

Metodologias de Investigação/Research Methodologies

Rute Sofia (rute.sofia@ulusofona.pt)

ULHT, Campo Grande
14.12.2011
(Aula 8)

Agenda

Fontes de Pesquisa, e factores de qualidade

Conferências e Jornais – seleccionar

Padronização/Estandardização

Fontes de Pesquisa

- **Google** ☺
- **ACM – Association for Computing Machinery**
 - ACM é uma associação organizada em SIGs - Special Interest Groups
 - SIGCOMM – data communication
 - SIGARCH – computer architecture
 - SIGCHI – multimedia, interacção homem-máquina
 - SIGMIS - Management Information Systems
 - <http://www.acm.org/conferences>
- **IEEE – Institute of Electronics and Electrical Engineers**
 - IEEExplore – base de dados online
- **Elsevier Science Direct**
 - [http:// www.sciencedirect.com](http://www.sciencedirect.com)
- **ISI Web of Science**
- **Wiki Call for Papers (<http://www.wikicfp.com/cfp/>)**



Algumas Fontes – ISI Web of Science

- **Thomson Reuters detém uma base de dados científica – ISI Web of Science**
 - Mais de 8000 jornais científicos
 - Mais de 3000 ciências sociais
 - Mais de 1800 em Artes e Humanidades
 - WoS permite
 - Manter track record de referências citadas
 - Manter track record de número de citações
 - Pesquisar autores mediante factor de impacto
 - ...

Thomson, Factor de Impacto

- **Aplicado a jornais**
- **Journal Citation Report**
 - Editado anualmente
 - Dados da WoS, mas apresentados de um modo mais legível
 - Apresenta factor de impacto de jornais

Para um Jornal de 2010, factor de impacto é calculado como:

Número de vezes que artigos publicados em 2008 e 2009 foram citados em jornais indexados durante 2010

Número de artigos passíveis de serem citados (indexados) em 2007 e 2008

Factor de Impacto, problemas

- **Thomson Reuters detém uma base de dados científica – ISI Web of Science**
 - Apenas um “pequeno” número de jornais é indexado pela ISI
 - Conferências na área de ciências informáticas só começaram a ser indexadas há aprox. 2 anos
 - Maioria do nosso trabalho é publicado em conferências
 - Jornais/revistas são maioritariamente em língua inglesa
 - Método de calcular citações não é homogéneo – outros ambientes (e.g. GoogleScholar) utilizam outros métodos
 - Por exemplo, inclui citações do autor
 - Subscrição (efectuada por instituições) é cara...

Eigenfactor – um outro mecanismo qualitativo

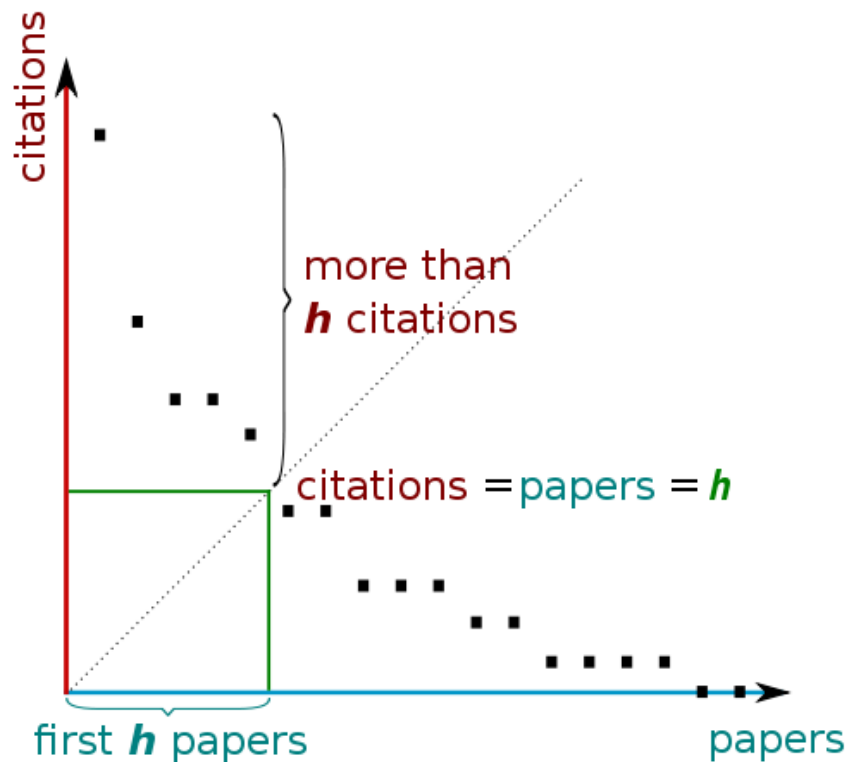
- Acessível em eigenfactor.org
 - Tem dados de 1995 a 2009
- Apenas indexa jornais da ISI mas considera todos os dados armazenados
- Algoritmo de pesquisa semelhante ao PageRank da Google
- Considera 5 anos

Journal	JCR Journal Impact Factor (2 year)	JCR Journal Impact Factor (5 year)	Eigenfactor Article Influence Score (5 year)	SCImago Journal Rank (3 year)
Science	26.372	30.631	16.539	3.726
Cell	29.887	28.779	18.188	10.735
Nature	28.751	28.751	16.996	4.636
PNAS	9.598	10.369	4.929	2.689
BMC Bioinformatics	3.493	4.221	1.608	.750
Bioinformatics	5.039	6.649	2.406	1.225



Avaliação de Autores – h-index

- **H de...Herfindahl index**
 - Herfindahl–Hirschman Index
 - Indicator de origem económica, medida de competição
- **Thomson Reuters detém uma base de dados científica – ISI Web of Science**
 - Método de quantificar valor do output científico de autorse
 - Proposto por Jorge E. Hirsch (Universidade da Califórnia, San Diego)
 - Combina quantidade (número de jornais) e qualidade (impacto e/ou citações)
 - H-index é automaticamente calculado nas ferramentas de pesquisa (ISI WoS; Scopus, etc)



Avaliação de autores - Publish or Perish (PoP)

- Desenvolvido por harzing.com
- Utiliza GoogleScholar e fornece
 - Média de citações por artigo
 - Média de autores por artigo
- Fornece h-index
- Versão grátis para academia: harzing.com/pop.htm

The screenshot displays the 'Author impact analysis' tool interface. At the top, there are navigation tabs: 'Author impact', 'Journal impact', 'General citations', 'Multi-query center', and 'Web Browser'. The main heading is 'Author impact analysis - Perform a citation analysis for one or more authors'. Below this, there is a search form with the following fields: 'Author's name' (containing '*martinis sa'), 'Exclude these names' (empty), and 'Year of publication between' (0 and 0). To the right of the search form are several checkboxes for subject categories: 'Biology, Life Sciences, Environmental Science' (checked), 'Business, Administration, Finance, Economics' (unchecked), 'Chemistry and Materials Science' (checked), 'Engineering, Computer Science, Mathematics' (checked), 'Medicine, Pharmacology, Veterinary Science' (checked), 'Physics, Astronomy, Planetary Science' (unchecked), and 'Social Sciences, Arts, Humanities' (unchecked). There are buttons for 'Lookup', 'Lookup Direct', and 'Help'. Below the search form is a 'Results' section with a summary table:

Papers:	86	Cites/paper:	20.65	h-index:	23	AWCR:	164.69
Citations:	1776	Cites/author:	533.47	g-index:	41	AW-index:	12.83
Years:	25	Papers/author:	31.15	hc-index:	14	AWCRpA:	49.93
Cites/year:	71.04	Authors/paper:	3.23	hI-index:	6.53	e-index:	29.98
				hI,norm:	13	hm-index:	11.00

Below the summary table is a list of citation results with columns: 'Cites', 'Per year', 'Rank', 'Authors', 'Title', 'Year', and 'Publication'. The first row is selected (checked) and shows 163 citations, a per-year average of 7.09, rank 1, authors 'SA Martinis, WM Atkins, PS Stayton...', and a title 'A conserved residue of cytochrome ...' published in 1989 in 'Journal of the American ...'. Other rows show various other papers with their respective citation counts and details.

Avaliação de autores - Publish or Perish (PoP)

- **JE Hirsch:**

“Based on typical h .. values found, I suggest (with large error bars) that for faculty at major research universities, $h \approx 12$ might be a typical value for advancement to tenure (associate professor) and that $h \approx 18$ might be a typical value for advancement to full professor. Fellowship in the American Physical Society might occur typically for $h \approx 15$ – 20 . Membership in the National Academy of Sciences of the United States of America may typically be associated with $h \approx 45$ and higher, except in exceptional circumstances.”

- As with other metrics, it's best to compare h -index values within a discipline.

Avaliação de autores - Comparação

Author	From Scopus	From Web of Science	From Google Scholar via “Publish or Perish”
Robinson, GE	54 [193 docs]	55 [202 docs]	55 [381 docs]
Wise, PM	44 [178 docs]	51 [204 docs]	50 [333 docs]
Garfield, E	22 [211 docs]	288 [815 docs]	45 [>1000 docs]
Hogan, MJ	-- [1 doc]	6 [33 docs]	14 [65 docs]

Agenda

Fontes de Pesquisa, e factores de qualidade

Conferências, factores de qualidade

Padronização/Estandardização

Conferências

- **Qualidade das conferências mede-se por...**
 - Indexação ISI WoS
 - Technical Programme Committee e Chairs
 - Historial de “acceptance rate” E número de artigos submetidos...
- **Acceptance rate...**
 - Número de artigos aceites/artigos submetidos
 - Acceptance rate acima de 20% - excelente conferência
 - Acceptance rate entre 20 e 40% - conferência boa
 - Acceptance rate 40-50% - Ok
 - Acceptance rate acima 50% - não muito boa (aceita maioria)
- **DICA: ao desenvolver vosso CV, devem**
 - **Colocar número de citações frente ao artigo**
 - **Colocar acceptance rate conferência**
 - **Colocar factor Impacto jornal onde artigo foi submetido...**

Conferências

- **Verificar Acceptance Rate**
 - Website de Kevin Almeroth
<http://www.cs.ucsb.edu/~almeroth/conf/stats/>
 - Website de Pradeep Pradala
 - <http://ppadala.net/conferences/>

- Página de K. Almeroth:

This page is an attempt to gather year-over-year statistics for some of the major systems and networking conferences. For a graphical version of these stats, see Pradeep Padala's page.

For a broader range of CFPs, see Tim Moors's WWW page.

For Architecture conference stats (ISCA, Micro, HPCA, ASPLOS), see the Prichard, Scopel, Hill, Sohi, and Wood Excel File.

For Software Engineering conference stats, see Tao Xie's Stats Page.

For Database conference stats, see Peter Aper's Stats Page.

For Graphics/Interaction/Vision conference stats, see Rob Lindeman's Stats Page.

For Computer Security conference stats, see Guofei Gu's Computer Security Conference Ranking and Statistics Page or Jianying Zhou's Crypto and Security Conferences Ranking.

Agenda

Fontes de Pesquisa, e factores de qualidade

Conferências, factores de qualidade

Padronização/Estandardização



Padronização/Estandarização - Standardisation

- **Padronizar é uma forma de disseminar ciência e de proteger resultados**
 - Permite auxiliar a passar de conceitos para produtos...
 - Padrões permitem assegurar que tecnologia desenvolvida por diferentes entidades funciona e é compatível...
- **Padronizar implica ter**
 - Uma solução única – passa a ser uma proposta de um standard
- **No contexto de um mestrado, como contribuir?**
 - Desenvolver propostas de standards
 - Participar (monitorizar e discutir) em grupos de padronização
- **Algumas entidades**
 - IETF – Internet Engineering Task Force
 - IRTF – Internet Research Task Force
 - IEEE – In
 - 3GPP –
 - Open-source – Open Networking Foundation

Padronização - IETF

- **Iniciada em 1986**
 - Agregou esforços da ARPA (Internet Configuration Control Board) e da Internet Activities Board
- **Movimento voluntário**
 - Cresceu de modo lento
 - Sem posição governamental – forneceu-lhe independência
 - Feito de indivíduos e não de entidades!
- **Organismos de gestão**
 - IETF Chair
 - Area Director for General Area, porta-voz
 - Seleccionado pela comunidade IETF (todos nós)
 - Area Directors (ADs)
 - Gestão de cada área (2 por área)
 - Internet Engineering Steering Group (IESG)
 - Cada director de área e também o chair do IETF
 - Internet Architecture Board
 - IETF chair also
 - IETF Chair, ADs & IAB members
 - Posições de 2 anos, nomeações
- **Todas as posições são voluntárias!**



Padronização - IETF

- *“We reject kings, presidents and voting. We believe in rough consensus and running code” (Dave Clark 1992)*
- **Não se trata de uma associação!**
 - Qualquer indivíduo pode fazer parte – basta participar
- **Desenvolve reuniões anuais (3 por ano)**
 - 1000 a 2000 pessoas em diferentes grupos de trabalho
 - Mais em mailing-lists
- **Desenvolve standards**
 - Qualquer pessoa pode participar

IETF, desenvolver um Standard

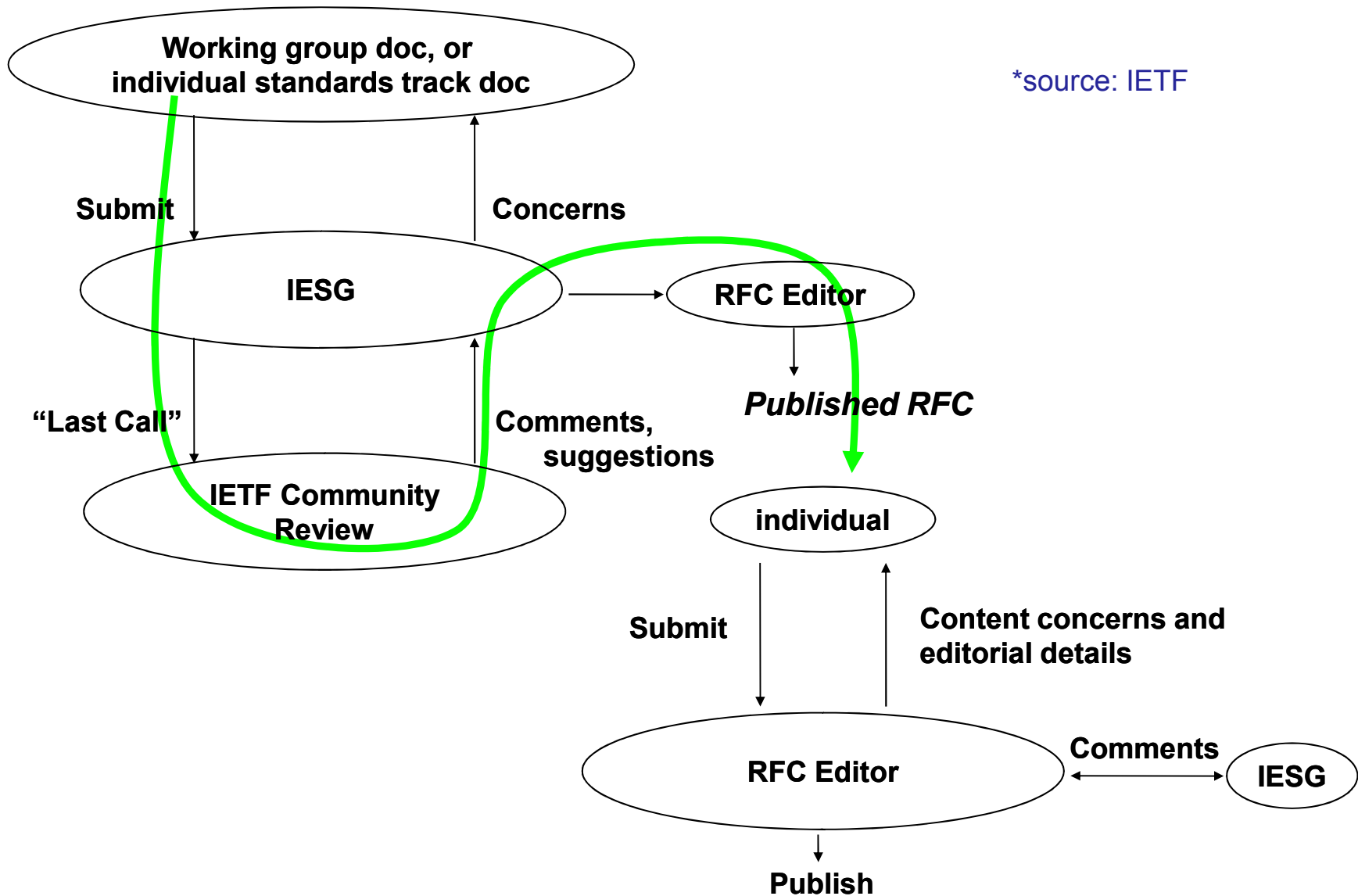
- **Proposta por uma pessoa ou grupo numa área (ou fora dela)**
 - Começa por ser um Internet draft (ID) – documento em desenvolvimento
 - Desenvolvido (comentado pelo grupo)
 - Necessário convencer área de que se trata de uma ideia relevante...
 - Mediante comentários, longo processo de revisão
 - Drafts vão alterando versão
 - Se não houver interesse, draft desaparece
 - Se houver interesse, desenvolve-se até atingir maturidade
 - Request for comments – mediante IESG
 - Documento para arquivar...
 - Sem draft, não há RFC...
 - RFC 1 Host Software - Apr 7 1969
 - Actualmente, mais de 5000 RFCs...
- Standards
 - Descritos em inglês
 - Escritos em ASCII
 - Sem figuras! (só esquemas ASCII)



Padronização - IETF

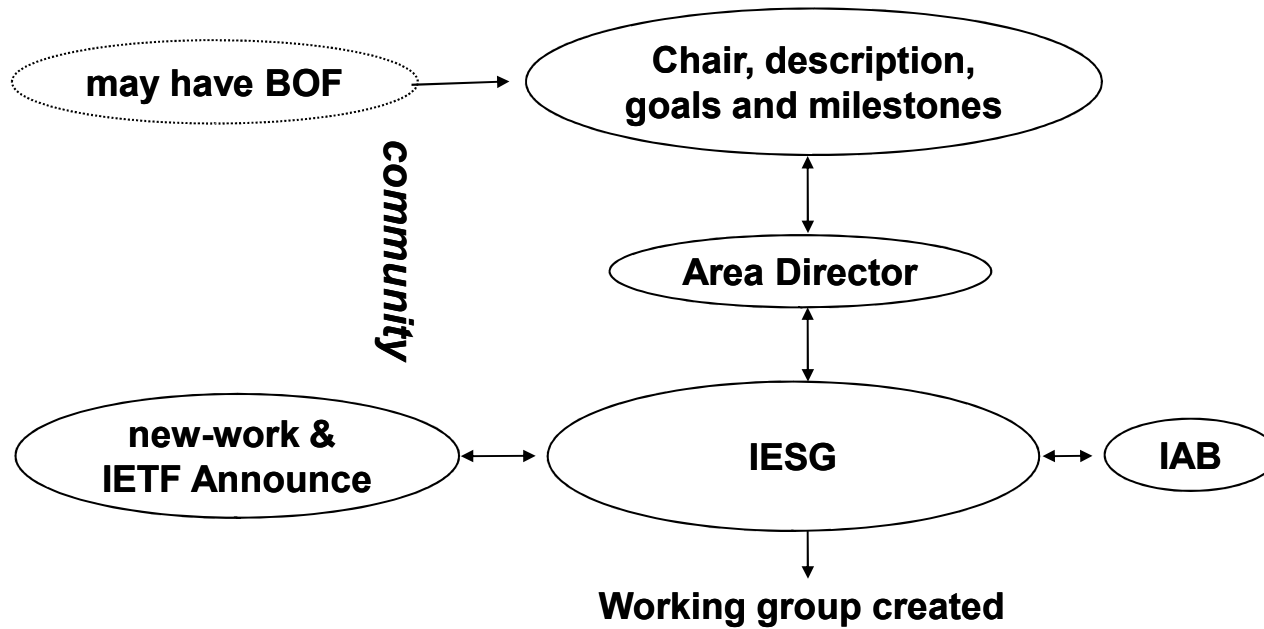
- **Quem define um standard**
 - TODAS as pessoas envolvidas em determinado grupo de trabalho
 - Não há qualquer entidade reguladora – movimento voluntário
- **Em que áreas? “above the wire and below the application”**
 - IP, TCP, email, routing, IPsec, HTTP, FTP, ssh, LDAP,
 - SIP, mobile IP, ppp, RADIUS, Kerberos, secure email,
 - streaming video & audio, ...
 - Mas...MPLS, etc, são também definidos no IETF
 - Acima do “wire” – noção de canal alterada desde o início da Internet...
- **Como participar?**
 - Verificar grupos de trabalho – IETF Working Group
 - Subscrever a mailing-lists
 - Ouvir, ler standards, comentar
 - Participar nos encontros do IETF

IETF, desenvolver um Standard



*source: IETF

IETF, propôr um Working Group*



*source: IETF

IETF e Intellectual Property Rights

- **Se uma entidade/pessoa propuser tecnologia já patenteada ou em submissão...**
 - Standards são abertos
 - Pode-se (RFC 2026)
 - Reter direitos de implementação
 - Reter direitos de licenciamento “fair and non-discriminatory”
- Existe actualmente receio sobre patenteamento não declarado, ou omissão
 - RFC 3978 e RFC 3979
 - Indica ser necessário apresentação IPR na submissão de standard
 - RFC 3669 fornece algumas directivas
 - Alguns WGs não querem patentes – assegurar que standards estão livres de IPRs
- Autores fornecem direito publicação não-exclusivo ao IETF
 - Também direitos com trabalho derivado
 - Retêm todos os outros direitos
 - Draft inclui afirmação que inclui:
 - Indicação de acordo para publicação de IPR
 - Opcionalmente, pode indicar que não pode haver seguimento (avançar para RFC)
 - Copyright

Padronização - IRTF

- **IETF: padrões, mas soluções concretas**
- IRTF: INVESTIGAÇÃO – problemas relacionados com Internet, visão a longo-prazo
- **Exemplos:**
 - **Anti-Spam Research Group (ASRG)**
 - **Crypto Forum Research Group**
 - **Delay-Tolerant Networking Research Group (DTNRG)**
 - **End-to-End Research Group**
 - **Host Identity Protocol Research Group (HIP)**
 - **IP Mobility Optimizations (Mob Opts) Research Group**
 - **Network Management Research Group Charter (NMRG)**
- **Modelos de Standards - IETF**

Referências

- Bakkalbasi, Nisa, et al. (2006) Three options for citation tracking: Google Scholar, Scopus and Web of Science. Biomedical Digital Libraries. DOI: 10.1186/1742-5581-3-7
- Harzing, Anne-Wil. (2010) The Publish or Perish Book: your guide to effective and responsible citation analysis. Tarma Software Research Pty Ltd, Melbourne, Australia. 246 pp. [U of I Library call number: 025.324 H269p]
- Monaskersky, Richard. (2005) The number that's devouring science. Chron High Educ. tinyurl.com/76ernc7
- Wikipedia: <http://en.wikipedia.org/wiki/H-index> http://en.wikipedia.org/wiki/Impact_factor
- Williamson, James R. (2009) My h-index Turns 40: My Midlife Crisis of Impact. ACS Chemical Biology. DOI: 10.1021/cb9001014

- Website de Kevin Almeroth <http://www.cs.ucsb.edu/~almeroth/conf/stats/>
- IETF: <http://www.ietf.org>
- IRTF: <http://www.irtf.org>



Informatics Systems and Technologies Research Unit
Unidade de I&D em Sistemas e Tecnologias Informáticas



UNIVERSIDADE LUSÓFONA
de Humanidades e Tecnologias
Humani nihil alienum