



The Scholarly Impact of CLEF (2000-2009)

Theodora Tsikrika, Birger Larsen

Royal School of Library and Information Science, University of Copenhagen, Denmark

Henning Müller

University of Applied Sciences Western Switzerland (HES-SO), Sierre, Switzerland

Stefan Endrullis, Erhard Rahm

University of Leipzig, Germany

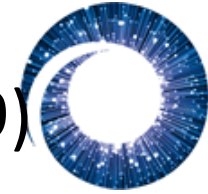


Motivation



- Evaluating evaluation campaigns
- Assess the impact of research not possible otherwise
 - Scientific impact
 - Scholarly impact, filed patents, etc.
 - Economic impact
 - Technology transfer, time & effort saved for researchers, etc.
- Goal: Assess the scholarly impact of CLEF (2000-2009)
 - Thornley, Johnson, Smeaton, & Lee. The scholarly impact of TRECVID (2003–2009). JASIST 2011.
 - Tsikrika, Seco de Herrera, & Müller. Assessing the scholarly impact of ImageCLEF. CLEF 2011.
 - Rowe, Wood, Link, & Simoni. Economic impact assessment of NIST's Text REtrieval Conference (TREC) Program. Technical Report, 2010.

CLEF evaluation campaign (2000-2009)



- Structured in Labs and/or Tasks
 - 14 Labs:
 - Adhoc, CL-SR, CLEF-IP, Domain-Specific, GeoCLEF, GRID@CLEF, iCLEF, ImageCLEF, INFILE, LogCLEF, MorphoChallenge, QA@CLEF, VideoCLEF, WebCLEF
 - 17 Tasks: AdHoc (4), ImageCLEF (7), QA (6)
- CLEF publications
 - CLEF Working Notes (NoteBooks)
 - **CLEF proceedings**
 - Papers describing or using CLEF resources published elsewhere
- CLEF proceedings publications types
 - Participant
 - Overview
 - Evaluation

Scholarly impact analysis method



Bibliometric analysis

1. Publication data collection

- Source: DBLP
- **Annotation** (automatic + manual validation)
 - Type : evaluation, overview, participant
 - Lab + Task

2. Citation data collection

- Thomson Reuters (ISI) Web of Science
- Scopus**
- Google Scholar**
 - Publish or Perish (**Pop**) – Anne-Wil Harzing
 - Online Citation Service (**OCS**) – University of Leipzig

3. Data analysis

Google Scholar shortcomings



- Several entries for the same publication
 - Due to misspellings
 - May deflate citation count
 - PoP: supports manual merging of publications
 - OCS: multiple matches merged manually
- Grouping citations of different papers
 - With similar titles and author lists
 - E.g., conference and journal versions of a paper
 - Manual data cleaning
- Inability to correctly identify publication year

CLEF proceedings publications



Number of CLEF 2000-2009 publications : 873

- Scopus indexes 92% (2000 – 2001 missing)
- Google Scholar indexes 97.5%

Country		Affiliation		Author	
Spain	178	Universidad de Alicante	44	Jones G.J.F.	29
Germany	105	UNED	33	Mandl T.	25
United States	93	Dublin City University	30	Llopis F.	24
France	67	University of Amsterdam	29	de Rijke M.	24
United Kingdom	61	Universidad de Jaen	27	Garcia-Cumbreras M.A.	20
Italy	55	Universität Hildesheim	25	Urena-Lopez L.A.	20
Netherlands	54	Universidad Carlos III de Madrid	24	Clough P.	19
Switzerland	52	UC Berkeley	23	Penas A.	18
Ireland	41	Universidad Politecnica de Madrid	22	Rosso P.	18
Canada	25	University of Sheffield	21	Leveling J.	17

* CLEF 2002-2009 proceedings (source: Scopus)

CLEF proceedings citations



	# labs # publ.		OCS			PoP			Scopus		
			# cit.	avg.	h-index	# cit.	avg.	h-index	# cit.	avg.	h-index
2000	3	27	501	18.56	15	507	18.78	15	-	-	-
2001	2	37	904	24.43	17	901	24.35	17	-	-	-
2002	4	44	636	14.45	14	634	14.41	14	74	1.68	4
2003	6	65	787	12.11	15	776	11.94	15	87	1.34	5
2004	6	81	989	12.21	17	942	11.63	16	137	1.69	5
2005	8	112	1231	10.99	18	1207	10.78	17	133	1.19	5
2006	8	127	1278	10.06	18	1250	9.84	18	133	1.05	5
2007	7	116	1028	8.86	16	902	7.78	15	119	1.03	5
2008	10	131	1002	7.65	16	989	7.55	16	78	0.60	3
2009	10	133	781	5.87	12	770	5.79	12	144	1.08	5
Total	14	873	9,137	10.47	41	8,878	10.17	41	905	1.04	10

CLEF proceedings citations



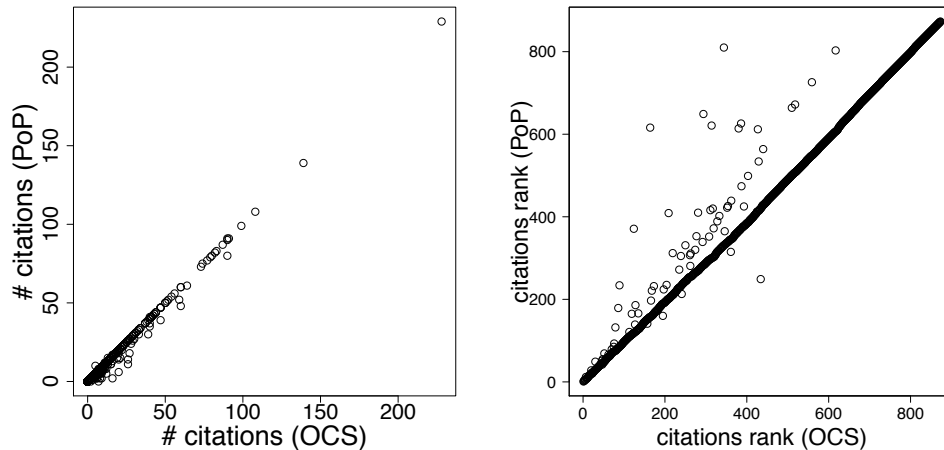
	# labs # publ.		OCS			PoP			Scopus		
			# cit.	avg.	h-index	# cit.	avg.	h-index	# cit.	avg.	h-index
2000	3	27	501	18.56	15	507	18.78	15	-	-	-
2001	2	37	904	24.43	17	901	24.35	17	-	-	-
2002	4	44	636	14.45	14	634	14.41	14	74	1.68	4
2003	6	65	787	12.11	15	776	11.94	15	87	1.34	5
2004	6	81	989	12.21	17	942	11.63	16	137	1.69	5
2005	8	112	1231	10.99	18	1207	10.78	17	133	1.19	5
2006	8	127	1278	10.06	18	1250	9.84	18	133	1.05	5
2007	7	116	1028	8.86	16	902	7.78	15	119	1.03	5
2008	10	131	1002	7.65	16	989	7.55	16	78	0.60	3
2009	10	133	781	5.87	12	770	5.79	12	144	1.08	5
Total	14	873	9,137	10.47	41	8,878	10.17	41	905	1.04	10

OCS finds 259 more citations (+3%)

OCS vs. PoP

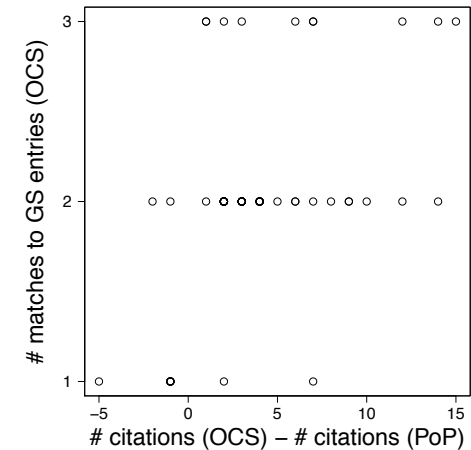


Correlations



- Strong correlations for number of citations and rankings
- Overlap in top $k = \{100, 200, \dots, 500\}$ > 96%

Differences



- Difference per publication: 1-15 citations
- Higher difference → More likely OCS found more matches

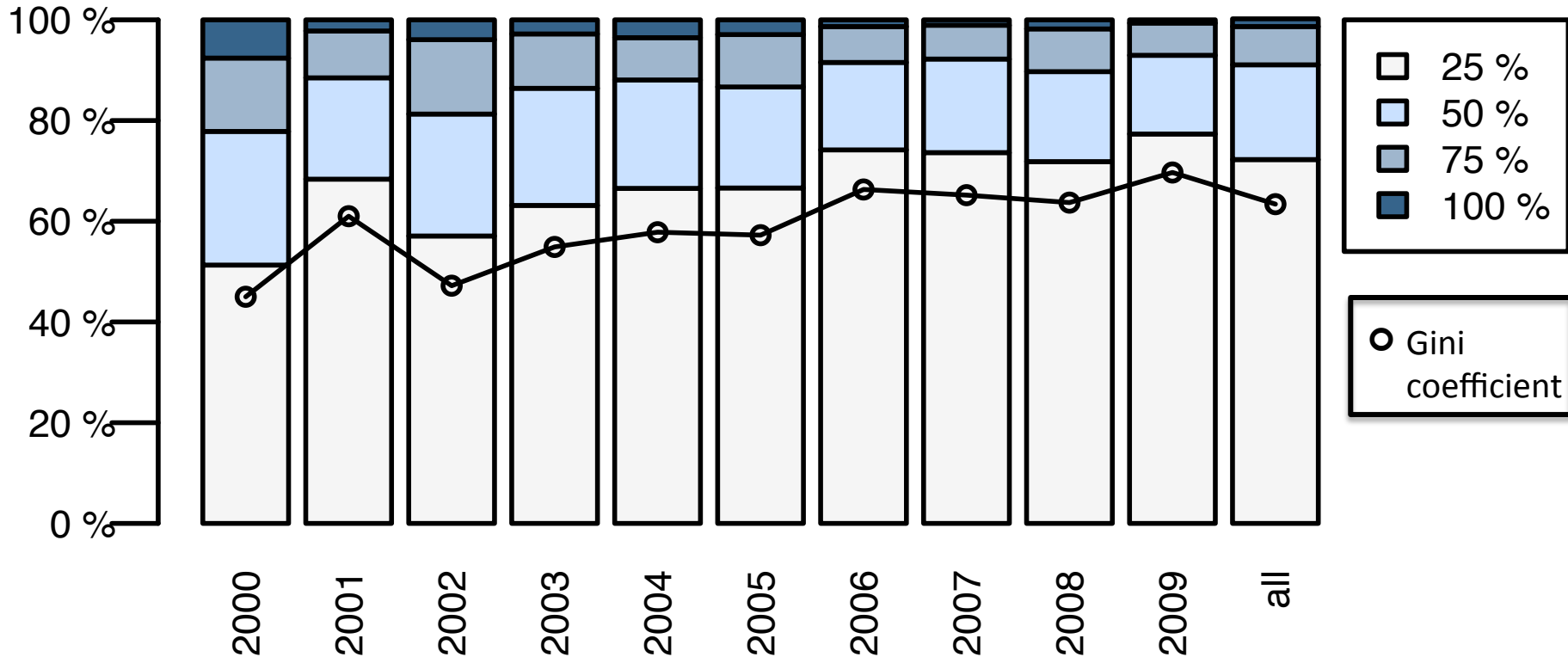
In essence comparable, with different querying facilities, may prove advantageous in different situations

Top CLEF proceedings publications



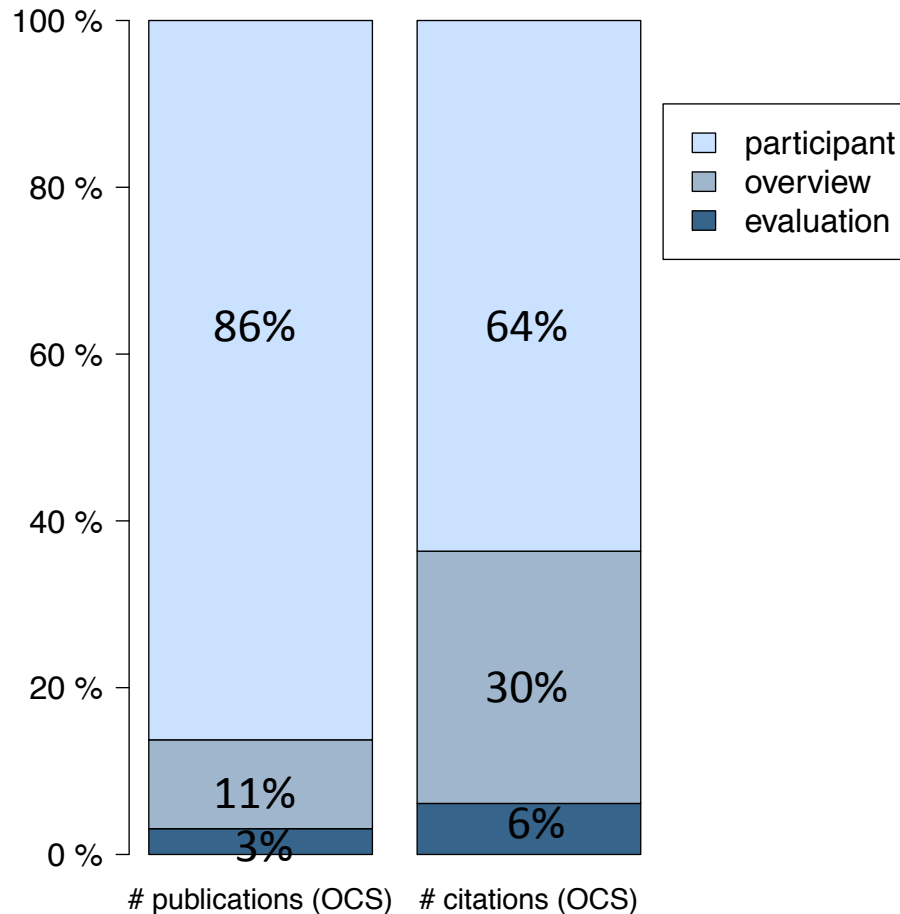
rank	# citations	Author(s)	Title	Year	Type
1	228	Voorhees	The Philosophy of Information Retrieval Evaluation.	2001	E
2	139	Müller et al.	Overview of the ImageCLEFmed 2006 Medical Retrieval and Medical Annotation Tasks.	2006	O
3	108	Clough et al.	The CLEF 2005 Cross-Language Image Retrieval Track.	2005	O
4	99	Clough et al.	The CLEF 2004 Cross-Language Image Retrieval Track.	2004	O
5	91	Vallin et al.	Overview of the CLEF 2005 Multilingual QA Track.	2005	O
6	90	Chen	Cross-Language Retrieval Experiments at CLEF 2002.	2002	P
7	90	Grubinger et al.	Overview of the ImageCLEFphoto 2007 Photographic Retrieval Task.	2007	O
8	90	Monz & de Rijke	Shallow Morphological Analysis in Monolingual Information Retrieval for Dutch, German and Italian	2001	P
9	87	Müller et al.	Overview of the CLEF 2009 Medical Image Retrieval Track.	2009	O
10	83	Magnini et al.	Overview of the CLEF 2004 Multilingual QA Track.	2004	O

Citation distribution



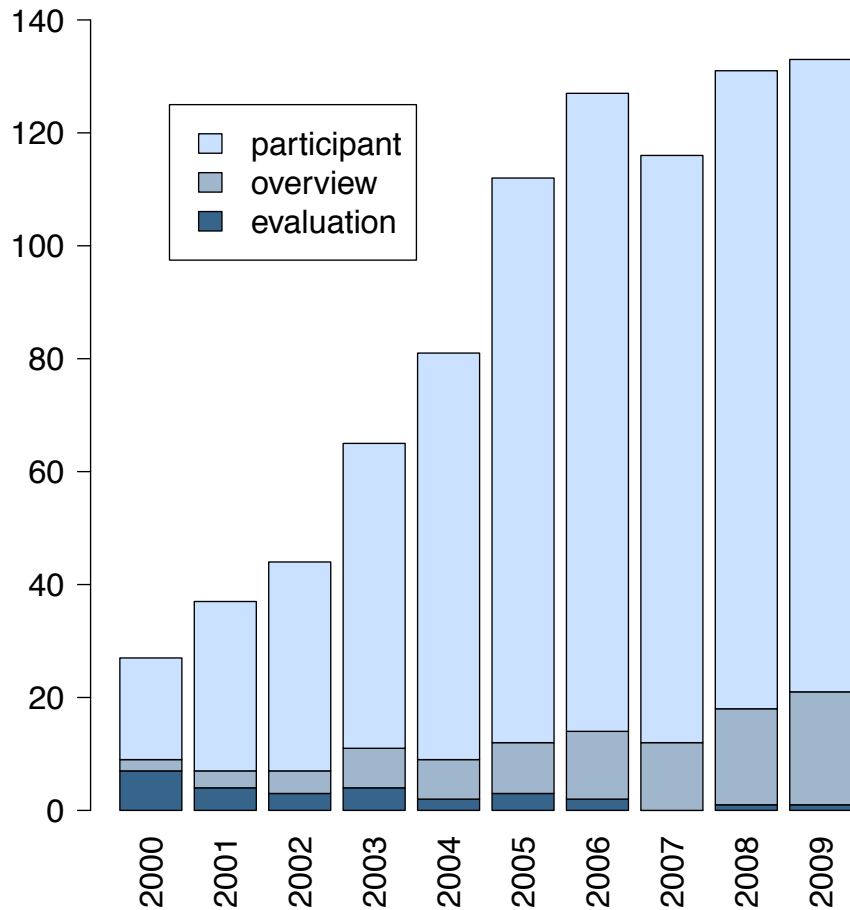
- Top 25% cited publications → 50-75% of all citations (72% avg.)
- Bottom 25% cited publications → 0.5-7.5% of all citations (1.5% avg.)
- Citation skewness increasing over the years

CLEF publications types

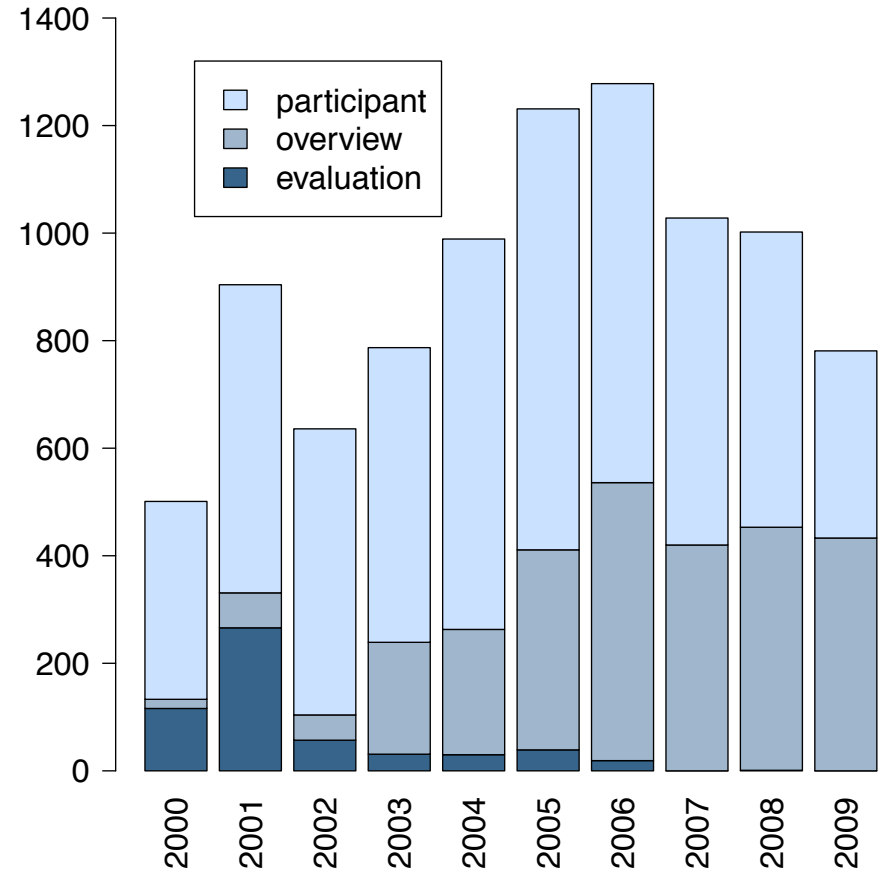


- **Overviews:**
relative citation frequency =
3 x relative publication frequency
- 7 of top 10 most cited are
overviews
- 60% of the papers with
 $\#citations \geq h\text{-index}$ are
overviews
- Impact of overviews
increases over the years

CLEF publications types

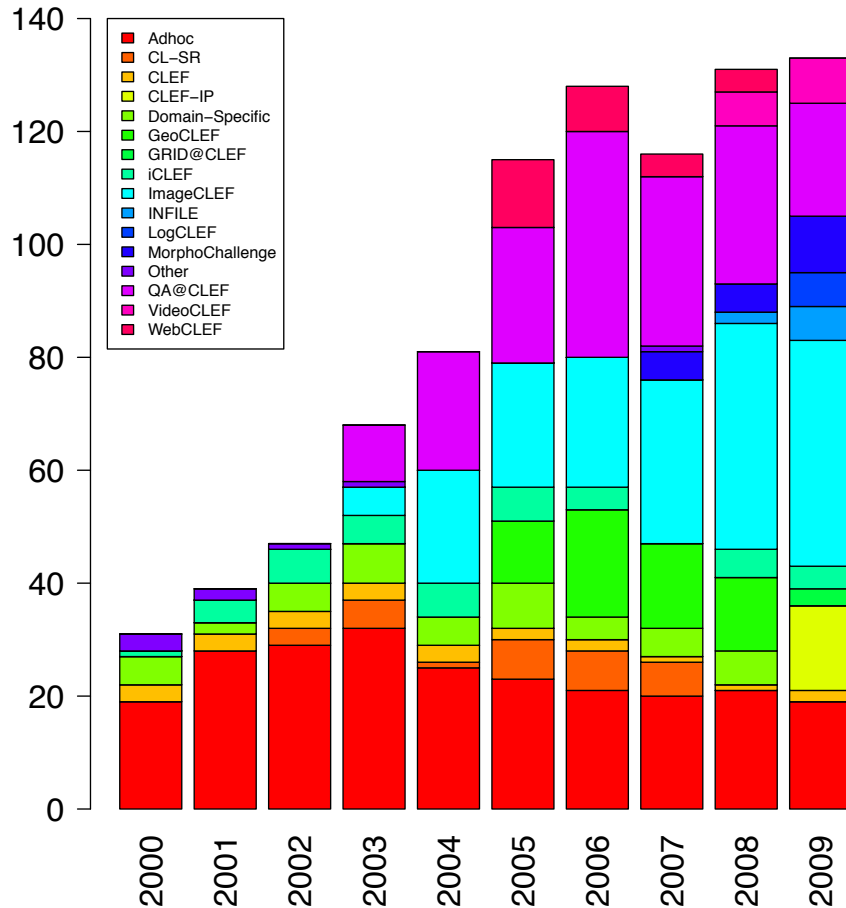


publications

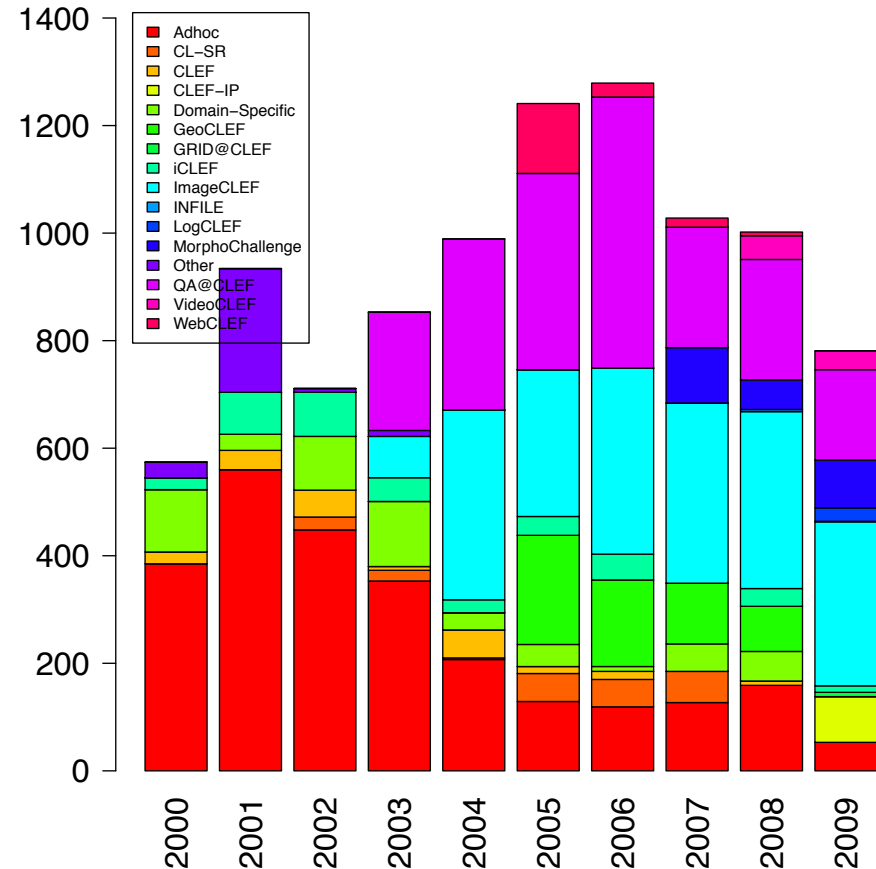


citations

CLEF Labs scholarly impact



publications



citations

CLEF Labs scholarly impact



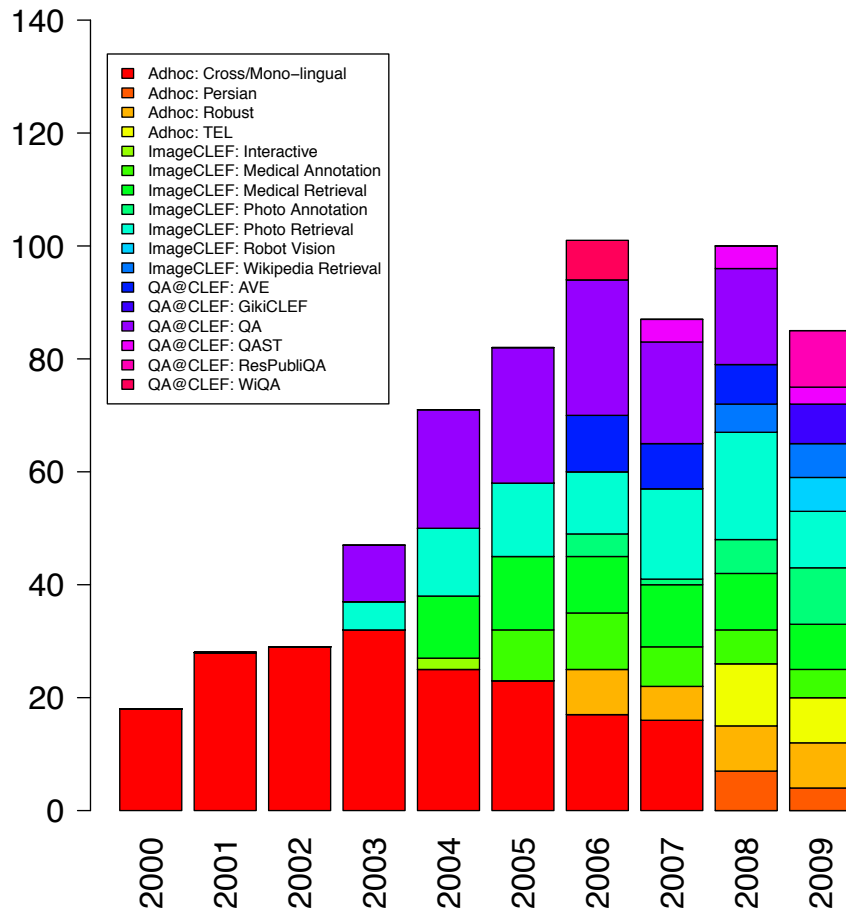
	#years	# publications	# citations	# citations / publication
Adhoc	10	237	2,540	10.72
CL-SR	6	29	208	7.17
CLEF	10	23	203	8.83
CLEF-IP	1	15	85	5.67
Domain-Specific	9	47	555	11.81
GeoCLEF	4	58	561	9.67
GRID@CLEF	1	3	8	2.67
iCLEF	9	41	378	9.22
ImageCLEF	7	179	2,018	11.27
INFILE	2	8	5	0.62
LogCLEF	1	6	25	4.17
MorphoChallenge	3	20	247	12.35
Other	5	8	277	34.62
QA@CLEF	7	173	2,023	11.69
VideoCLEF	2	14	79	5.64
WebCLEF	4	28	180	6.43
All	10	873	9,137	10.47

CLEF Labs scholarly impact

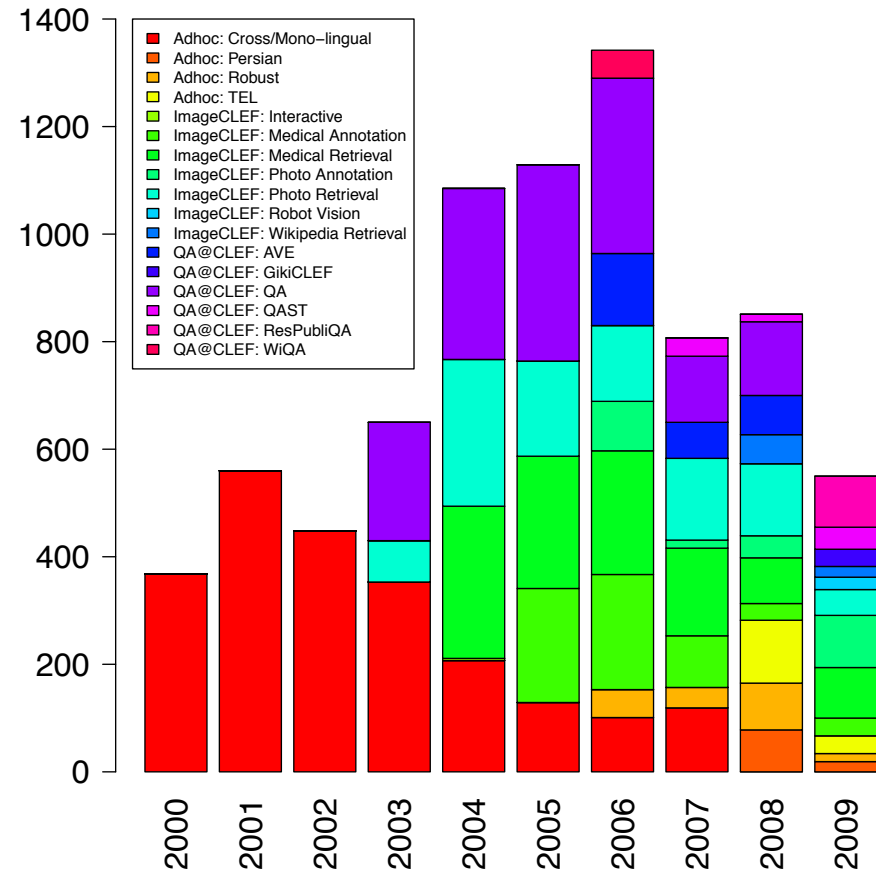


	#years	# publications	# citations	# citations / publication
Adhoc	10	237	2,540	10.72
CL-SR	6	29	208	7.17
CLEF	10	23	203	8.83
CLEF-IP	1	15	85	5.67
Domain-Specific	9	47	555	11.81
GeoCLEF	4	58	561	9.67
GRID@CLEF	1	3	8	2.67
iCLEF	9	41	378	9.22
ImageCLEF	7	179	2,018	11.27
INFILE	2	8	5	0.62
LogCLEF	1	6	25	4.17
MorphoChallenge	3	20	247	12.35
Other	5	8	277	34.62
QA@CLEF	7	173	2,023	11.69
VideoCLEF	2	14	79	5.64
WebCLEF	4	28	180	6.43
All	10	873	9,137	10.47

CLEF Tasks scholarly impact



publications



citations

CLEF Tasks scholarly impact



	#years	# publications	# citations	# citations / publication
Adhoc	10	237	2,540	10.72
Cross/Mono-lingual	8	188	2285	12.15
Persian	2	11	97	8.82
Robust	4	30	192	6.40
TEL	2	19	150	7.89
ImageCLEF	7	179	2,018	11.27
Interactive	1	2	4	2.00
Medical Annotation	5	37	586	15.84
Medical Retrieval	6	62	1002	16.16
Photo Annotation	4	21	245	11.67
Photo Retrieval	7	86	1002	11.65
Robot Vision	1	6	23	3.83
Wikipedia Retrieval	2	11	74	6.73
QA@CLEF	7	173	2,023	11.69
AVE	3	25	274	10.96
GikiCLEF	1	7	32	4.57
QA	6	114	1489	13.06
QAST	3	11	89	8.09
ResPubliQA	1	10	95	9.50
WiQA	1	7	52	7.43

CLEF Tasks scholarly impact



	#years	# publications	# citations	# citations / publication
Adhoc	10	237	2,540	10.72
Cross/Mono-lingual	8	188	2285	12.15
Persian	2	11	97	8.82
Robust	4	30	192	6.40
TEL	2	19	150	7.89
ImageCLEF	7	179	2,018	11.27
Interactive	1	2	4	2.00
Medical Annotation	5	37	586	15.84
Medical Retrieval	6	62	1002	16.16
Photo Annotation	4	21	245	11.67
Photo Retrieval	7	86	1002	11.65
Robot Vision	1	6	23	3.83
Wikipedia Retrieval	2	11	74	6.73
QA@CLEF	7	173	2,023	11.69
AVE	3	25	274	10.96
GikiCLEF	1	7	32	4.57
QA	6	114	1489	13.06
QAST	3	11	89	8.09
ResPubliQA	1	10	95	9.50
WiQA	1	7	52	7.43

CLEF Tasks scholarly impact



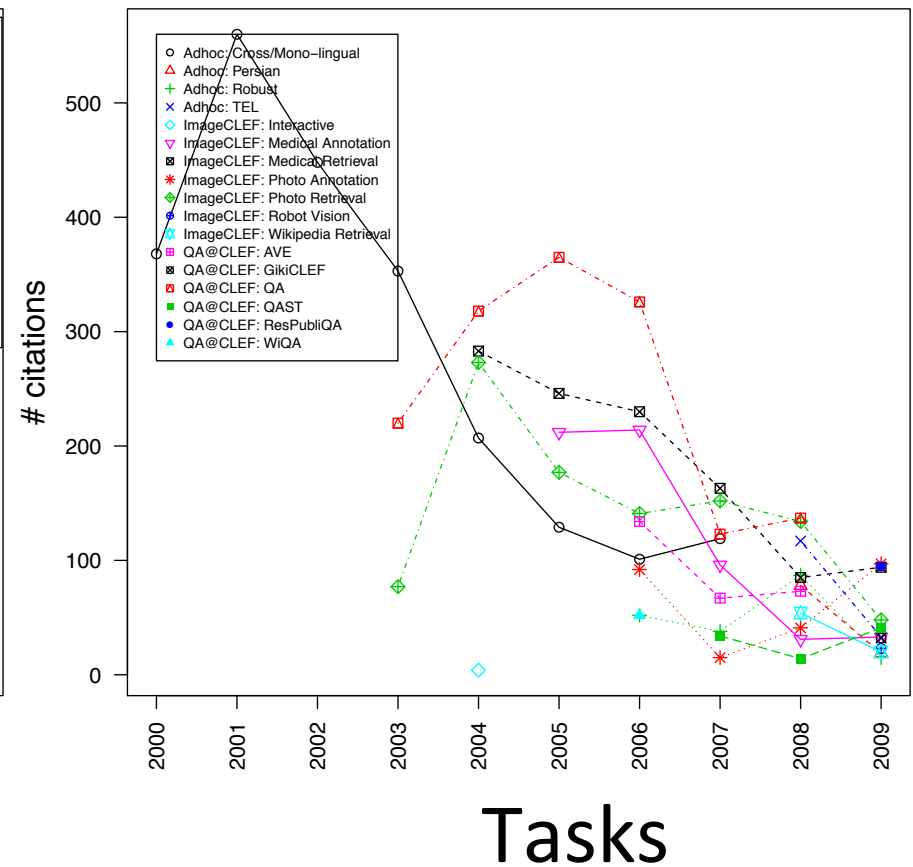
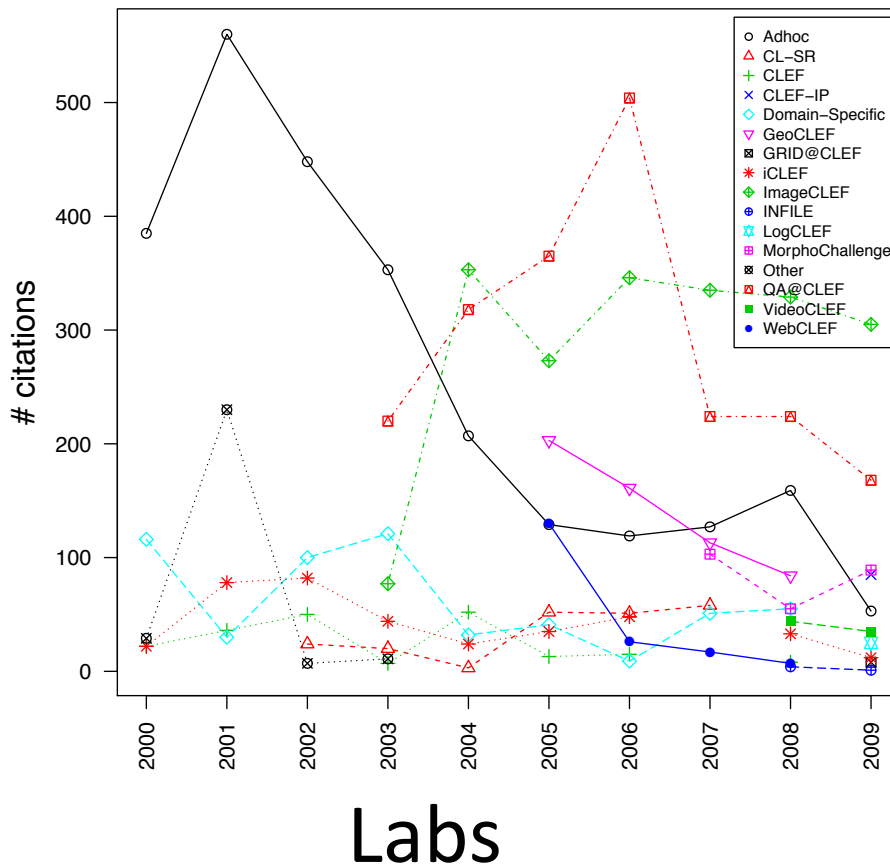
	#years	# publications	# citations	# citations / publication
Adhoc	10	237	2,540	10.72
Cross/Mono-lingual	8	188	2285	12.15
Persian	2	11	97	8.82
Robust	4	30	192	6.40
TEL	2	19	150	7.89
ImageCLEF	7	179	2,018	11.27
Interactive	1	2	4	2.00
Medical Annotation	5	37	586	15.84
Medical Retrieval	6	62	1002	16.16
Photo Annotation	4	21	245	11.67
Photo Retrieval	7	86	1002	11.65
Robot Vision	1	6	23	3.83
Wikipedia Retrieval	2	11	74	6.73
QA@CLEF	7	173	2,023	11.69
AVE	3	25	274	10.96
GikiCLEF	1	7	32	4.57
QA	6	114	1489	13.06
QAST	3	11	89	8.09
ResPubliQA	1	10	95	9.50
WiQA	1	7	52	7.43

CLEF Tasks scholarly impact



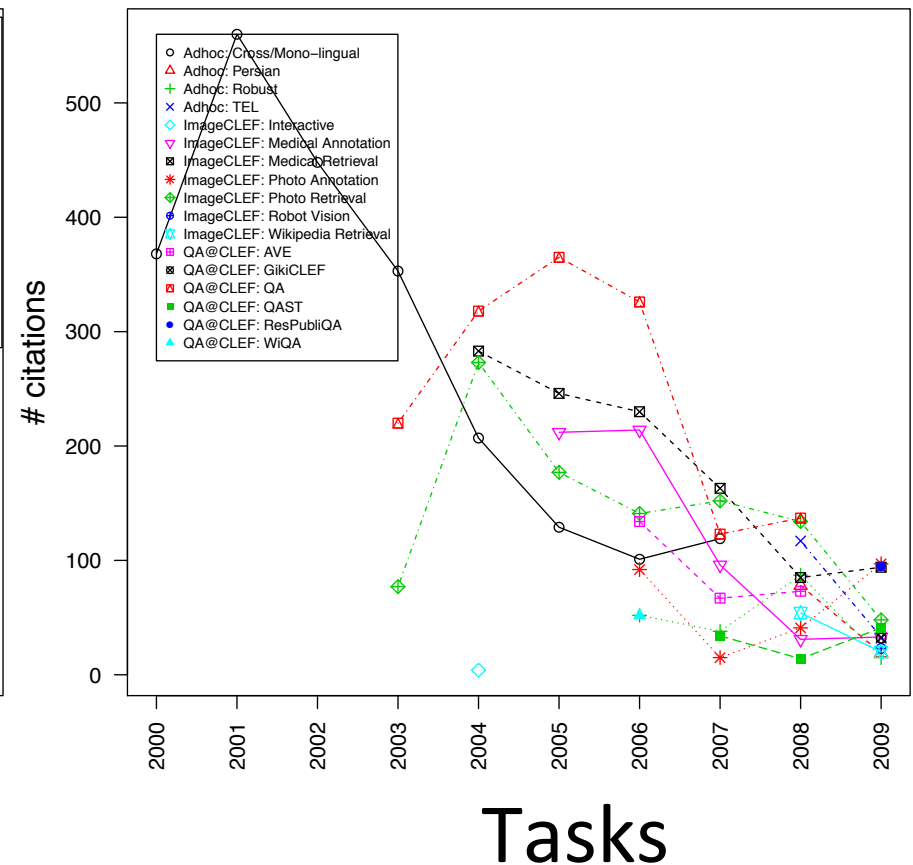
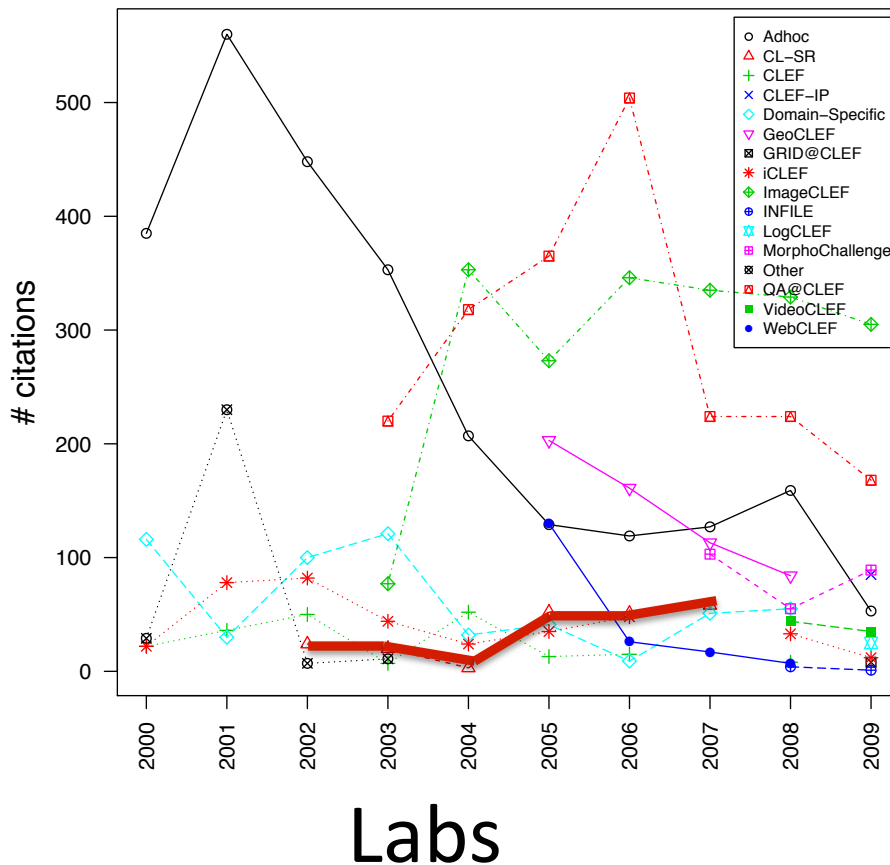
	#years	# publications	# citations	# citations / publication
Adhoc	10	237	2,540	10.72
Cross/Mono-lingual	8	188	2285	12.15
Persian	2	11	97	8.82
Robust	4	30	192	6.40
TEL	2	19	150	7.89
ImageCLEF	7	179	2,018	11.27
Interactive	1	2	4	2.00
Medical Annotation	5	37	586	15.84
Medical Retrieval	6	62	1002	16.16
Photo Annotation	4	21	245	11.67
Photo Retrieval	7	86	1002	11.65
Robot Vision	1	6	23	3.83
Wikipedia Retrieval	2	11	74	6.73
QA@CLEF	7	173	2,023	11.69
AVE	3	25	274	10.96
GikiCLEF	1	7	32	4.57
QA	6	114	1489	13.06
QAST	3	11	89	8.09
ResPubliQA	1	10	95	9.50
WiQA	1	7	52	7.43

CLEF Labs & Tasks trends



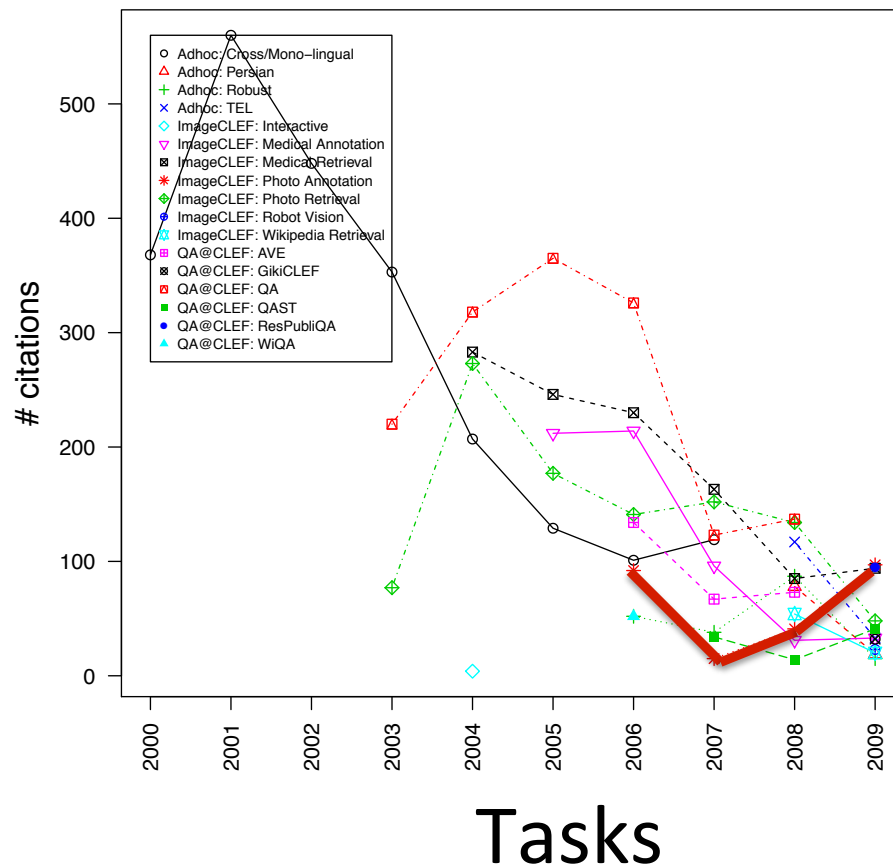
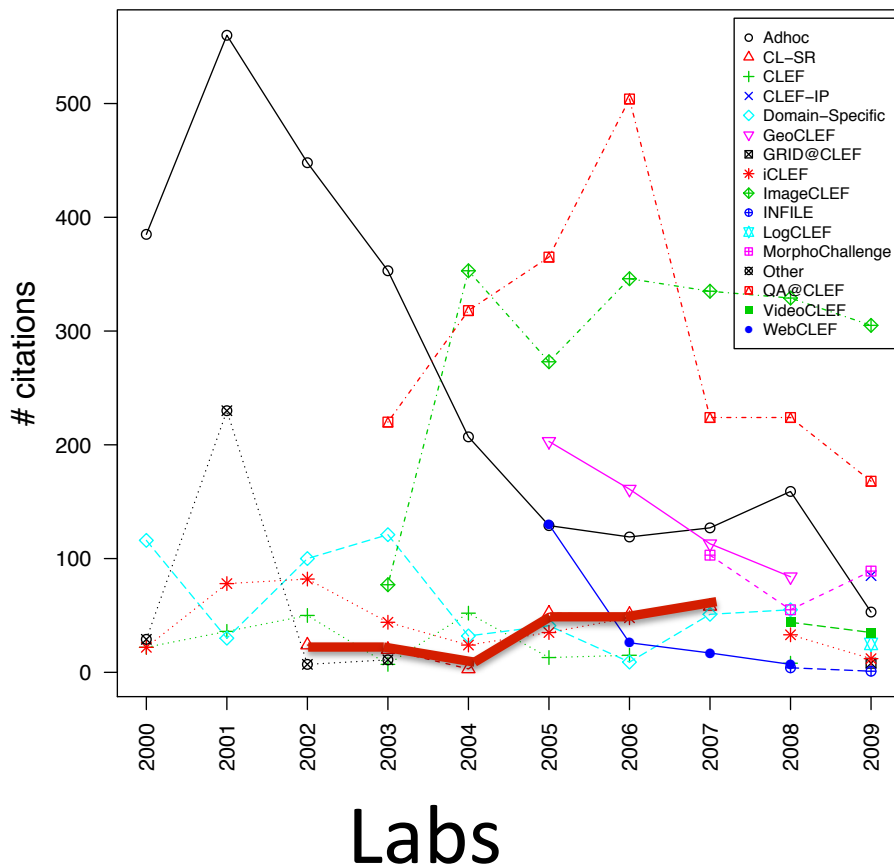
Peak in second or third year of operation, followed by a decline, ... unless there is a major overhaul of the task

CLEF Labs & Tasks trends



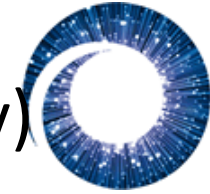
Peak in second or third year of operation, followed by a decline, ... unless there is a major overhaul of the task

CLEF Labs & Tasks trends

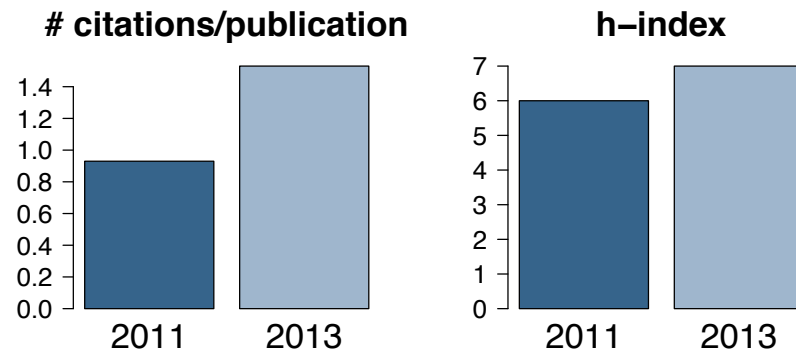
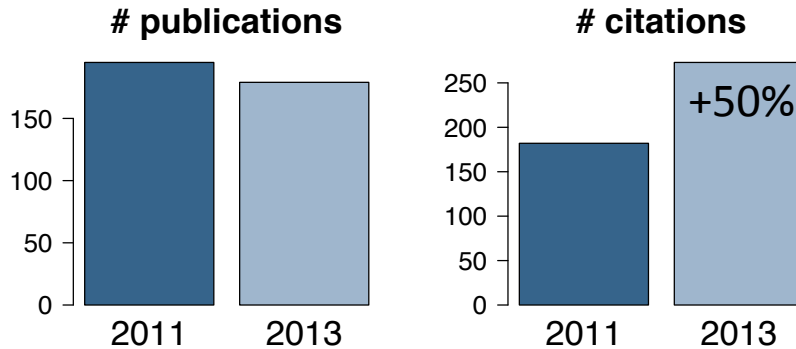


Peak in second or third year of operation, followed by a decline, ... unless there is a major overhaul of the task

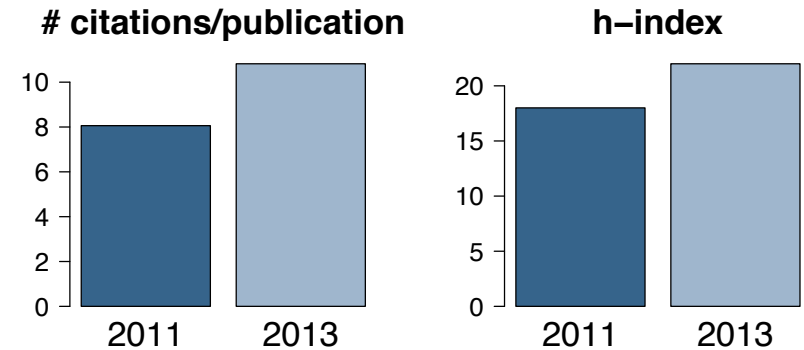
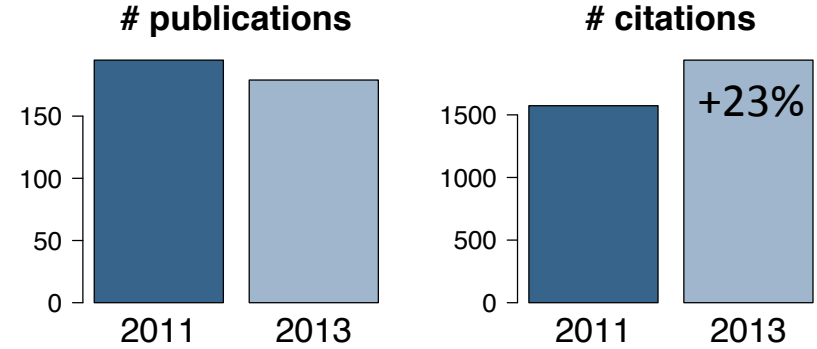
ImageCLEF impact (2011* vs. 2013 study)



Scopus



PoP





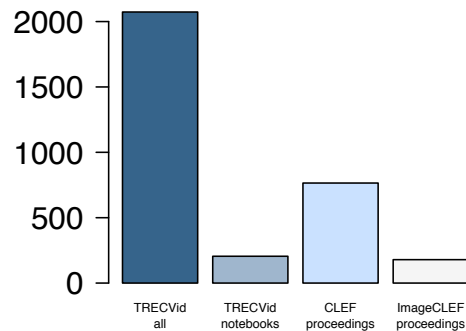
What to compare against?

- Different fields → Different citation practices
 - CLEF is multi-disciplinary
- Identify similar fields
 - Level of granularity? Lab? Task? Other?
 - Bibliometric studies not readily available
- Other evaluation campaigns
 - In similar fields

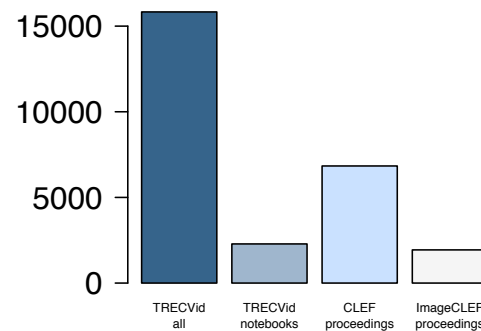
Comparison to TREC Vid* (2003-2009)



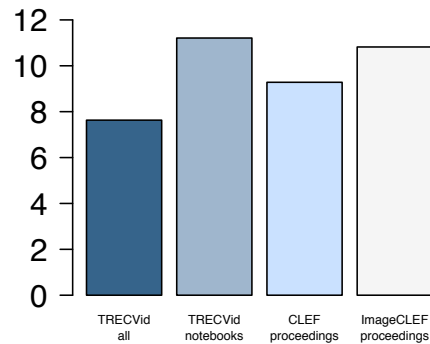
publications



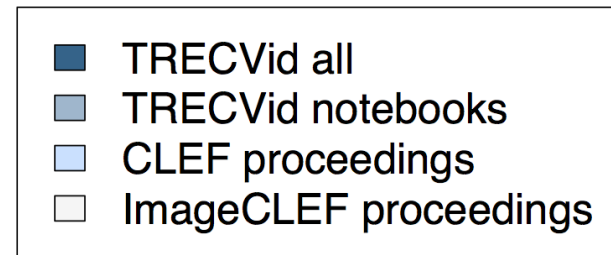
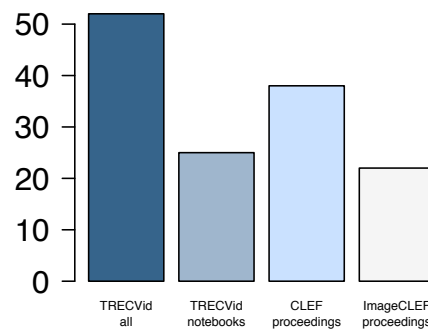
citations



citations/publication



h-index



Conclusions & Future work



- Impact assessment supports research policy decisions
- **Bibliometric analysis of CLEF proceedings**
 - 873 publications with 9,137 citations (10.47 citations / publication)
 - Skewed citation distribution
 - High impact of overview papers
 - Labs/Tasks need to update every 2-3 years
 - TRECVID vs. ImageCLEF : comparable #citations/publication
- Next steps:
 - Include CLEF working notes + CLEF derived publications
 - Automatically learn the task/dataset associated with a publication
 - Compare against other baselines



Thank you!