Name: Prof Mike Thelwall   Email address: m.thelwall @ wlv.ac.uk   Date of portfolio: 19 January 2014

Part 1: Narrative and academic age calculation

Part 1a: Narrative

I am a professor at the University of Wolverhampton in the UK on a contract that is 45% teaching and 55% research and research-related administration. My main achievement is founding and leading the Statistical Cybermetrics Research Group (SCRG) at the University of Wolverhampton in 2000. Before the foundation of this group there had been no library and information science research ever at the University of Wolverhampton, so this group started with no researchers and no tradition of research in this field. By the end of 2007, the official UK national Research Assessment Exercise (the most recent) rated three quarters of the research produced by the group as "World leading" in the area, which was the joint highest for any UK group in the library and information management category. Since its formation, the group has made substantial contributions to research in webometrics, altmetrics, scientometrics and sentiment analysis, has attracted about £1.5million in research funding, has graduated 7 PhD students and has produced over 200 articles in international (ISI-indexed) journals, over a quarter of which are in the top journal of the discipline.

My own research has two related goals: (a) the development of methods for extracting and analysing web data to solve research problems from information science and (b) the development of methods for extracting and analysing web data as a goal in itself, for use in the wider social sciences. These goals have been realised in three books, hundreds of journal articles and conference papers and four computer programs distributed free on the web. These four programs are: the sentiment strength detection software SentiStrength; the altmetrics and webometrics data gatherer and analyser Webometric Analyst; the Twitter time series analyst Mozdeh; and the web crawler and hyperlink analyser SocSciBot. The main evidence in support of the value of this research is my Google Scholar h-index of 51.

Although SentiStrength is free to researchers I also sell it commercially and its biggest user is Yahoo! for its question answering service. SentiStrength is also used in major media events to harness the sentiment in tweets from the public for light shows, including on the London Eye throughout the London Olympics.

My teaching experience has mainly included undergraduate mathematics, statistics and operational research, and postgraduate statistics and research methods. In addition, I have developed undergraduate and postgraduate cybermetrics courses, which I believe to be the first of their kind in the world. For example the MSc Introduction to Cybermetrics course combines mathematical, information science and computing theoretical foundations with a practical component that involves using webometric software for practical tasks. This course has been attended not only by MSc students but by PhD students, postdoctoral researchers and lecturers from around Europe. Versions of this course have also been delivered by me in two other countries

and I have also given courses on sentiment analysis to PhD students at summer schools in Spain, Russia and China.

Part 1b: Academic age calculation	Information [delete this and/or replace with extra academic age explanations]
Start date of PhD: September 1986 *Date of PhD defence: July 1989* *Number of children raised after PhD defence: 0* *Special allowances (describe below): 0* ***Academic Age*** *= Number of full-time years worked (count % of full time for part-time years) since PhD defence – Number of children raised – special allowances =* ***24.5 years*** *(min. 1 year)*	The academic age calculation helps the evaluator to mentally adjust their expectations based upon someone's academic age. The minimum permitted Academic Age is 1 in all cases, even for those without a PhD.

Part 2: Expertise sub-portfolio

Expertise	Sub-factor	Claim and evidence [delete the help text in most cases and replace it with your text]
Scientific/ scholarly expertise	Theoretical	I do not conduct theory-driven research but use a wide range of theories to drive research questions and help explain the results. I have helped to develop two theoretical contributions, one for link analysis and one for big data "information-centred" research, as cited below. Thelwall, M. (2006). Interpreting social science link analysis research: A theoretical framework. Journal of the American Society for Information Science and Technology, 57(1), 60-68.

		Thelwall, M., Wouters, P., & Fry, J. (2008). Information-centred research for large-scale analysis of new information sources, Journal of the American Society for Information Science and Technology, 59(9), 1523-1527.
	Subject	Link analysis. My early contributions are summarised in the following book but numerous subsequent journal articles have updated this. Thelwall, M. (2004). Link analysis: An information science approach. San Diego: Academic Press. Webometrics My main contributions are summarised in the following book but numerous subsequent journal articles have updated this. Thelwall, M. (2009). Introduction to webometrics: Quantitative web research for the social sciences. San Rafael, CA: Morgan & Claypool Web research methods for social scientists. These are summarised in the following book. Thelwall, M. (2013). Webometrics and social web research methods [free online in-progress draft copy]. University of Wolverhampton. Altmetrics. Methods to evaluate metrics derived from the social web for impact assessment and related purposes. The following article is an example of this. Thelwall, M., Haustein, S., Larivière, V. & Sugimoto, C. (2013). Do altmetrics work? Twitter and ten other candidates. PLOS ONE, 8(5): e64841. doi:10.1371/journal.pone.0064841
	Methodological	Because I research methods development, this section is the same as the one above.
	Originality / independence	I value originality perhaps above everything else in research. I demonstrate originality in my research by attempting to get it published in competitive journals that use originality as one of their criteria for inclusion. Almost all of my publications are in outlets that value originality. As an example, the paper below was original for (a) the first dual classification system for positive and negative sentiment, (b) the

		first sentiment analysis program with explicitly social science goals, and (c) the first sentiment analysis program to draw upon social network sites for its raw data. Thelwall, M., Buckley, K., & Paltoglou, G. (2012). Sentiment strength detection for the social Web. Journal of the American Society for Information Science and Technology, 63(1), 163-173.
Knowledge transfer	Reviewing	Provide the total number of (a) conference papers and (b) journal articles you have reviewed and (c) give examples of journals or conferences for which you reviewed (list top 3). 88 conference papers reviewed 690 journal articles reviewed Journal of the American Society for Information Science and Technology Journal of Informetrics Scientometrics
	Entrepreneurship	List entrepreneurship activities undertaken, such as launch or participation in spin-offs, and joint projects with industry, NGOs or government (list top 3). I was contracted to provided twice-yearly webometric reports for the UK NGO NESTA for five years.
Educational expertise	Courses taught or developed	(a) Approximate number of hours spent lecturing to a class (not including preparation or 1-1 supervision) 9000 – currently 250 per year. (b) Description of the types of courses prepared and taught *excluding* online courses and MOOCs (these are listed as outputs). This might be a list of courses prepared or an overall description of the topics and levels of the courses. List only those that are most important to you, (list top 3). MSc Introduction to Webometrics. Developed from scratch by me and taught by me. Gives an introduction to a range of webometric techniques, including mathematical, theoretical and computing

		foundations. The first of its kind in the world.
		MSc Advanced Statistics. Jointly developed from scratch by me (50%) mainly with new topics not taught at the University before, such as structural equation modelling.
		Final year undergraduate topics in pure maths: Developed from existing notes from a previous lecturer. I love teaching this course and had to learn all of the topics in it (knot theory, continued fractions, combinatorics) in order to teach it for the first time.
	Other educational expertise	I am able to give courses to PhD students on current research topics. For example, I have given sentiment analysis courses to three PhD summer schools, in China, Russia, Germany and Spain and have given lecture courses to information science PhD students on webometrics in Spain and the USA.
Technological expertise	Methods	I use advanced statistical methods, with the most advanced probably being zero inflated negative binomial regression, as used in the article below. Thelwall, M. & Maflahi, N. (in press). Are scholarly articles disproportionately read in their own country? An analysis of Mendeley readers. Journal of the Association for Information Science and Technology. Other than statistics, my research mainly develops methods rather than applying existing methods.
	Tools + lab equipment	Nothing except lots of computers.
	Software	I use statistical software such as R in the following paper Thelwall, M. & Maflahi, N. (in press). Are scholarly articles disproportionately read in their own country? An analysis of Mendeley readers. Journal of the Association for Information Science and Technology. And SPSS in the following:

		Thelwall, M. (2012). Journal impact evaluation: A webometric perspective, Scientometrics, 92(2), 429-441
		And I wrote my own statistical programs to analyse the data in the following
		Thelwall, M., Haustein, S., Larivière, V. & Sugimoto, C. (2013). Do altmetrics work? Twitter and ten other candidates. PLOS ONE, 8(5): e64841. doi:10.1371/journal.pone.0064841
		I also use social network analysis software such as UCINET and Pajek but prefer my own similar software, within Webometric Analyst (http://linkanalysis.wlv.ac.uk)
	Data management or data curation	I archive my datasets and put them free online, when possible and practical. I have done this for 14 years at http://cybermetrics.wlv.ac.uk/datasets/ but now use figshare for longer term stability.
Communication expertise	Languages	Native English speaker. English shouter (foreigners understand it). Very poor French, Dutch and Spanish. Abysmal spoken Urdu/Punjabi (greetings, very common words, swearing).
	Presentations	Keynote/invited talks at institutions other than your own or at conferences (list top 3): CICLing 2013 conference keynote talk on sentiment analysis, Greece. The Nordic Library conference keynote talk on webometrics, Turku. NLDB 2011 conference keynote talk on sentiment analysis, Spain.
	Writing	Awards for papers (list top 3): Kousha, K. & Thelwall, M. (2007). Google Scholar citations and Google Web/URL citations: A multi-discipline exploratory analysis, Journal of the American Society for Information Science and Technology, 57(6), 1055-1065. [Highly cited paper award in social sciences and art and humanities from Iran's Ministry of Science and Technology, 2008] Thelwall, M. (2004). Can the web give useful information about commercial uses of scientific research?

		Online Information Review, 28(2), 120-130. [Emerald 'Outstanding Paper' award] Thelwall, M. (2007). Blog searching: The first general-purpose source of retrospective public opinion in the social sciences? Online Information Review, 31(3), 277-289. [publisher version] ['Highly Commended' award from the publisher Emerald]
	Public engagement (media interview and other)	Examples of video or audio media interviews (list top 3): Czech TV News 24 live interview on sentiment analysis and hacking in the Girl with the Dragon Tattoo trilogy. Radio 4 recorded interview on swearing in social network sites. Radio 4 recorded interview on the mathematics of search engines.
Organisational expertise	Management	Description of management roles undertaken (list top 3): Head of the Statistical Cybermetrics Research Group (2000-) Chair of the Student Management Board of the Research Institute for Information and Language Processing (2011-2013). Coordinator of Wolverhampton's contribution to numerous external funding projects.
	Advising	Visits to other institutions (universities or other) and the type of advice given (list top 3):
	Project leadership	Projects and teams led (list top 3). Include the size of the team, and whether this is an international collaboration: Led Wolverhampton's contribution to European projects NETREACT, RINDICATE and WISER

	Collaboration	Projects and teams involved in but not led (list top 3). Include the size of the team, whether this is an international collaboration, and the portfolio owner's position in the team: EU project WISER: 15 members; leader of Wolverhampton's component, international collaboration EU project CyberEmotions: 20 members; leader of Wolverhampton's component, international collaboration EU project ACUMEN: 15 members; leader of Wolverhampton's component, international collaboration
	Administration and committee work	Administrative roles undertaken, including committee membership, chair or secretary roles, organising workshops or conferences, organising online discussions (list top 3): Wolverhampton professoriate member. Co-organised free ASIST European Chapter workshop in Wolverhampton Organised free Webometrics workshop in Wolverhampton.
Other		Other types of relevant expertise not covered above (list top 3, explaining each one): Programmer (Java, dot net, R, lots of others)

Part 3: Output sub-portfolio

Output	Sub-factor	Claim and evidence
Scholarly outputs	Books	Number of scholarly books or theses published (exclude self-published): _2_ List of books published (list top 3): Thelwall, M. (2009). Introduction to webometrics: Quantitative web research for the social sciences. San Rafael, CA: Morgan & Claypool (Synthesis Lectures on Information Concepts, Retrieval, and Services, 2009,

		Vol. 1, No. 1). Webometrics book blog. and Webometrics book web site. Also available in Persian, translated by M. Hassanzadeh, M. Hosseini and F. Navidi. Thelwall, M. (2004). Link analysis: An information science approach. San Diego: Academic Press. Thelwall, M. (2013). Webometrics and social web research methods [free in-progress draft copy]. University of Wolverhampton. [**self-published**]
	Book chapters	Number of book chapters published: _22__ List of book chapters published (list top 3): Thelwall, M. (2013). Society on the Web. In: Dutton, W., (ed.), The Oxford Handbook of Internet Studies. Oxford, UK: Oxford University Press (pp. 69-85). Thelwall, M. (2011). Privacy and gender in the Social Web. In: Sabine Trepte, Leonard Reinecke (Eds), Privacy online: Perspectives on Privacy and Self-Disclosure in the Social Web, New York: Springer (pp. 255-269). Thelwall, M. (2011). Investigating human communication and language from traces left on the web. In: Malcolm Williams, W Paul Vogt, (Eds), The SAGE Handbook of Innovation in Social Research Methods, London: Sage. (pp. 167-181).
	Reviews	Number of book reviews published: 3
	Editorials	Number of editorials published: 1
	Journal articles	Number of refereed journal articles or fully refereed complete conference papers published: 244 List of refereed journal articles or fully refereed full conference papers published (list top 3) 1. Thelwall, M. & Maflahi, N. (in press). Are scholarly articles disproportionately read in their own country? An analysis of Mendeley readers. *Journal of the American Society for Information Science and Technology*. 2. Thelwall, M. & Kousha, K. (in press). ResearchGate: Disseminating, communicating and measuring scholarship? *Journal of the American Society for Information Science and Technology*. 3. Thelwall, M., Haustein, S., Larivière, V. & Sugimoto, C. (2013). Do altmetrics work? Twitter and ten other candidates. *PLOS ONE*, 8(5), e64841. doi:10.1371/journal.pone.0064841

	Conference papers	Number of conference abstracts, panel discussions or posters published: 25 Ignore unpublished conference papers.
Communication to the general public	Press stories	Number of magazine or newspaper articles published (written by you, not about you): 0 List of magazine or newspaper articles published (by you, not about you) (list top 3) :
	Encyclopedia articles	Number of encyclopedia articles published (excludes Wikipedia and similar): 0 List of encyclopedia articles published (list top 3) Can give examples of contributions made to Wikipedia as part of these:
	Popular books / articles	Number of popular books or articles published: 0 List of popular books or articles published (list top 3).
Teaching	Books	Number of textbooks published (exclude self-published): 0 List of textbooks published (list top 3):
	Online courses	List of online courses created (includes MOOCs), including creation date, type of materials generated and number of students per year (list top 3): Introduction to the Internet (1996) Introduction to Visual Basic (1996) Project management (1996)
	Students completed	Undergraduate students supervised as main supervisor that have now graduated: 8 students Master's students supervised as main supervisor that have now graduated: 2 students PhD students supervised as main supervisor that now have their doctorate: 7 students
Web and social	Online presence	List your online presence: accounts in social media used for academic purposes, academic network

media academic communication		accounts, digital repository accounts, websites that you created or use to create output. If it applies, mention per site how active you are in posts per year or month (e.g., twitter, blogs, ResearchGate, SlideShare) (list top 3 https://twitter.com/mikethelwall (2 posts per year) https://wlv.academia.edu/MikeThelwall https://www.researchgate.net/profile/Mike_Thelwall
	Online contributions	Give examples of other online contributions to scholarly discussions that you made here. Do not repeat information given above (list top 3):
Datasets, software, tools, instruments	Datasets	Number of datasets published: 20+ depending how you count Brief description of datasets published (list top 3): Sets of the hyperlinks between the universities of the UK, Australia, China, New Zealand for several years Six human-classified sentiment analysis data sets
	Software, tools, instruments	Number of software, tools and instruments developed: 4 Brief description of software, tools and instruments developed (list top 3): the sentiment strength detection software SentiStrength; the altmetrics and webometrics data gatherer and analyser Webometric Analyst; the Twitter time series analyst Mozdeh
Registered intellectual or industrial rights	Patents	Number of patents, standards, guidelines published: _____ Brief description of patents, standards, guidelines published (list top 3):
	Discoveries	Number of registered discoveries, such as animal species, celestial bodies, DNA sequences, algorithms: _____ Brief description of registered discoveries (list top 3):
Funding & Grants	Funding	Number of projects funded: 22 Total grant funding received (do not count funding allocated to other universities, and if there were multiple applicants in your own university, divide the funding by the number of applicants): £1.1 million Brief description of funded projects (list top 3): **CyberEmotions** - This large FP7 EU project on sentiment analysis in social media involved nine European

		organisations and with a multi-million Euro budget. **CREEN** - EU FP6 project developed the physics of complexity science, and developed web tools to track, model and predict science debates on the web (2005-8). **Rindicate** - "The use of webometrics for the analysis of knowledge flows within the European Research Area". In relation to (DG-RTD-2005-M-02-01): "Multiple Framework Service Contract for Expert Support with the Production and Analysis of R&D Policy Indicators" together with IDEA CONSULT (coordinator), NIFU STEP, and SPRU. The project is set to be published in the EU research information portal.
Other		Other types of relevant output not covered above (list top 3, explaining each one):

Part 4: Influence sub-portfolio

Influence	Sub-factor	Claim and evidence [delete the help text and replace it with your text]
Influence on science	Total and average citations	Total citations received to all publications, as listed in Google Scholar: 9,487 and average number of citations per paper: 31 Total citations received to all publications, as listed in Web of Science or Scopus: 2,667 and average number of citations per paper: 12.6
	Article citations	Total citations to one of your top 3 articles, as listed in Google Scholar: 220 and as listed in Web of Science: 101. Article title: Extracting macroscopic information from web links Publication year: 2001 Authors (in order): Thelwall, M. Total citations to another of your top 3 articles, as listed in Google Scholar: 218 and as listed in Web of Science: 56. Article title: Social networks, gender, and friending: An analysis of MySpace member profiles Publication year: 2008 Authors (in order): Thelwall, M. Total citations to another of your top 3 articles, as listed in Google Scholar: 188 and as listed in Web of

		Science: 35. Article title: Sentiment strength detection in short informal text Publication year: 2010 Authors (in order): M Thelwall, K Buckley, G Paltoglou, D Cai, A Kappas
	h-index	h-index, as listed in Google Scholar: 51 h-index, as listed in Web of Science: 27 [The h-index is the largest number h such that at least h articles have received at least h citations.]
	Book citations	Total citations to one of your top 3 books, as listed in Google Books: 15; Book name: Introduction to webometrics: Quantitative web research for the social sciences Publication year: 2009 Authors (in order): Thelwall, M. Total citations to another of your top 3 books, as listed in Google Books: 29; Book name: Link Analysis, An information science approach. Publication year: 2004 Authors (in order): Thelwall, M.
	Age-corrected h-index	m-quotient for Google Scholar: 2.01 m-quotient for Web of Science: 1.10 [For this portfolio, the m-quotient is the h-index divided by academic age.]

	Multi-authorship compensation	To compensate for multi-authorship, either report (a) or (b) below, which the evaluator will take into account when assessing your citations.
		
		  a) Average number of authors on your publications listed in Google Scholar (including yourself):2
		  b) Proportion of publications listed in Google Scholar for which you were the first author: _____
		
		If monographs, Web of Science publications or Scopus publications are more important to you than Google Scholar publications, you can report for these instead, but please state it clearly.

	Scholarly prizes	Scholarly prizes and awards received (local, national and international) (list top 3):
	Editing and reviewing	Your main reviewer, editor or editorial board member tasks (list top 3): Journal of the Association for Information Science and Technology (assoc. editor) Journal of Informetrics (board member) Scientometrics (board member)
	Committees	Your main conference/program committee memberships (list top 3): ISSI, STI, NLDB
	Online discussions - social web followers	Number of followers, if substantial, in your web presences (e.g., Academia, Blogs, Twitter) (list top 3): Social website name Twitter Number of followers: ___ 271 Social website name Academia Number of followers: ___ 103 Social website name ResearchGate Number of followers: ___ 129 Also, report up to 3 interesting web mentions of you or your work that are not already elsewhere in the portfolio:
	Downloads	Article name Social networks, gender, and friending: An analysis of MySpace member profiles Number of downloads: 3rd Most Accessed JASIST Article of 2012: (no numbers given) Article name Sentiment in Twitter events Number of downloads: 4th Most Accessed JASIST Article of 2012: (no numbers given) Article name Do altmetrics work? Twitter and ten other candidates. Number of downloads: 7330 (top 3

		downloaded only)
	Mendeley readers	Article name Sentiment strength detection in short informal text Number of Mendeley readers: 136 Article name Sentiment strength detection for the social web Number of Mendeley readers: 91 Article name Data Mining Emotion in Social Network Communication : Gender differences in MySpace Number of Mendeley readers: 142 To count Mendeley readers, go to Mendeley.com and search for each publication, recording how many readers it has (list top 3):
	Invited talks	Number of invited keynote talks at conferences outside your country: 5 Number of invited keynote talks at conferences inside your country: 0 Number of invited talks at universities outside your country: 15 Number of invited talks at other universities inside your country: 10 List of invited talks of all kinds [include name and venue] (list top 3): Yandex, Moscow: Sentiment analysis talk. Oxford Internet Institute: Webometrics guest lecture to MA students every year University Complutense Madrid: Lectures to PhD students and undergraduates
Influence on society	General public	Number of magazine or newspaper articles published (written about your research, not by you): 15 Examples of magazine or newspaper articles published (about your research, not by you) (list top 3): - Time Magazine: Want to Light Up the London Eye? Just Tweet That the Olympics Are 'Totes Amazeballs', July 27, 2012. - UK Daily Telegraph article, p. 27, 19 July 2012, "Happy Olympic tweeters to light up London Eye" in "the world's first social media driven light show". - BBC News Article: 20 July 2012, London Eye Olympic Twitter positivity lightshow launched. Examples of web pages published (about your research, not by you) (list top 3, including title and who

		wrote them):
	Tweets or blog posts about publications.	Article name Do Altmetrics Work? Twitter and Ten Other Social Web Services Number of Tweets of it: 104 (one article only) Tweets can only be monitored in real time but can report them if they are reported in the publisher website or by the Altmetric Bookmarklet, available free at: http://www.altmetric.com/bookmarklet.php. Article name Do Altmetrics Work? Twitter and Ten Other Social Web Services Number of Blog posts about: 6 (one article only)
	Advice	Number of times asked for specialist evidence outside academic, economic and educational contexts, including membership of non-academic, non-educational committees: 0 Examples of giving specialist evidence outside academic, economic and educational contests, including committee memberships (list top 3):
	Professional practice	Examples of professional practice using your subject expertise (e.g., working as a lawyer, nurse) (list top 3):
	Laws, regulations, guidelines	Laws, regulations, guidelines and so forth that have been initiated, developed or amended, at least partly based on your research. Briefly explain how and refer to projects, papers and other evidence of this influence (list top 3):
Influence on economy	Income	Total 3rd stream income (money generated for commercial activities): £24,000
	Consultancies	Number of consultancy or advisory positions for companies: 0

	Citations from patents	Number of citations to your work from patents: 0
		Names of patents citing your work (list top 3):
		[Citations from patents, if any, may be listed in the Google Scholar citations to a paper.]
	Citations *to* patents	Number of citations to your patents (if any) from scholarly documents: 0
		[Citations to your patents, if any, can be found by searching Google Scholar for the patent.]
	Spin-offs	Number of spin off companies created: 0
Influence on teaching	Awards	Teaching awards, including both within and outside the host institution (list top 3): 0
	Online views	Number of views of your top 3 SlideShare or YouTube presentations, if substantial.
		Presentation URL: _____ views: _____
		Presentation URL: _____ views: _____
		Presentation URL: _____ views: _____
		[Could also report any similar view counts for other sites, such as Vimeo, or online learning environments.]
	Syllabus mentions:	Number of online syllabuses or course notes pages listing the academic's works (list top 3).
		Publication: Link analysis, an information science approach Syllabuses mentioning: 3
		Publication: Introduction to Webometrics Syllabuses mentioning: 2

	Textbook sales	Total sales of your textbooks: _____ copies. Amazon Bestsellers Rank: 1,162,256 in Books [Can also report Amazon sales ranks in comparison to similar books instead, if sales figures unavailable (list top 3). To do this, search for your book by title or ISBN in Amazon.com and find "Amazon Best Sellers Rank" in the Product Details section and report this number.]
	Invited lectures	Number of invited lectures to undergraduates at other universities: 9
	Dataset or software downloads	Number of downloads of datasets or applications created by the portfolio owner (list top 3). Name of software/dataset: SentiStrength Number of downloads 950 Name of software/dataset: SocSciBot Number of downloads 1023 Name of software/dataset: Webometric Analyst Number of downloads 3254 [can also report citations to the software or datasets from Google Scholar, if any]
Other		Other types of relevant influence not covered above (list top 3, explaining each one): The official UK national Research Assessment Exercise 2008 (the most recent) rated three quarters of the research produced by the group as "World leading" in the area, which was the joint highest for any UK group in the library and information management category.