

*13: Would you be able to replicate the results based on the information given in the paper?

- Yes; No

	# Submissions	# Accepted	Acceptance Rate
# Reviews say 'Yes' >=1	362	88	24.3%
# Reviews say 'Yes' >=2	128	37	28.9%

*14: Are the data and implementations publicly available for possible replication?

- Yes; No

	# Submissions	# Accepted	Acceptance Rate
# Reviews say 'Yes' >=1	475	107	22.5%
# Reviews say 'Yes' >=2	190	59	31.1%

Significantly higher than the overall acceptance rate 18.1%



Academic & Industry Papers

	# Submissions	Percentage	# Accepted	Acceptance Rate
Pure academic	484	60.0%	89	18.4%
Pure industry	101	12.5%	16	15.8%
Academic + Industry	222	27.5%	41	18.5%

- Collaboration between academic and industry is helpful
- More helpful for industrial authors



Hot Topics

Top 11 submitted

Top 11 accepted

Rank	Topic	Rank	Topic
1	1.1	1	1.1
2	2.5	2	2.5
3	1.5	3	1.5

4	2.9	X	4	1.2
5	1.2	X	5	3.1
6	2.7	X	6	1.6
7	3.1	X	7	2.9
8	1.6	X	8	2.7

9	1.4	X	9	2.1
10	4.4	X	10	1.4
11	2.1	X	11	4.4

39 topics

Top 11: 70.4% submissions

Top 11: 81.5% accepted papers

Novel data mining algorithms in traditional areas

Mining and link analysis in networked settings

Mining spatial and temporal datasets

Models and algorithms for new, structured, data types

Data pre-processing

Mining textual and unstructured datasets

Mining in emerging domains – Other

Data mining in electronic commerce

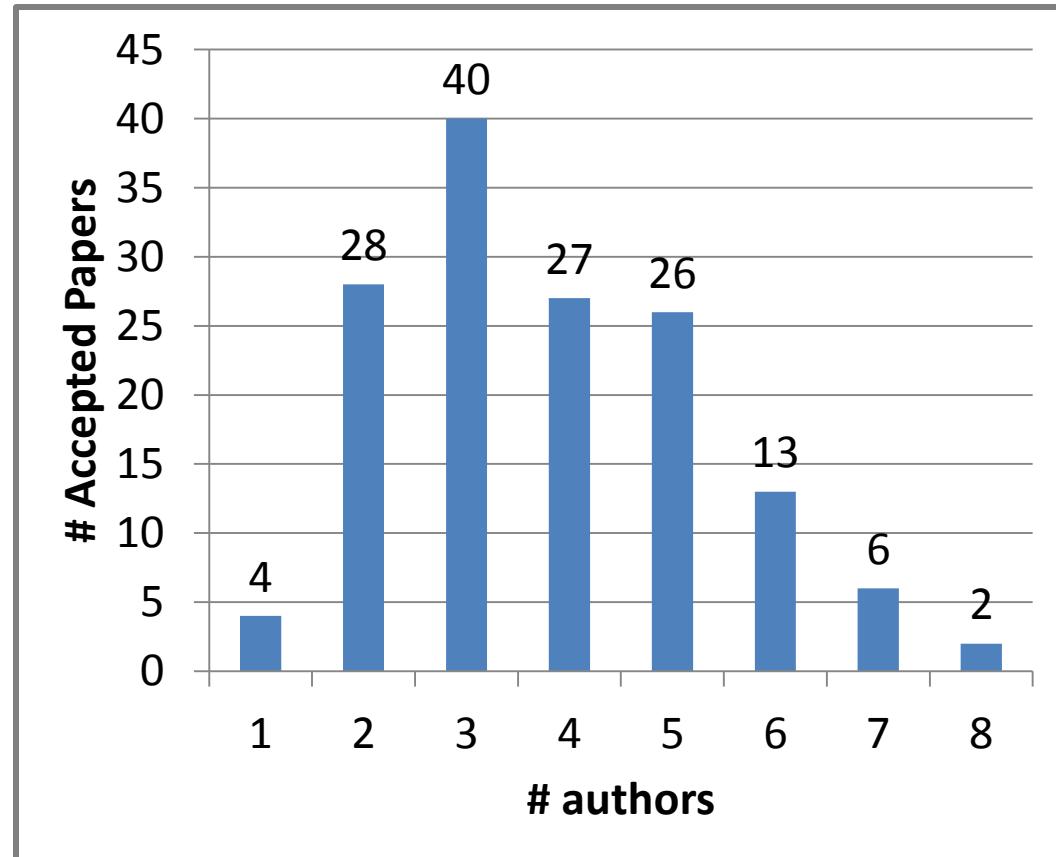
Stream Data Mining

Mining sequences and sequential data

Healthcare, epidemic modeling, and clinical research



Authors per Paper

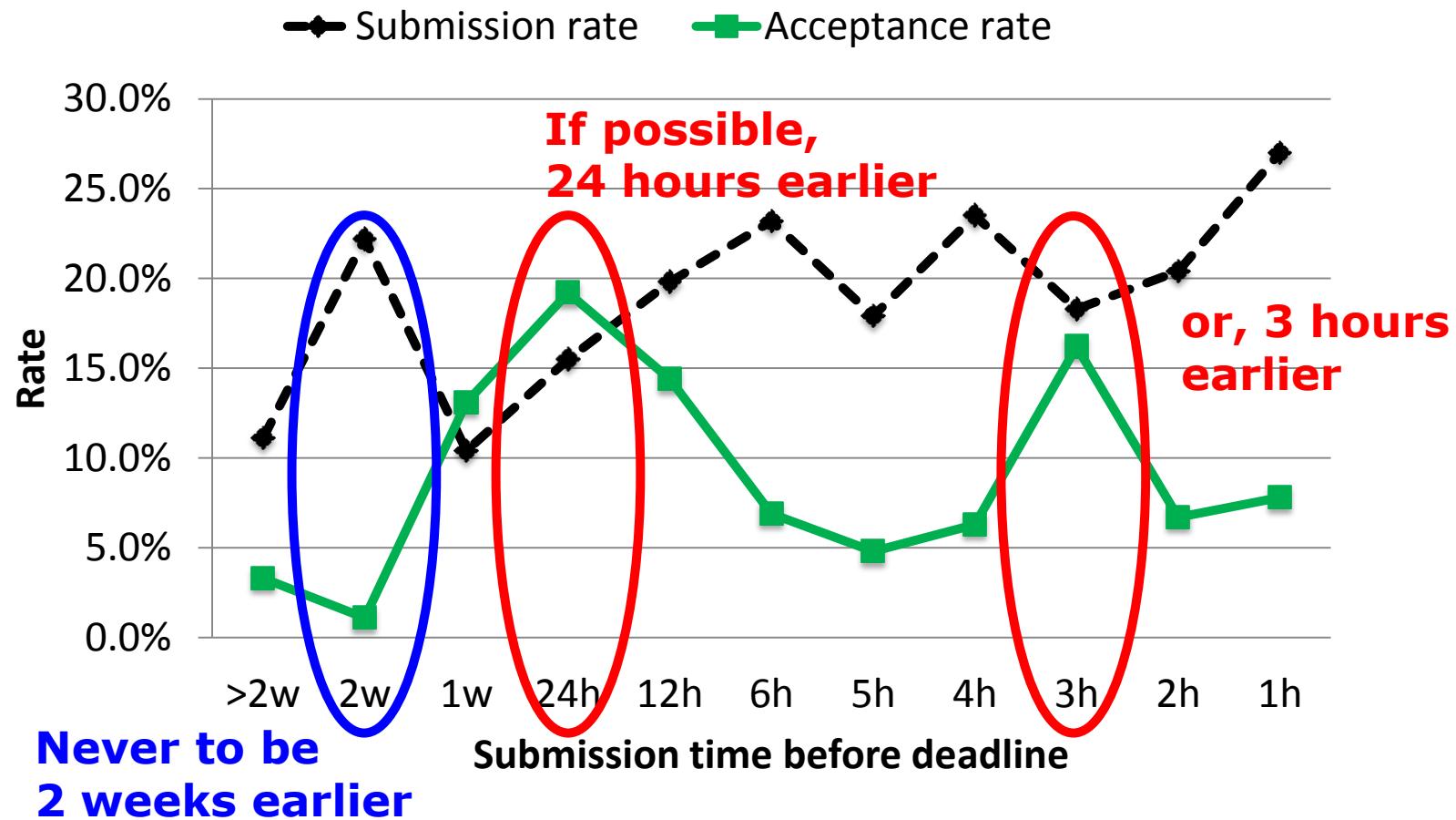


On average

- Submitted papers:
3.3 / paper
- Accepted papers
3.8 / paper



Submission Time Analysis





Submission Title Words Analysis

- Classes: *positive* - acceptance, *negative* - reject
- Words: $\{w_1, \dots, w_n\}$
- Probability estimation

$$P(w_i|C) = \frac{1 + \#w_i \text{ in } C}{2 + \#w_i}$$

- Regularization: to remove bias for rare words

$$\text{Reg}(w_i) = 1 - e^{-\sqrt{\frac{\#w_i}{\max_j \#w_j}}}$$

- Significance

$$P(w_i|C) \cdot \text{Reg}(w_i)$$



Top 10 **Positive** Title Words

Rank	Word	Probability	Regularization	Significance
1	systems	63.64%	0.0409	0.0260
2	model	32.43%	0.0791	0.0256
3	dynamics	75.00%	0.0335	0.0252
4	large-scale	47.06%	0.0525	0.0247
5	hierarchical	42.11%	0.0558	0.0235

Probably the best title:

Online dynamics of large-scale sparse model across hierarchical heterogeneous information systems

(probability of acceptance: 99.99%)



Top 10 **Negative** Title Words

Rank	Word	Probability	Regularization	Significance
1	social	85.00%	0.1006	0.0855
2	clustering	81.03%	0.0990	0.0802
3	detection	85.71%	0.0911	0.0780
4	approach	88.89%	0.0873	0.0776

Probably the worst title:

Graph clustering and classification approach for detection, prediction and recommendation in social networks mining

(probability of reject: 99.99%)

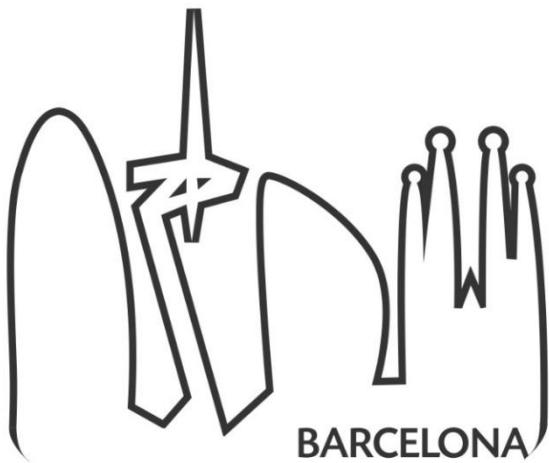


ICDM 2015

IEEE International Conference on Data Mining

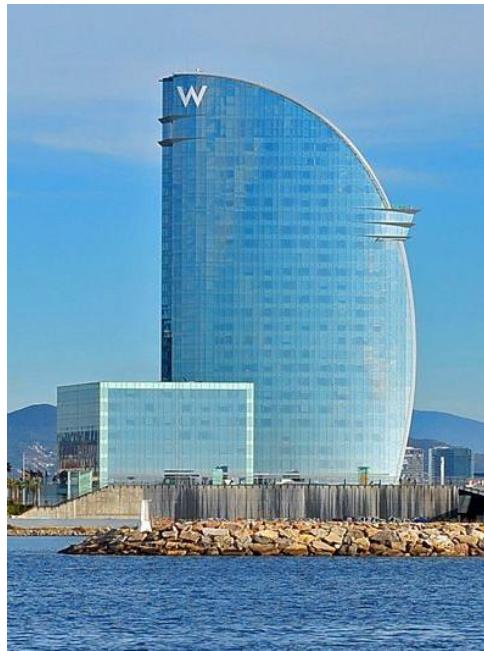


Thank you !



ICDM 2016

IEEE International Conference on Data Mining





ICDM 2016

IEEE International Conference on Data Mining



Ricardo Baeza-Yates,
General Chair



Michael Berthold,
PC Co-Chair



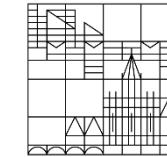
Francesco Bonchi,
PC Co-Chair

YAHOO!

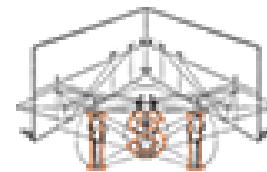


Universitat
Pompeu Fabra
Barcelona

Universität
Konstanz

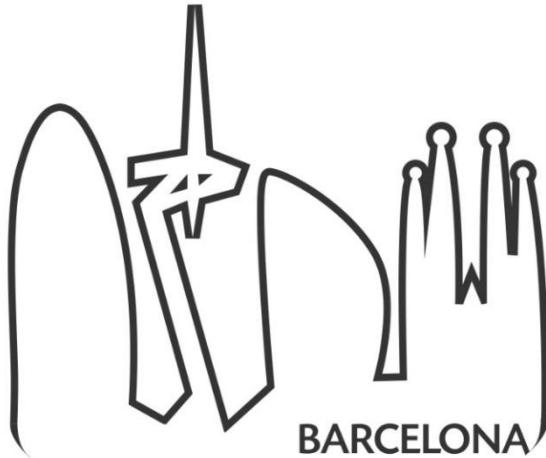


Open for Innovation
KNIME



ISI Foundation

eurecat
Centre Tecnològic de Catalunya



ICDM 2016

IEEE International Conference on Data Mining

December 12-15, 2016
World Trade Center,
Barcelona, Spain

(NIPS is 4-9 December in Barcelona)



ICDM 2016

IEEE International Conference on Data Mining

Paper submission deadline:
17th June 2016

BCN-ICDM 2016

12th -15th December 2016



BARCELONA 2016



- 1. compact city**
- 2. enjoys a very agreeable and mild climate all yearlong (DECEMBER AVG: 14° C, 57° F)**
- 3. offers a very intense cultural life.**
- 4. has an excellent network of public transport and flights.**
- 5. boasts a varied accommodation.**



BARCELONA 2016



- ✓ Perfectly linked to the main cities in all five continents
- ✓ BARCELONA-PRATS AIRPORT

137 scheduled destinations:

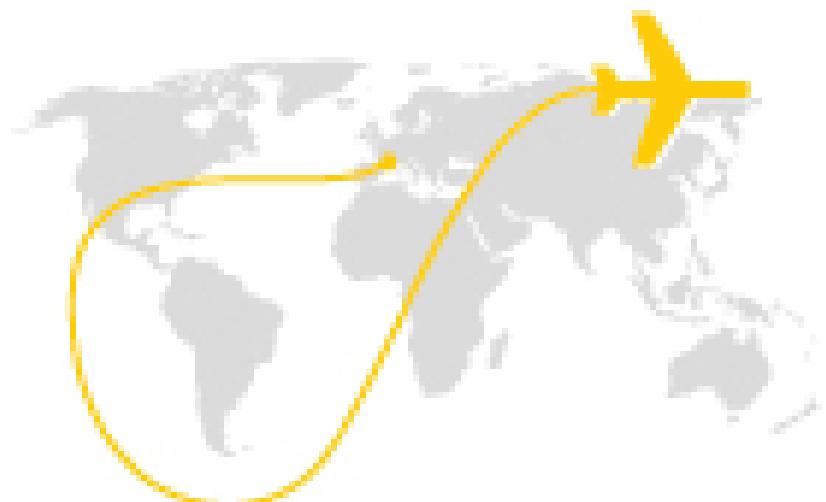
33 Domestic

72 European Union

32 International

- ✓ 19 Low-Cost Companies from/to B C N
- ✓ Prepared to attend 70 million passengers
- ✓ 1,300 daily flights coming from 130 destinations
- ✓ **GIRONA/REUS AIRPORTS (SECONDARY AIRPORTS)**
- ✓ National highway and train crossroad

an accessible city





Barcelona

WIDE PUBLIC TRANSPORTATION NETWORK

a comfortable city

- 6 metro lines
- 1,065 urban buses
- Tram
- Ferrocarril
- 15,637 taxi cabs
- 74 Touristic Buses

Touristic tickets for the days you need: **Hola BCN!**



BARCELONA
2016



Why

b a r c e | o n a

**BARCELONA
2016**



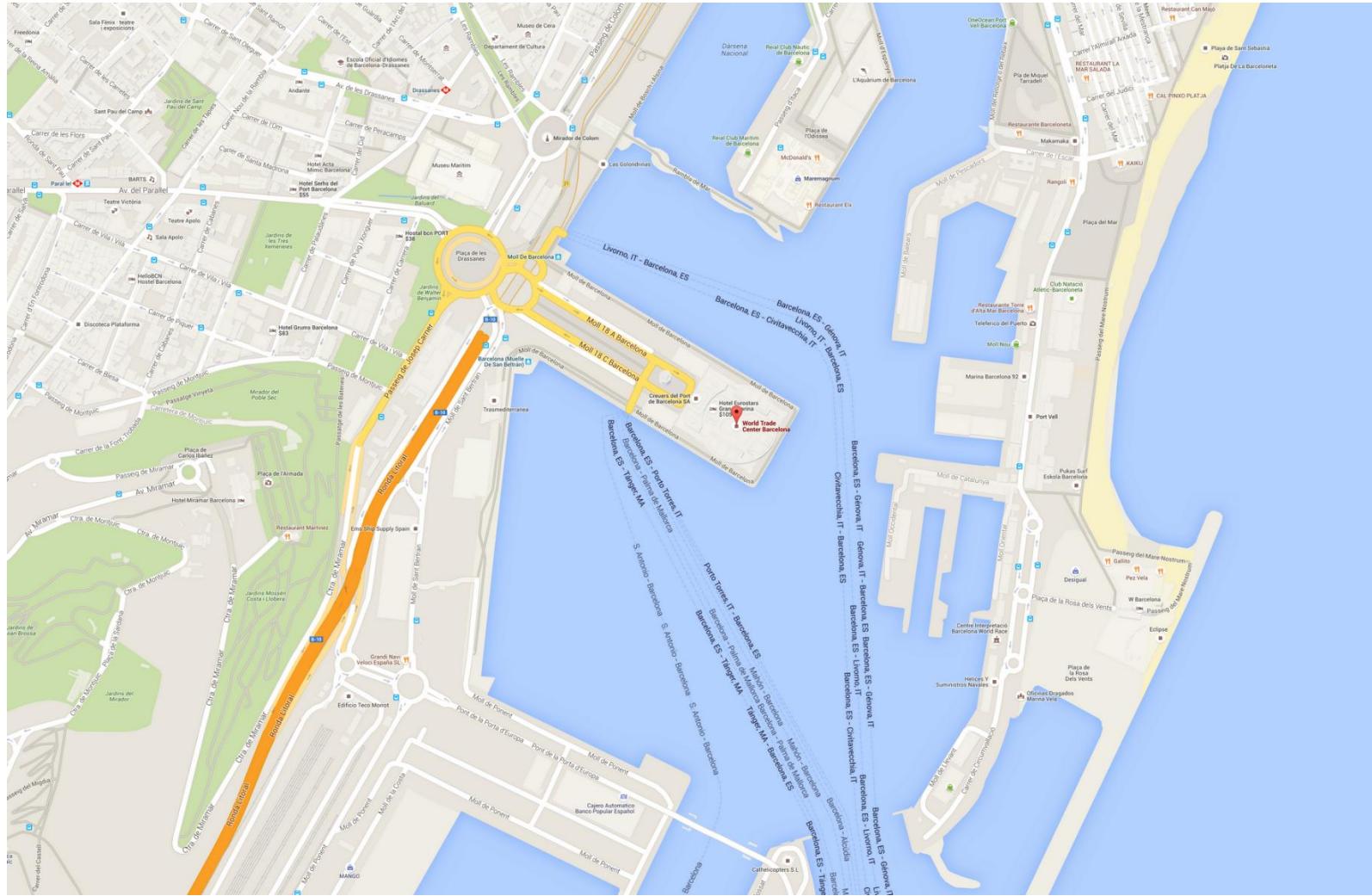
Video Barcelona



The Venue

World Trade Center Barcelona















AUDITORIUM (430 PAX)





AUDITORIUM (430 PAX)





BREAK-OUT 1(120 PAX), BREAK-OUT 2 (80 PAX), BREAK-OUT 3 (80 PAX)





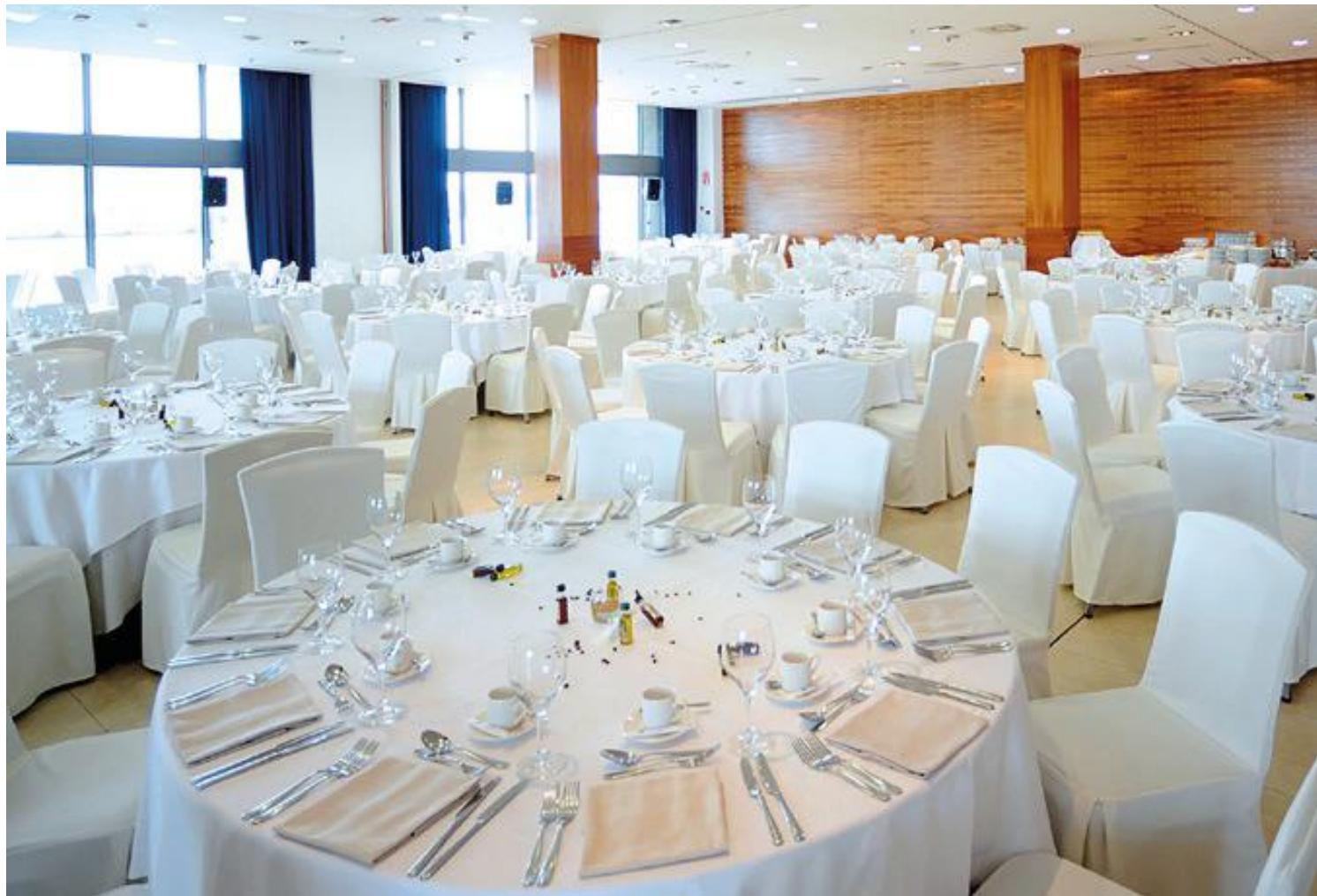
BREAK-OUT 1(120 PAX), BREAK-OUT 2 (80 PAX), BREAK-OUT 3 (80 PAX)





Room	Hauteur m ²	Théâtre	École	U	Banquet	Cocktail
Room 1	2,76m 24 m ²	30	—	6	—	—
Room 2	2,76m 40 m ²	30	—	30	—	+
Room 3	2,76m 57 m ²	30	20	16	—	—
Room 4	2,76m 52 m ²	30	20	16	—	—
Room 5	2,76m 52 m ²	30	20	10	—	—
Room 6	2,76m 57 m ²	30	20	16	—	—
Room 7	2,76m 49 m ²	20	—	10	—	—
Room 8	2,76m 24 m ²	10	—	6	—	—
Room 9	2,76m 75 m ²	50	35	22	—	—
Room 10	2,76m 60 m ²	40	25	18	—	—

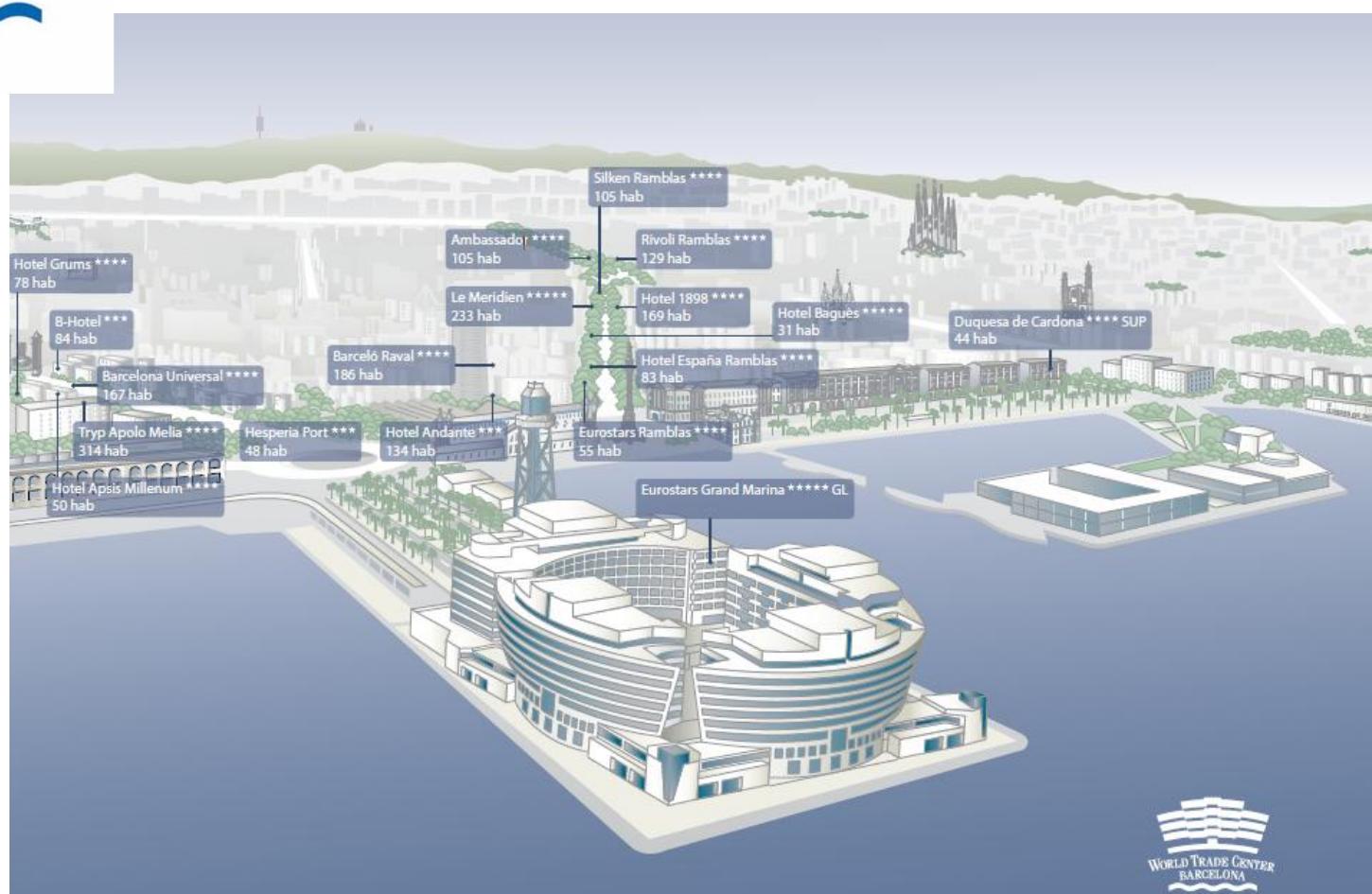
Room	Hauteur m ²	Théâtre	École	U	Banquet	Cocktail
Room 11	2,76m 60 m ²	40	25	18	—	—
Room 12	2,76m 76 m ²	50	35	22	—	—
Room 2+3	2,76m 90 m ²	50	25	20	—	—
Room 3+4	2,76m 111 m ²	60	40	32	—	—
Rooms 2+3+4	2,76m 151 m ²	90	45	36	—	—
Room 5+6	2,76m 111 m ²	60	40	32	—	—
Room 6+7	2,76m 98 m ²	50	25	29	—	—
Rooms 5+6+7	2,76m 151 m ²	90	45	36	—	—
Room 9+10	2,76m 136 m ²	90	60	38	—	—
Rooms 11+12	2,76m 196 m ²	90	60	38	—	—





Other RESTAURANTS & INTERNATIONAL CAFES





More than 1.000 bedrooms within walking distance

Excursion + Gala dinner



- Visit to Cavas Codorníu
 - Historic cava producer
 - Cava is the Catalan Champagne
-
- Designed by the architect Josep Puig i Cadafalch, the cellar was declared Historic and Artistic National Monument in 1976 and represents one of the most outstanding examples of modernist architecture in our country.

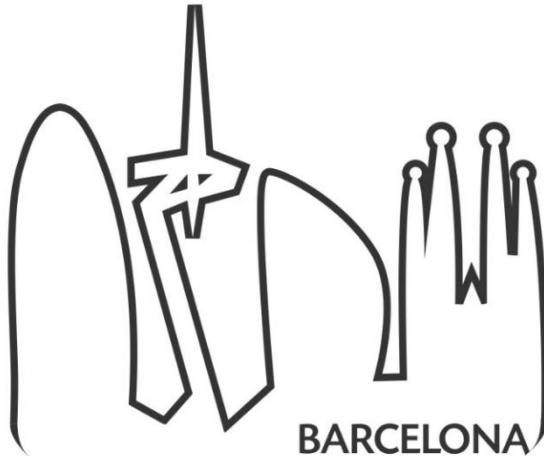


Excursion
+
Gala dinner



Excursion + Gala dinner





ICDM 2016

IEEE International Conference on Data Mining

December 12-15, 2016

World Trade Center, Barcelona

Web: <http://icdm2016.eurecat.org/>

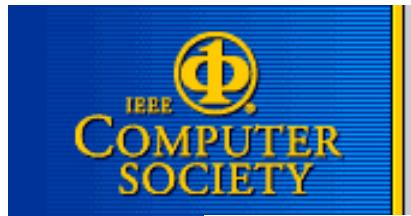
Mail: ICDM2016Chairs@eurecat.org

ICDM2016Sponsors@eurecat.org

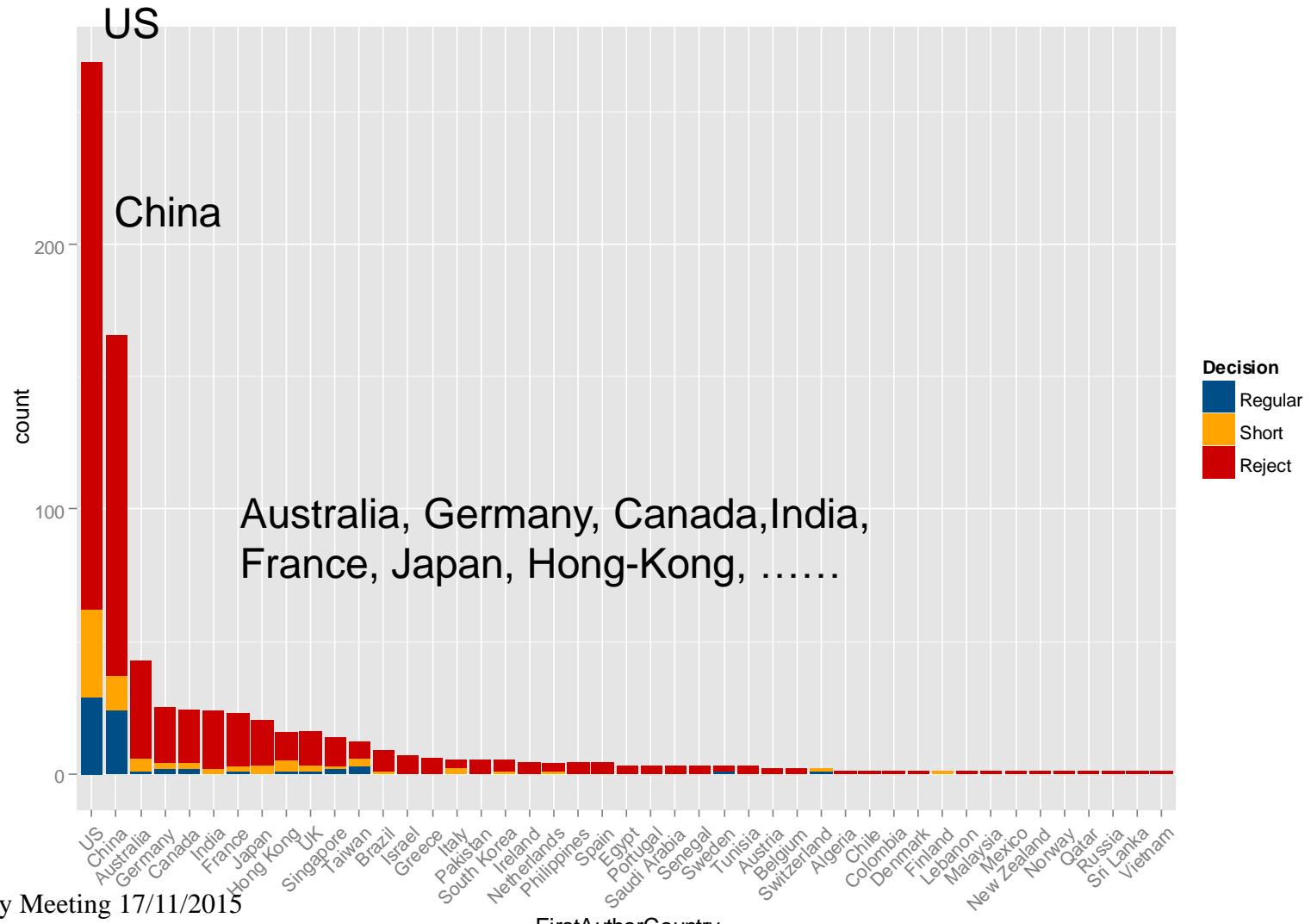


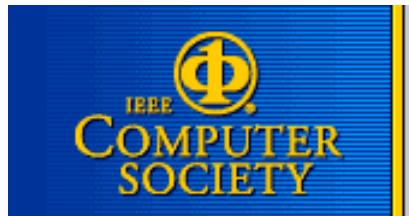
Data Mining on ICDM Submission Data (2015)

Shusaku Tsumoto, Shoji Hirano,
Ning Zhong and Xindong Wu



Country of First Author (2015)





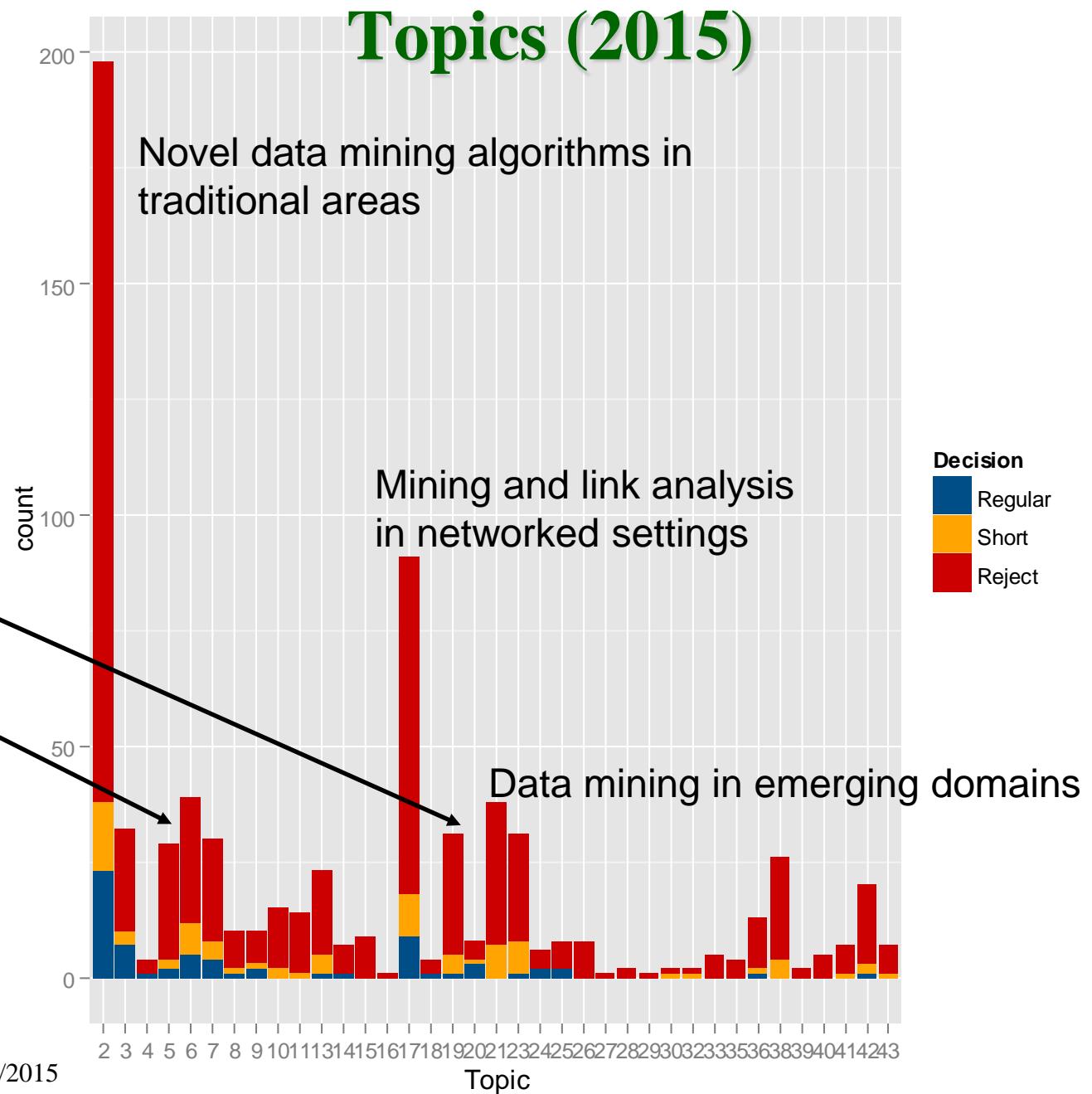
Topics (2015)

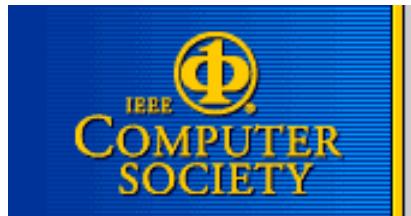
Data Mining in E-commerce

Mining spatial and temporal datasets

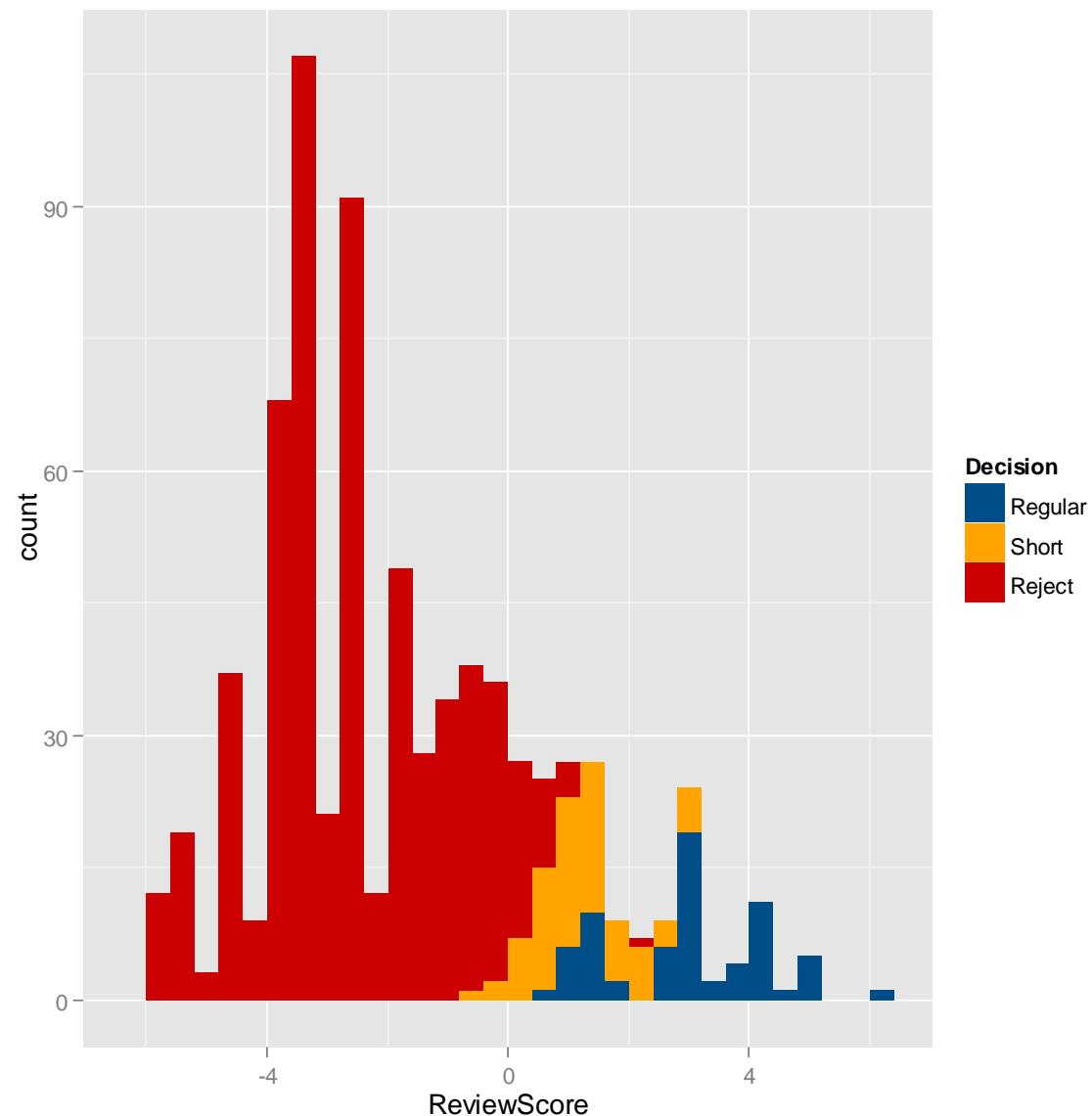
Mining and link analysis in networked settings

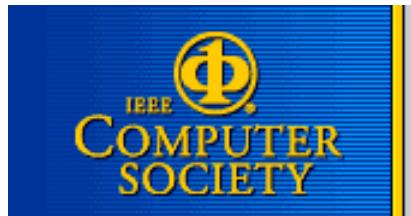
Data mining in emerging domains



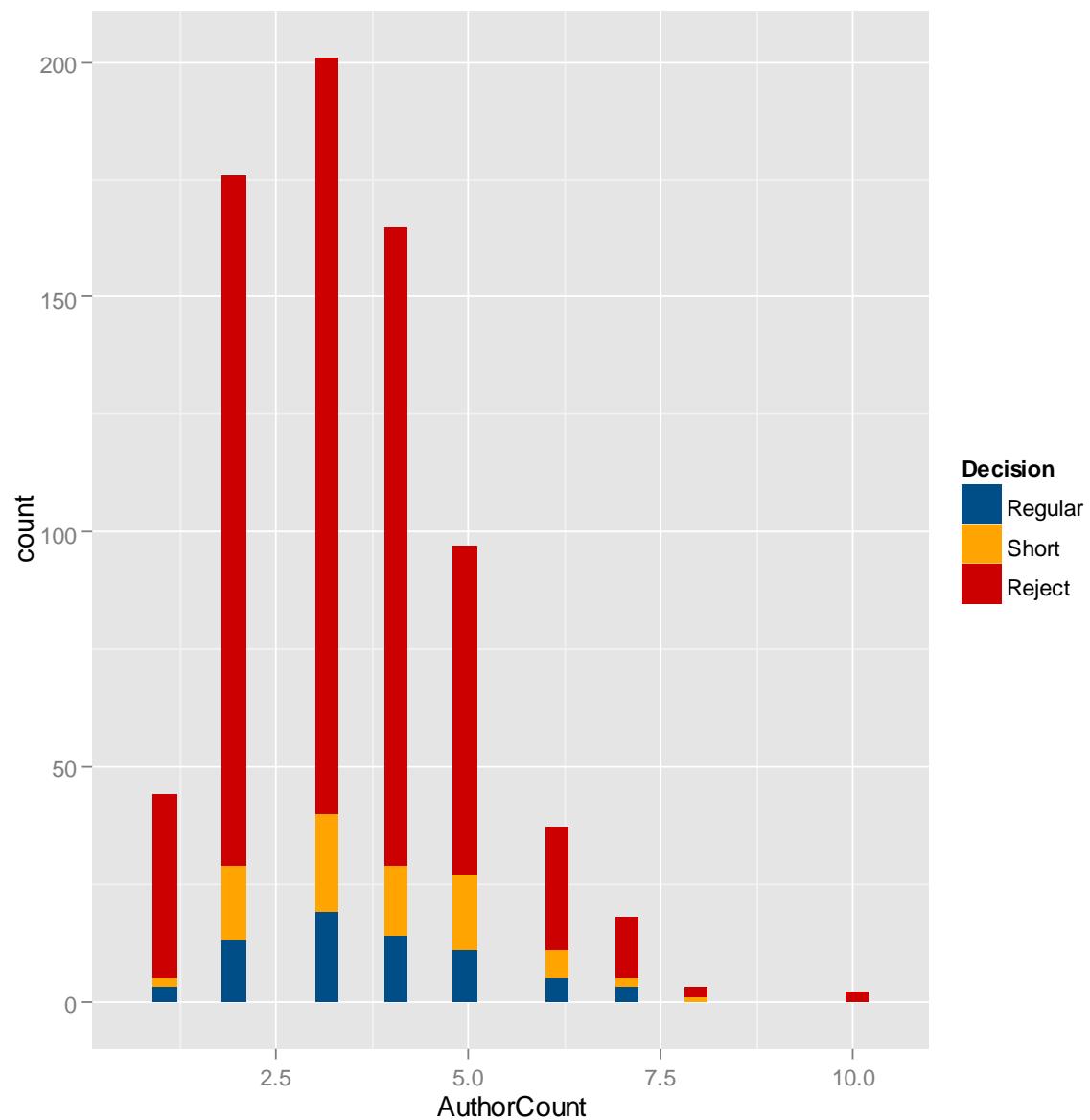


Review Score (2015)



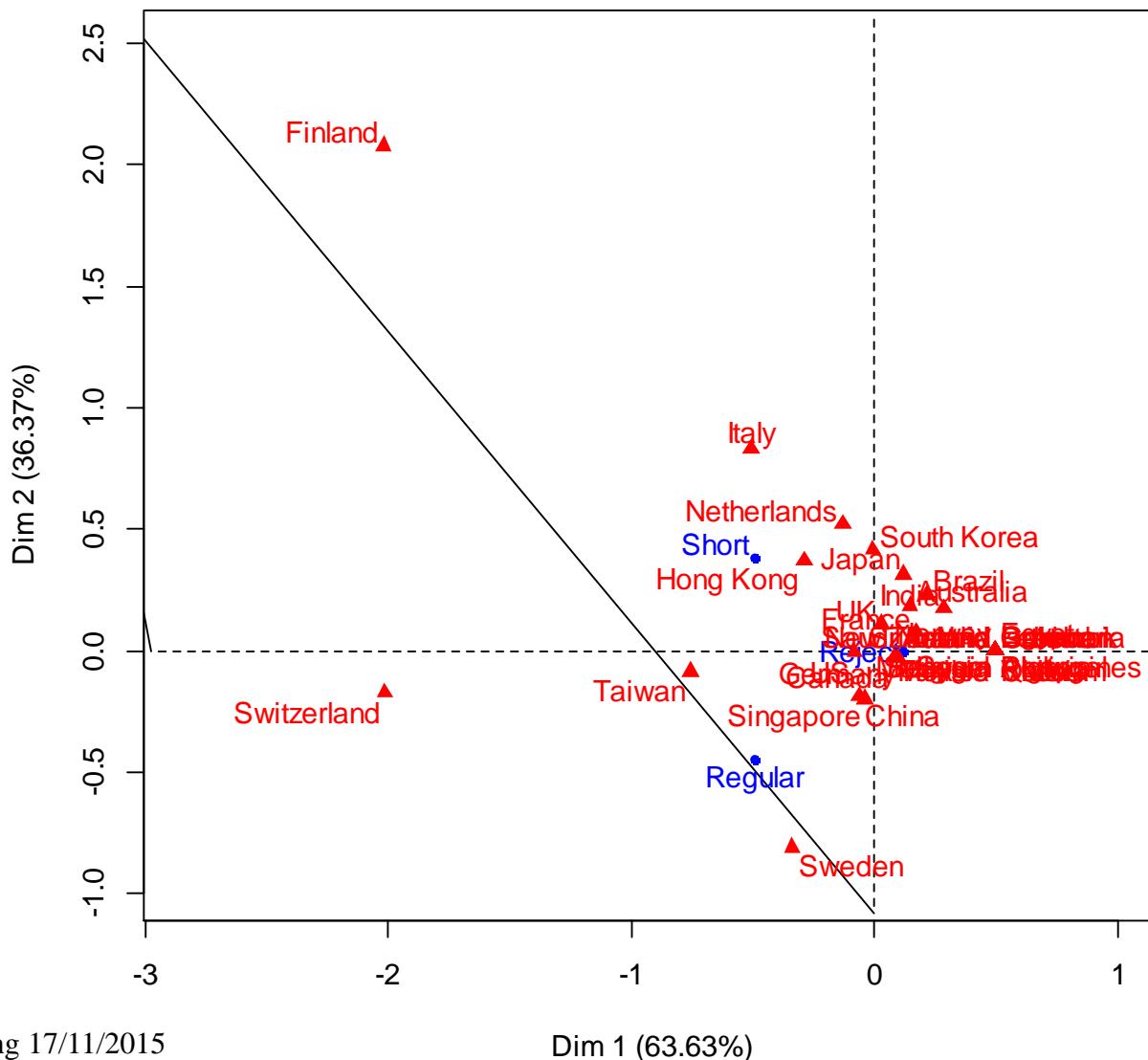
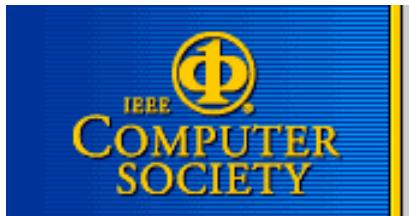


#Authors (2015)

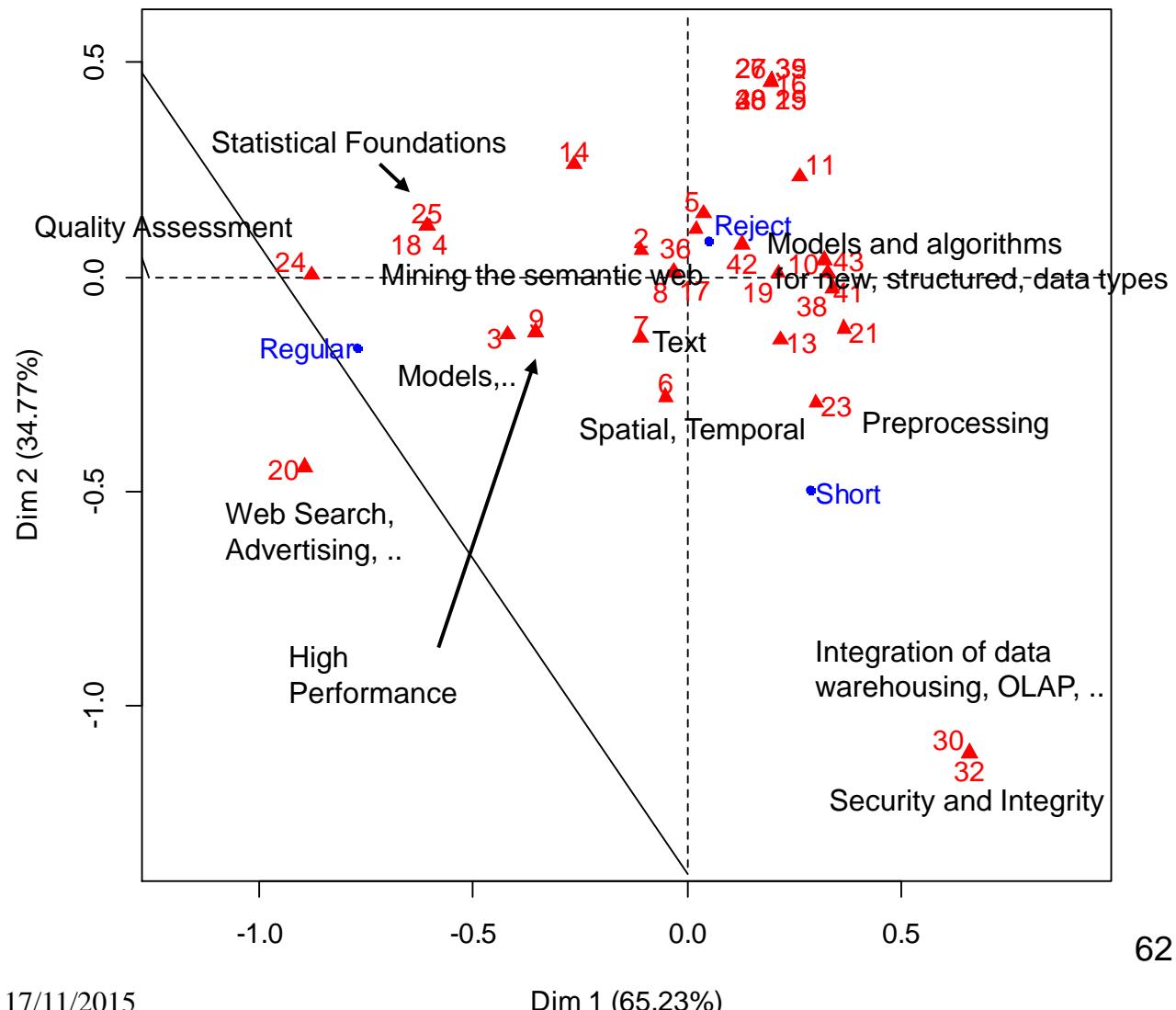


Correspondence Analysis (Country, 2015)

CA factor map



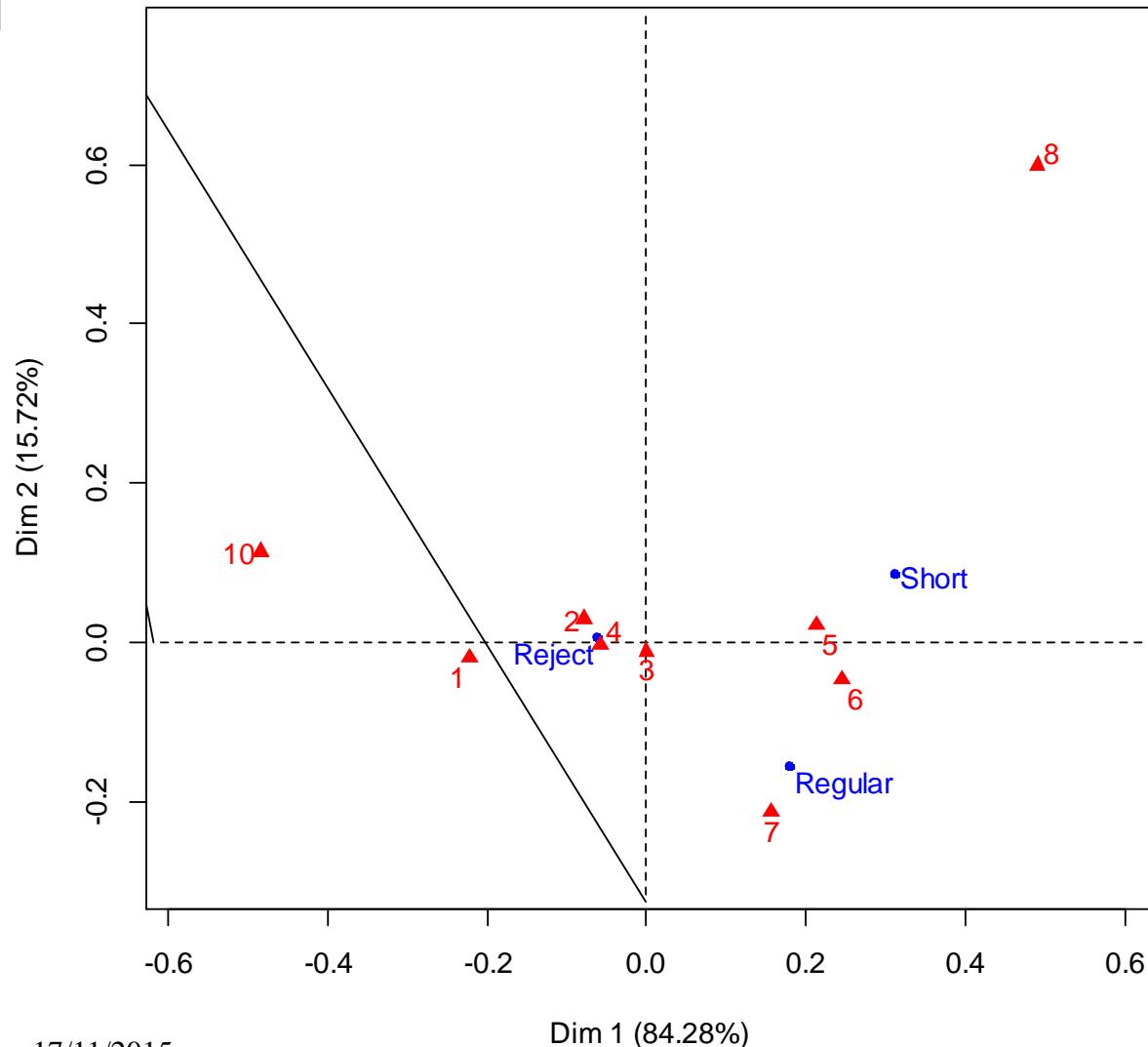
Correspondence Analysis (Topics, 2015)





Correspondence Analysis (#Authors, 2015)

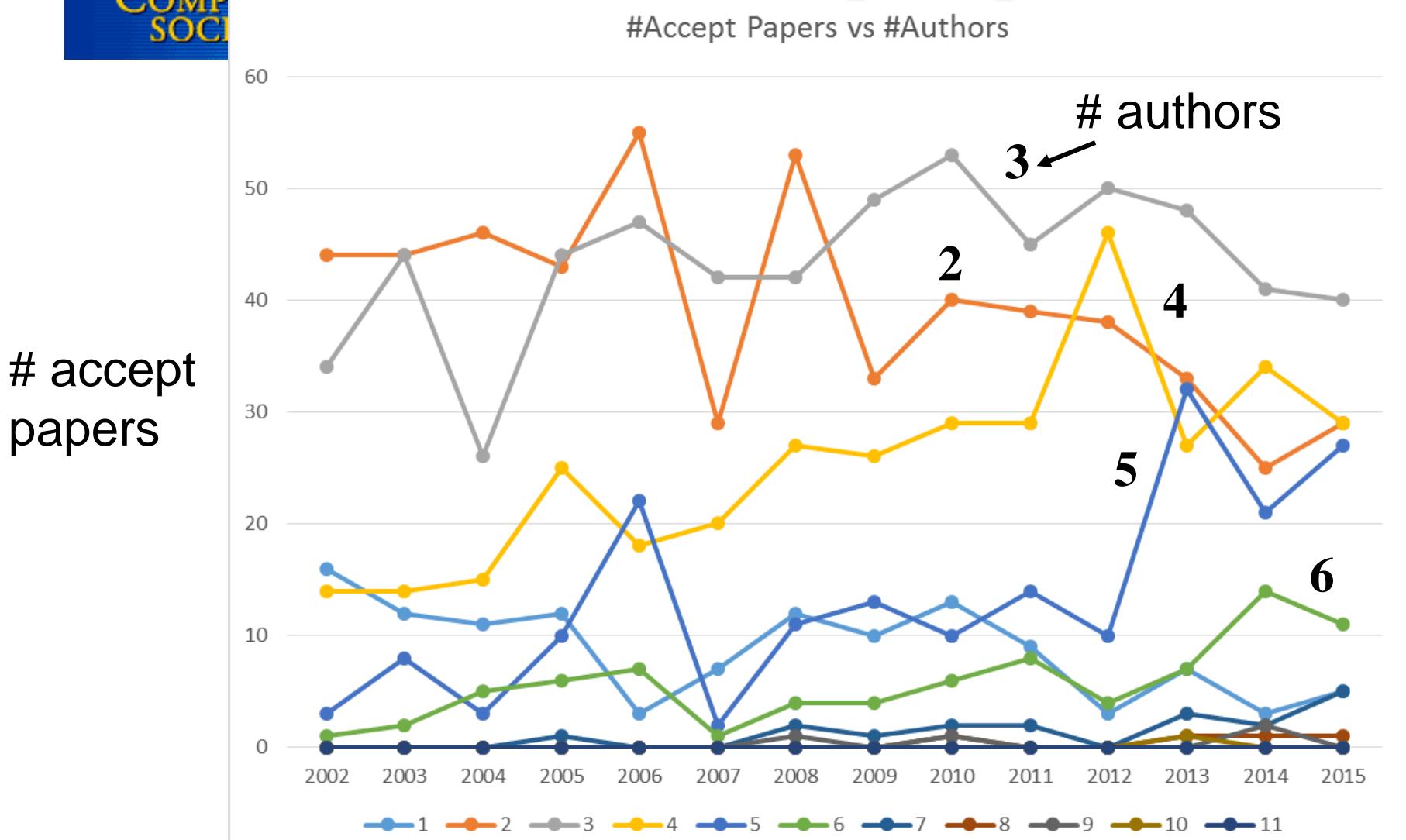
CA factor map





Comparison between 2002-2015

#Accept Papers





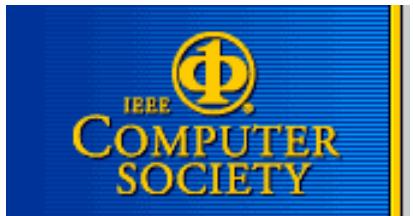
WordCloud (R) Accept Papers (Title)

2002

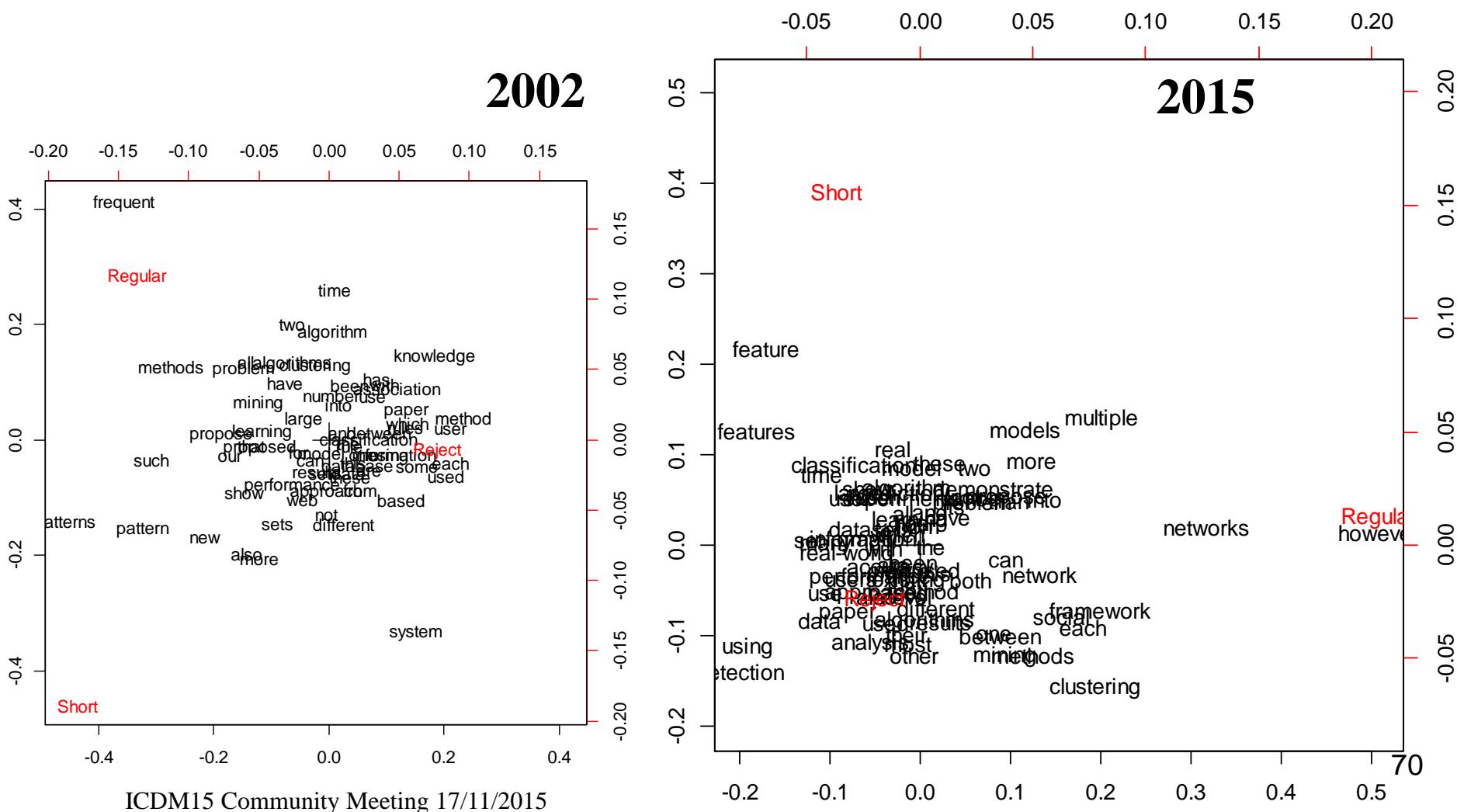
A word cloud visualization for the year 2002. The most prominent word is "mining" in a large, dark gray font. Other significant words include "data" (yellow), "learning" (orange), "classification" (purple), "rules" (pink), "clustering" (blue), "patterns" (blue), "discovery" (orange), "series" (orange), "text" (orange), "high" (orange), "analysis" (orange), "based" (orange), "using" (pink), "algorithm" (orange), "web" (orange), "time" (blue), "frequent" (blue), "selection" (orange), "models" (blue), "networks" (purple), "via" (orange), "sparse" (blue), "detection" (blue), "association" (blue), and "approach" (orange).

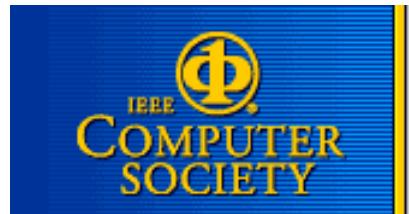
2015

A word cloud visualization for the year 2015. The most prominent word is "learning" in a large, dark gray font. Other significant words include "data" (gray), "mining" (green), "classification" (pink), "graphs" (orange), "feature" (orange), "ensemble" (orange), "dynamics" (orange), "multiple" (orange), "information" (orange), "systems" (orange), "matrix" (orange), "model" (pink), "clustering" (pink), "largescale" (blue), "detection" (blue), "sparse" (blue), "online" (blue), "time" (blue), "via" (orange), "networks" (purple), "hierarchical" (blue), "based" (blue), "social" (blue), "recommendation" (blue), "heterogeneous" (orange), "prediction" (orange), and "network" (orange).



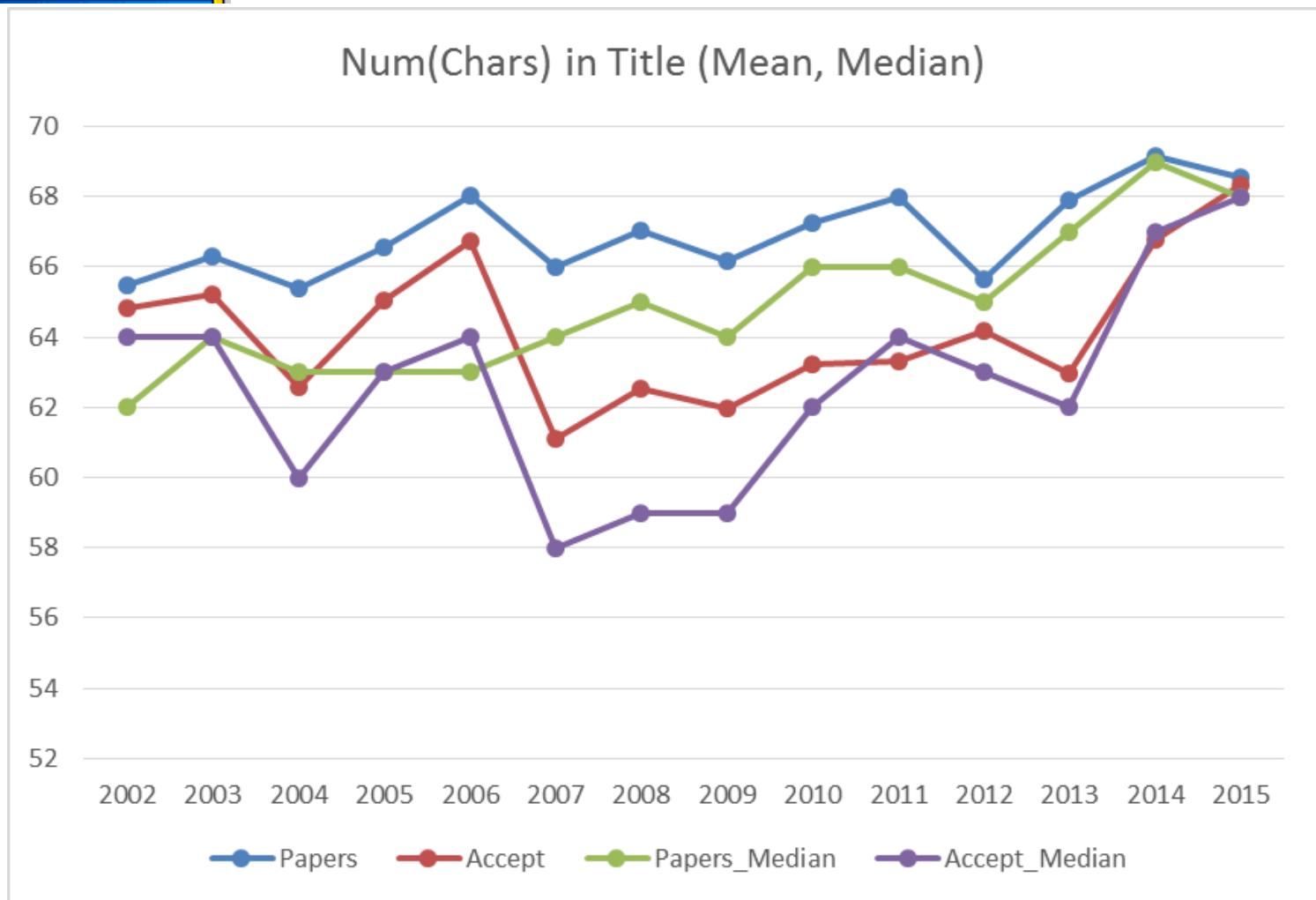
Text Mining by Correspondence Analysis (Abstract: 2002, 2015)





ICDM rule ?

Num of Char: larger => Reject ?





Data Mining on ICDM Submission Data

- Acknowledgements
 - Many thanks to
 - PC Chairs, Area Chairs and PC members
 - All the authors
 - All the contributors to ICDM 2015
 - See you next year in Barcelona!



Thank You!

- In the afternoon,
 - 1 Panel
 - 1 Tutorial
 - 6 Technical Sessions