

教育資料與圖書館學

*Journal of Educational Media & Library Sciences*

<http://joemls.tku.edu.tw>

---

Vol. 54 , no. 1 (2017) : 111-129

他山之石，能否攻錯？談「舊金山  
宣言」與「萊登宣言」所帶來的啓示

What We Have Learned from  
San Francisco Declaration on  
Research Assessment and Leiden Manifesto?

陳 明 俐\* Carey Ming-Li Chen\*

Assistant Researcher

Doctoral student

E-mail : [carey.mlchen@gmail.com](mailto:carey.mlchen@gmail.com)

林 雯 瑤 Wen-Yau Cathy Lin

Associate Professor

[English Abstract & Summary see link](#)

[at the end of this article](#)



# 他山之石，能否攻錯？談「舊金山宣言」與「萊登宣言」所帶來的啟示

陳明俐<sup>ab\*</sup> 林雯瑤<sup>c</sup>

## 摘要

近年在政府或機構層級對於學術社群成員的研究表現評估多以各種指標配合同儕審查的方式進行，但對於研究評估指標的設計與應用均有諸多爭議。本文介紹舊金山宣言與萊登宣言的發展過程，並概述其準則內容，最終並比較兩者的差異，希望透過本文能喚起台灣學術社群對於各種與研究評估指標相關爭議的重視與省思，凝聚使用研究評估指標的共識，進而對未來發展具在地特質的宣言或有所助益。

**關鍵詞：**舊金山宣言，萊登宣言，資訊計量學，研究評估指標

## 前 言

2015年4月，*Nature*期刊刊登了一篇名為*Bibliometrics: The Leiden Manifesto for research metrics*的文章，多位來自資訊計量學領域的學者闡述了現今研究評估指標濫用、指標分析與評鑑評比結果互相掣肘、過分重視被引用次數相關指標等現象，並提出十項學術研究評估的準則(Hicks, Wouters, Waltman, de Rijcke, & Rafols, 2015)。該份文件隨即引發各界重視，在近幾年所舉行的資訊計量學領域重要研討會的主題討論場合，亦有對於該議題的後續討論。

該篇文章又可簡稱為「萊登宣言」(Leiden Manifesto)，最初的內容源自2014年於荷蘭萊登大學科學與技術研究中心(Centre for Science and Technology Studies, CWTS)所舉辦的19<sup>th</sup> International Conference on Science and Technology Indicators (STI)研討會上的專題討論場次*Quality standards for evaluation indicators: Any chance for the dream to come true?* 在該場討論中，多位資訊計量學、科技政策領域的知名學者共同探討評估計量的指標適切性，以及在現行制度上實際運用時所遭遇的問題。隨後來自美國喬治亞理工學院中公共政策學院的

<sup>a</sup> 財團法人國家實驗研究院科技政策研究與資訊中心助理研究員

<sup>b</sup> 國立台灣大學商學研究所博士生

<sup>c</sup> 淡江大學資訊與圖書館學系副教授

\* 本文主要作者兼通訊作者：carey.mlchen@gmail.com

Diana Hicks教授在會議上發表了七項準則(Halevi, 2014)，其內容立即引發在場人士的熱烈討論。最後在隔年(2015年)，由Diana Hicks、CWTS主任Paul Wouters、Ludo Waltman研究員、Sarah de Rijcke助理教授，以及來自西班牙國家研究委員會的Ismael Rafols共同具名發表萊登宣言，並刊載於*Nature*期刊。

事實上，這已經不是第一次由學術社群提出有關研究評估指標的宣言。2012年12月，在一場由The American Society for Cell Biology(ASCB)所舉辦的年度會議上，一群學術期刊編輯與出版商代表已共同發表對於研究評估的建言，名為San Francisco Declaration on Research Assessment(DORA)，又稱「舊金山宣言」。該份文件提出18項研究評估準則，且分為整體性準則，以及針對不同利害關係者所給予的特定性準則。

本文將從兩個面向梳理，首先彙整兩項宣言的發展過程，以及概述其準則內容，最後再比較兩者間的差異。希望透過本文的分析，能喚起台灣學術社群對使用研究評估指標相關爭議的重視與省思，對於凝聚關於使用研究評估指標的共識，進而發展「在地版」的萊登宣言或有所助益。

## 二、「舊金山宣言」的起源與內容

舊金山宣言乃2012年12月16日在ASCB年會中倡議，當時學術社群中有155位以個人身分簽署，另有82個單位亦以機構名義簽署，顯示各單位對於研究評估的重視。隨著外界充斥著對評估指標的不滿聲浪，以及體認到濫用評估指標恐將對學術研究發展產生不良影響，使得越來越多的單位開始重視研究評估的真諦為何，並認同舊金山宣言所提及之準則。截至2017年3月，已有840個機構簽署舊金山宣言，以個人身分簽署的更是大幅成長，已有超過12,550位來自全球各地的研究人員或學術社群的利害關係者，表達對於這份文件的認同。檢視舊金山宣言的起始簽署者其國家與領域的分布情形，2013年時，來自歐洲的占比最高，共有46.8%，其次為北美洲與中美洲，占比為36.8%，南美洲為8.9%，亞洲則僅有5.1%。在領域分布上，多數簽署者來自科學領域，人文領域者僅有6%。

舊金山宣言共包含18項準則，分別針對整體及不同利害關係者提出各自所需注意的事項，包含了整體建議、對補助機構、研究機構、出版商、提供計量指標機構，以及研究者本身，然而該份文件並未對各項準則提出進一步完整的說明文字。以下為18項準則的內容節錄。<sup>1</sup>

### (一) 整體建議

1. 勿使用以期刊為基礎的計量指標(如期刊影響係數)作為判斷單一論文的

<sup>1</sup> 舊金山宣言內容：<http://www.ascb.org/files/SFDeclarationFINAL.pdf?x30490>

品質、衡量個人層級的研究表現之依據，或是影響招募、升遷、補助的決定過程。

### (二) 對補助機構的建議

1. 對於新進研究人員的表現衡量，研究論文的內容遠比發表指標或其所出版的期刊更為重要。
2. 在進行研究評估時，應考量其所有研究產出的價值與影響，甚至更應檢視其研究是否對政策或實務上造成影響。

### (三) 對研究機構的建議

1. 明確闡述在招聘、終身職、升遷的決策過程所使用標準為何，特別是針對新進研究人員的評估，研究論文的內容遠重要於該篇論文的計量指標結果。
2. 在進行研究評估時，應考量其所有研究產出的價值與影響，甚至更應檢視其研究是否對政策或實務上造成影響。

### (四) 對出版商的建議

1. 應大幅降低使用期刊影響係數作為促銷期刊的工具，理想上，最好能避免談及影響係數，或改以其他相關指標（五年的期刊影響係數、EigenFactor、SCImago、*h*-index、編輯與發行頻率等）作為期刊績效的說明，以提供更完整的面貌。
2. 提供一系列以論文為基礎的衡量指標，以鼓勵各方將評估的重點轉向於論文本身的內容。
3. 鼓勵負責的作者行為，並提供每位作者貢獻情形的訊息。
4. 無論期刊為開放取用或採訂閱模式，均應移除論文中對於參考文獻列表的再利用限制，使其以創用CC的授權方式呈現。
5. 移除或降低研究型論文對於參考文獻數目的限制，或適當要求作者應用初始發現研究結果的論文。

### (五) 對提供計量指標組織的建議

1. 計算指標的數據與方法皆須公開透明。
2. 若可能的話，在經授權的狀況下，提供可不受限制、重新使用的數據，並提供數據的計算過程。
3. 表明無法接受對於計量指標的不當操弄，並明確說明哪些行為屬於不當操弄，且該採用哪些量測方式予以反擊。
4. 需考量不同文獻類型或不同領域在計量指標使用、聚合、比較時會產生的差異。

## (六) 對研究者的建議

1. 當身為研究資金補助、招募、終身職、升遷委員會的成員時，謹記衡量過程需基於研究內容本身，而非以發表計量指標為主。
2. 不論在何時，適當的文獻引用應是引用第一次針對該主題報告研究結果的論文，而非引用綜述；應將功勞給予真正有功勞的學者。
3. 藉由使用一系列的論文計量指標作為評估個人層級的研究產出與影響力的證據。
4. 挑戰不適當仰賴期刊影響係數的研究評估結果，且採用著重於研究產出的價值與影響力的最佳實踐。

## 三、「萊登宣言」的起源與內容

萊登宣言的產生過程可謂精彩，是多位來自資訊計量學領域學者腦力激盪、反覆論證而得的成果。該份文件雖基於 2014 年 STI 研討會上專題討論的結果，但其討論議題與契機可追溯至 2013 年，分別於維也納與柏林舉行的 14<sup>th</sup> International Conference of the International Society for Scientometrics and Informetrics (ISSI) 研討會和 18<sup>th</sup> International Conference on Science and Technology Indicators (STI) 研討會 (de Rijcke, 2014)。在 2013 年的 ISSI 研討會中，Wolfgang Glänzel 與 Paul Wouters 提出了 “The dos and don'ts in individual level bibliometrics” (Wouters, Glänzel, Gläser, & Rafols, 2013)，本文將其整理如下。

### 十件不可做之事：

1. 不可將個人績效簡化為單一數字。
2. 不可使用期刊影響係數作為衡量個別研究者的研究品質。
3. 不可使用隱藏的「書目計量篩選」作為選擇依據。
4. 不可對共同作者使用武斷的權重。
5. 不可使用單一指標對科學家進行排名。
6. 不可合併無法比較衡量的指標。
7. 不可使用錯誤的統計數據。
8. 不可盲目地相信一擊命中的事物 (one-hit wonders)。
9. 不可拿蘋果與橘子來比較。
10. 不可讓截止期限與工作量迫使你放棄良好的實踐作法。

### 十件可以做之事：

1. 基本指標如出版量與被引用數仍是相關的統計衡量指標。
2. 分析研究人員的合作模式。

3. 總是合併執行量化與質化的研究方法。
4. 使用引文的內容分析 (citation context analysis)。
5. 分析個別研究人員的研究主題概況。
6. 針對分析全部著作或對特定引用區間進行比較分析做出明確的選擇。
7. 在書目計量學中結合職涯分析。
8. 仔細的清理書目計量數據並使用外部資料來源。
9. 切勿對本列表所提的十項可做與不可做之事太過絕對地使用：有時不可做之事也能在正確的情境中被使用。
10. 幫助使用者闡述與使用分析後的結果。

而 2013 年在柏林所舉行的 STI 研討會，則持續針對此議題進行討論。在一場討論場次中，資訊計量研究學者認為提供更多與評估指標有關的訊息或準則給研究人員、管理者與政策制定者有其必要，資訊計量學的社群應扮演提供上述資訊的角色。這群學者甚至討論了應該由哪個組織肩負這樣的責任，然而科學計量與資訊計量重要的學會組織 ISSI 似乎不適合扮演這樣的角色，相反地，應該由主要的資訊計量研究機構擔任或協調工作 (Wouters et al., 2013)。

除此之外，在 STI 研討會中，亦有多位學者闡述指標評比結果引導績效評估或預算資源分配的現象。在資源有限的情況下，許多科學政策制定者會倚賴資訊計量評估結果作為資源分配的依據，甚至在以證據力為基礎的政策制定過程中，反而促使資訊計量指標作為績效評估的衡量指標。來自德國的社會學學者 Jochen Gläser 指出，許多政策制定者往往忽略精心製作的資訊計量分析結果報告，反而只以「摘櫻桃」的方式，選擇可合法解釋所欲推出的政策之評比結果 (Wouters et al., 2013)。換言之，強調以證據力為基礎的政策制定過程，一不小心卻轉變成以政策制定為基礎的證據分析過程，面對此現象不可不慎。因此，在場多位學者認為，在接下來幾屆的 ISSI 或 STI 研討會中，應該持續討論該議題，甚至應思考如何提出標準，讓各利害關係人能更正確解讀與使用資訊計量指標。

也因此，2014 年的 STI 研討會持續討論，並產生共識。資訊計量學領域的多位學者認為資訊計量學社群應肩負起支持、倡議正確使用資訊計量指標的責任。與會學者更認為評估準則若建立在提供特定指標或許過於狹隘，相反地，學者們倡議應該聯合發表文章，引導外界應如何正確看待「量化評量」一事。Diana Hicks 隨即提出了科技指標的品質標準準則，包括七點呼籲：

1. 指標並非用來取代評估。
2. 需投入時間與金錢以生產高品質的數據。
3. 指標需為透明且可取得的。
4. 數據可受到被評估者的驗證。

5. 對於領域的差異應有所警覺。
6. 使用標準化的數據可降低因領域與隨時間變化而造成差異的現象。
7. 指標應與策略目標有所結合。

從上述七點呼籲可觀察到，這份倡議已不侷限於告知外界何種指標才是適當指標，反而有形成整體學術評估 (research evaluation) 準則的企圖心，指引學術社群正確使用指標的觀念、評估時機，甚至導正看待學術評估的觀點。Diana Hicks 也認為這份準則應該發表於頂級期刊，如 *Nature* 或 *Science*，方能產生真正的影響 (de Rijcke, 2014)。而同時在 STI 研討會所舉行的 European Network of Indicator Designers (ENID) 年會裡，主席 Ton van Raan 認為應協調與會人員基於 Diana Hicks 所提出的七點準則，共同撰寫此份倡議文件。隨即在 2015 年 4 月 23 日，由五位資訊計量學領域學者聯合具名，在 *Nature* 期刊發表萊登宣言。

以下為萊登宣言的內容節錄。<sup>2</sup>

1. 量化評估應支援質化評估與專家評量。
2. 衡量績效表現需基於機構、團體或研究者的使命。
3. 保護與在地相關的卓越研究。
4. 數據蒐集與分析過程應為公開、透明與簡潔。
5. 允許被評估者檢驗數據與分析。
6. 留意論文發表數量與被引用次數在不同領域上的差異。
7. 個人層級的研究評量基礎應是對其研究成果歷程檔案進行質化的判斷。
8. 避免錯誤的具體性與虛假的精確性。
9. 體認評量和指標所造成的系統性影響。
10. 定期審視與更新指標。

#### 四、「舊金山宣言」刊出後的迴響與討論

在舊金山宣言推出前後，學術界濫用期刊影響係數作為評估指標一事已受到各界討論。包含 *Science*, *EMBO Reports* 與 *Proceeding of the National Academy of Sciences of the United States of America (PNAS)* 等知名學術期刊均透過刊登編輯評論 (editorial material) 文章，從期刊編輯觀點評析學術界過分依賴期刊影響係數的案例，並說明其對舊金山宣言的期許。

由於舊金山宣言並非以期刊論文的方式呈現，而是在刊登於 ASCB 的官方網頁中，因此較難追蹤其被引用狀況。因此本文以 “<http://www.ascb.org/dora/>”、 “San Francisco Declaration on Research Assessment” 在參考文獻欄位檢索，並於

<sup>2</sup> 萊登宣言繁體中文版「為研究計量指標而訂的萊登宣言」一文，全文可於該網站下載，[http://www.leidenmanifesto.org/uploads/4/1/6/0/41603901/leiden\\_manifesto\\_translated\\_traditional\\_chinese\\_final\\_20151127.pdf](http://www.leidenmanifesto.org/uploads/4/1/6/0/41603901/leiden_manifesto_translated_traditional_chinese_final_20151127.pdf)

題名、摘要、關鍵字欄位檢索“San Francisco Declaration on Research Assessment”、“DORA”，整合檢索後分析舊金山宣言自2013年發表以來的被引用狀況。在Scopus中，共有191篇文章引用或提及DORA，值得注意的是，其中有45篇為編者言；Web of Science (WoS)中共有32篇文章，13篇是編者言，顯示相對來說，主編較其他學術社群成員更重視這個議題，並透過編者言宣示所編輯期刊的立場。刊登相關文章的期刊清單與文章數量如表1所示。

表1 引用舊金山宣言之期刊論文

期刊	文章數量		期刊
	Scopus	WoS	
<i>Journal of Informetrics</i>	6	4	<i>Scientometrics</i>
<i>Scientometrics</i>	6	2	<i>Journal of Informetrics</i>
<i>Soziale Welt</i>	6	2	<i>Minerva</i>
<i>PLoS One</i>	5	--	
<i>Journal of Korean Medical Science</i>	4	--	
<i>Roczniki Psychologiczne</i>	4	--	
<i>Canadian Journal of Chemical Engineering</i>	3	--	
<i>Elife</i>	3	--	
<i>European Science Editing</i>	3	--	
<i>Insights</i>	3	--	
<i>Proceedings of ISSI 2015</i>	3	--	
Others	145*	24**	Others
合計	191	32	

\*少於三篇，\*\*少於兩篇

多位期刊編輯認為期刊影響係數確實存有其缺陷，包含被引用數分布實為偏態，使用平均值的概念予以計算將無法反映實際情況 (Jacobs, 2013; Pulverer, 2015)。Pulverer (2015)以EMBO Journal於2013年所累積被引用次數之分布情形為例，指出僅僅20%的論文即貢獻了該期刊80%的被引用次數，與Nature期刊相比，Nature期刊影響係數高於EMBO Journal，然而前者卻擁有較高比例的未被引用論文，故使用期刊影響係數評比期刊並不適切。Jacobs (2013)指出，舊金山宣言試圖消弭期刊影響係數作為研究作品品質優劣的評估指標，然而我們又該如何評斷各種指標的優劣與適用性？其觀察到有越來越多的科學家或研究機構會在其個人履歷或著作列表上標註每篇論文所獲得的被引用次數、使用h-index來呈現個人研究品質的高低，甚至斤斤計較每篇論文的作者序與被引用次數之間的評比關係，這些現象似乎並非舊金山宣言起草者當初所樂見的。Alberts (2013)以Science期刊的主編身分發表文章，認為期刊影響係數的設計初衷為呈現期刊品質，而非作為評斷個人研究品質高低的標準，然而學界對影響係數的著迷卻走火入魔，違反初衷。Alberts (2013)觀察到學者在個人履歷上每篇論文著作後面標註刊登期刊的影響係數，且數值精確到小數第三位；甚至若期刊影響係數的數值低於5，在部分國家就視該期刊毫無價值，這些誤用情

況不僅荒謬且毫無意義。*PNAS*的主編Verma (2015) 則撰文發表對濫用期刊影響係數的現象觀察，認為若預期學者在選擇投稿期刊時，期刊影響係數並未扮演重要角色，是不切實際的看法；事實上，許多令人激賞或重要的研究論文往往確實是發表在具有高影響係數的期刊。儘管如此，學術界仍須對獨尊期刊影響係數的文化進行反省，例如某些大學在審查教職申請人的著作時，除非申請者擁有以第一作者身分在高影響係數期刊發表的論文，否則便不列入考慮；或某些機構以所累計刊登期刊的影響係數點數作為學者或研究人員職涯升遷與否的門檻值；甚至祭出金額獎勵，以激勵學者於高影響係數的期刊發表論文。

對照上述種種畸形的制度或文化，多位期刊編輯均肯定舊金山宣言試圖打破學術界對於獨尊期刊影響係數一事的呼籲。Verma (2015) 指出，美國國家科學院 (National Academy of Sciences) 在2015年年會上，針對學術界濫用期刊影響係數進行討論，並進而減少對其過度依賴；而*PNAS*也強調應該重視的是學術論文的影響力，而非期刊影響係數。Alberts (2013) 則認為舊金山宣言所列出的建議有助於學術界建立「健康」且正確的學術評估態度，最低底線在於學術社群的領袖們應肩負起審慎檢視研究人員學術貢獻之責任，仔細閱讀研究人員所列出的出版品方具有意義，而非理所當然地將責任交付於期刊編輯手中。

隨著簽署舊金山宣言的機構數越來越多，也開始有大學與研究機構改變了學術評估的流程。例如Schmid (2013) 發表「雇主宣言」，說明所屬系所因簽署舊金山宣言而改變其招聘流程，不再根據申請者是否曾在頂級期刊發表論文作為是否可通過篩選而進入實質審查的標準，反之以申請者的求職信 (cover letter) 所描述的過往研究表現與對未來職涯規劃的內容作為篩選依據。此外，亦有學會宣示將施行變革，例如The European Molecular Biology Organization (EMBO) 決定不再依賴指標作為評選依據的模式，改為鼓勵申請者遞交重要著作清單，並提供該篇論文主要貢獻的說明文字，以作為審查委員會參考的依據，而非僅仰賴書目計量指標計算後的數值 (Pulverer, 2015)。儘管如此，仍有學者專家對舊金山宣言所帶來的改變持保守態度，認為仍需觀察。知名的資訊計量學專家Henk Moed在接受*Nature*期刊訪談時表示，期刊影響係數恐仍對招聘過程造成影響，其原因在於評估者可能依舊私下自行計算每位求職者列於個人履歷上每篇文章的期刊影響係數作出排名；抑或部分大學排名系統在評估構面上，仍計算各校發表在*Nature*或*Science*等高聲望期刊的論文篇數，還是會促使各校鼓勵研究人員將論文發表於頂級期刊，而忽略研究的本質 (Reich, 2013)。因此，即便舊金山宣言的主要訴求在於呼籲各界停止使用期刊影響係數作為任何層級的評估指標，因該指標的計算方式有其缺陷；或應檢視論文的內容與品質，盡量以質化方式進行評估，降低使用量化指標做為評估主要依據的作法，但對於指標的不正確認知與濫用仍然可能繼續發生。與其倡議應該使用或不應使用哪些指

標，或許我們更加需要的是較為全面的指導原則。換句話說，我們需要的是遵照舊金山宣言的「精神」，而非「設計更多更複雜或號稱更為客觀的指標系統」。Jacobs (2013) 認為我們應需要第二代的舊金山宣言，一套眾人有共識卻又可靈活運用的指導原則，以更有意義的評判科學研究系統下的產出，綜合許多主觀的審查意見或許仍優於任何僅有假象卻客觀的評判系統。

## 五、「萊登宣言」刊出後的迴響與討論

資訊計量學家 Henk Moed 接受 *Nature* 期刊訪問時曾指出，文獻計量學家正嘗試改善期刊品質的評估方式，同時亦正積極地教育研究人員有關評估指標的價值與限制 (Reich, 2013)。由文獻計量學領域的學者專家們所提出的萊登宣言，或可視為舊金山宣言的迴響與延伸。然而舊金山宣言引發的討論多限於在期刊編輯社群間，萊登宣言則在資訊計量學界與其他學術社群中均引發迴響。我們利用 Scopus 與 Web of Science (WoS) 分別進行分析<sup>3</sup>，截至 2017 年 3 月 21 日，刊登於 *Nature* 期刊的萊登宣言一文被引用次數分別在 Scopus 中為 130 次，WoS 中為 55 次。在兩個引文索引資料庫中，引用次數最多的期刊均為資訊計量學中重要的期刊 *Journal of Informetrics* 與 *Scientometrics*，另外跨領域的巨型期刊 (mega journal) *PLoS One* 也有多篇文章引用萊登宣言。刊登相關文章的期刊清單與文章數量如表 2 所示。

表 2 引用萊登宣言之期刊論文

期刊	文章數量		期刊
	Scopus	WoS	
<i>Journal of Informetrics</i>	24	11	<i>Scientometrics</i>
<i>Scientometrics</i>	20	10	<i>Journal of Informetrics</i>
<i>PLoS One</i>	10	8	<i>PLoS One</i>
<i>Professional de la Informacion</i>	4	3	<i>Professional de la Informacion</i>
<i>Journal of the Association for Information Science and Technology</i>	3	3	<i>Research Evaluation</i>
<i>PLoS Biology</i>	3	2	<i>Online Information Review</i>
<i>Research Policy</i>	3	2	<i>Research Policy</i>
Others	63*	16**	Others
合計	130	55	

\* 少於三篇，\*\*少於兩篇

萊登宣言甚至也引發了各國政府及相關單位重視「使用指標的態度」一事。舉例來說，英國高等教育撥款委員會 (Higher Education Funding Council for England, HEFCE) 於 2015 年 7 月出版一份名為 *The metric tide: Report of the Independent review of the role of metrics in research assessment and management* 的

<sup>3</sup> 檢索日期為 2017 年 3 月 21 日。

報告，指出負責任的指標 (responsible metrics) 應具備以下五點特徵：

1. 穩健性 (robustness)：盡可能在數據的準確性與範圍達到最佳的衡量。
2. 謙遜 (humility)：體認到量化評估應該是作為決策支援，而非用來取代質化、專家評估。
3. 透明 (transparency)：數據的收集與分析過程應公開且透明，讓被評估者可以測試並驗證其結果。
4. 多樣性 (diversity)：考量領域之間的差異，使用一系列的指標以反映並支持整個體系中研究人員的職涯路徑。
5. 反思性 (reflexivity)：識別與預測指標所造成的系統性與潛在性影響，並以更新做為回應。

事實上，這份報告的出版緣由即是來自於英國學術社群對於研究評估過分強調量化評估指標一事的重新檢討。報告的作者群認為即便在研究社群中，對於指標的錯誤認知與誤用仍持續不斷地發生，再加上不適當的評估指標恐將創造出不妥當的研究激勵方案 (Wilsdon et al., 2015)。這份報告也與舊金山宣言和萊登宣言被經濟合作暨發展組織 (Organization for Economic Co-operation and Development, OECD) 認為是與研究評估相關的三份重要文件，儘管研究社群在現階段並不可能自廢武功，屏除傳統的書目計量指標，然而這三份倡議文件卻可能是研究社群重新檢視指標使用的合理性轉捩點，將會對未來十年的研究評估模式造成影響 (OECD, 2016)。

此外亦有大學重新省思自身的研究評鑑制度是否需要調整。英國Loughborough University於2016年11月發表*Responsible Use of Metrics*的聲明，宣示該校體認萊登宣言的十項準則後，應以負責任的態度使用指標。Loughborough University認為外部的研究評估只會越來越頻繁的使用書目計量指標，然而相對於同儕審查，該校的研究產出若以書目計量指標進行評估，其結果並不佳，即便在已使用領域標準化的指標評估。該校也觀察到，部分書目計量指標並不能忠實地呈現該校的研究能見度，甚至在校內進行個人或機構層級的評估時，該校也應以負責的態度看待指標且合理的使用。有鑑於此，該校參考萊登宣言所提供的十項準則，作為該校研究評估操作之基準 (Loughborough University, 2016)。

荷蘭萊登大學CWTS的Ludo Waltman (亦是萊登宣言作者之一)，曾於2016年的STI研討會發表The Need for Contextualized Scientometric Analysis: An Opinion Paper一文。Waltman與van Eck (2016) 認為在進行學術評估時，指標可分為簡易指標與進階指標，各自擁有優缺點，例如簡易指標的計算過程較為簡潔，概念易於理解，然而在評估上可能有其盲點，未能達到較高的精確性；進階指標雖在計算過程中考慮較多實務的可能性，但往往也會使得計算公式較為複雜，讓人不易理解。有鑑於此，故建議應發展「具脈絡的科學計量分

析 (contextualized scientometric analysis)」，而這類分析模式可做為同儕審查的輔助工具，是基於三大原則予以運作，分別是內容脈絡 (context)、簡潔 (simplicity) 與多樣性 (diversity)。內容脈絡指的是在選用科學計量指標時，須明確告知是在何種情境脈絡下進行評估，方能使專家在進行同儕審查時能清楚為何選擇這樣的指標進行評估，如此一來，指標選擇方有意義且具有代表性；簡潔指的是評估指標最好使用既簡單又易理解的指標，例如若只是針對特定單一領域進行評估，選擇一般的平均引用數指標即可，並不需再使用進階指標，如針對不同領域進行標準化後而取得的平均引用數指標；多樣性則是指評估指標不一定要侷限在科學計量指標，畢竟許多引用相關的指標，是基於 Web of Science 或 Scopus 資料庫所收錄的書目資料進行計算；相反地，使用其他資料來源而計算的指標亦可協助專家進行同儕審查或展開評估工作，如投入面的資訊，包含研究人員數或研究經費投入規模，這些指標將可協助專家勾勒完整的面貌，進行詮釋 (Waltman & van Eck, 2016)。簡言之，若能妥善處理指標選擇與結果詮釋，運用資訊計量指標進行評估，作為同儕審查的輔助決策資訊，指標的使用仍具有其意義，不需以看待洪水猛獸之姿，對資訊計量分析做出負面評論，甚至一味排斥。

## 六、「萊登宣言」與「舊金山宣言」的差異

舊金山宣言與萊登宣言的最大差異在於，前者是以期刊主編的角度出發，針對不同利害關係者提出各自的研究評估準則呼籲，且特別在意學術社群對於期刊影響係數的濫用情形，並認為該指標有瑕疵；後者則是出自於資訊計量學領域學者的自省。相較於舊金山宣言，萊登宣言則是不分對象的全面宣示，在其所提出的十項準則中，除每項準則在提出後皆有所補充說明，充分讓讀者明白該準則產生的動機或預期可改善的現象為何外，萊登宣言的內容亦著重在各式資訊計量指標的使用時機與計算過程需留意之處。

就宣傳策略來說，舊金山宣言是依附於 ASCB 組織之中，萊登宣言則未棲身於任何資訊計量學領域組織學會 (如 ISSI Society) 的網站裡，反而是由一群資訊計量學領域的學者共同撰文後，投稿至 *Nature* 期刊，試圖在跨領域的頂級期刊中，取得更高的能見度。萊登宣言的目標群眾設立更為廣泛，企圖在全球的學術社群引發討論熱潮，而非「只」針對特定領域的群眾。如同 David 與 Frangopol (2015) 所述，萊登宣言所提及的十項準則，除部分準則為闡述資訊計量學的概念或蒐集資料、指標計算時應注意的事項外 (準則四到六)，有些準則更是屬於一般科學評估 (science evaluation) 的範疇，如準則二、準則九與準則十。可見萊登宣言的目標群眾已不僅鎖定資訊計量學領域的學者，而是更擴大至整個學術社群中從事科學評估的個體，甚至包含評估者與被評估者。除此之

外，由於萊登宣言選擇發表於*Nature*期刊，就其文體來說，屬於期刊論文，此舉或許也促使萊登宣言的擴散能更加迅速，藉由文章引用，取得更高的能見度。

另一方面，舊金山宣言與萊登宣言雖皆建置了網頁或網站，但兩者的介面排版上也略有差異。舊金山宣言的網頁<sup>4</sup>除提供宣言全文，更設置了簽署頁面。萊登宣言則是由發起者建置了單獨網站，提供全文連結。值得一提的是，萊登宣言網站<sup>5</sup>亦有宣傳影片，試圖提供更為多元的宣傳途徑(Leiden Manifesto for Research Metrics, 2015)。此外萊登宣言還具備多種語言的翻譯，截至目前為止，網站上已有16種語言，其中包含繁體中文版本。各種語言的翻譯工作完全是各國資訊計量學或相關領域學者的自發性舉動，在學者們完成翻譯後，聯繫萊登宣言的作者群，請其協助將翻譯版本置於網站中。此舉除可顯現各國皆有學者對萊登宣言一文感到認同外，更希望藉由文件翻譯，讓該國學術社群的參與者知曉此內容，並趁機宣達正確的指標使用態度，以達到萊登宣言最終的目的。

## 七、結語

近年來，政府或機構層級對於學術社群成員的研究表現評估，雖多以各種指標評比並搭配同儕審查的方式進行，然而對於研究評估指標的選擇，或應佔整體研究評估的比重多寡，仍有諸多討論。甚至，學術競爭力與研究評估指標的使用，是否屬於「胡蘿蔔與棍子」的關係，因而需要重新檢討整個學術評鑑制度，亦是最近多項與科技政策或學術研究政策有關的會議中，引發熱烈討論的議題之一。檢視外界對於現行學術評鑑制度的批評，最多的指責聲浪為國內學界過度重視SCI與SSCI論文，導致學者發表論文僅重視期刊影響係數的積點，而忽略了研究的本質；在過度追求論文引用所產生之影響力的同時，卻漠視了研究更應肩負具有社會影響力的使命。

然而，期刊影響係數或研究評估相關指標是否如此的萬惡不赦？在檢視舊金山宣言與萊登宣言的內容以及整理國際學術社群對這些倡議文件的回應後，可發現真正的問題其實不在於研究評估指標本身，而是研究評估指標的正確使用時機與對待方式。從舊金山宣言的內容可得知，研究評估不應獨厚單一指標，特別在期刊影響係數的計算易造成與事實偏誤的情形下，任何的研究評估都應回歸到研究的內容檢視。而萊登宣言更提出十項研究評估的準則，認為指標本身並未有錯，應被檢討的是使用指標的時機與看待方式，包含指標挑選應符合評估目標，指標計算過程應公開透明，資訊計量指標評比結果應作為研究評估的輔助而非取代質性的同儕審查。

在導正研究評估過度重視資訊計量指標評比結果，甚至學術研究只追求期

<sup>4</sup> 舊金山宣言網頁：<http://www.ascb.org/dora/>

<sup>5</sup> 萊登宣言網站：<http://www.leidenmanifesto.org/>

刊論文積點而忽略研究本質的畸形現象時，或許可從舊金山宣言與萊登宣言的內容獲得啟示。因此，本文建議未來資訊計量學界與學術研究政策領域應合作推廣舊金山宣言與萊登宣言的內容，以讓學界與相關利害關係人重新體認應如何正確看待資訊計量指標，瞭解其發展脈絡；甚至也應提升國內學界對於資訊計量指標的認知，包含指標的名詞統一與定義上的介紹。唯有正確認識，才有正確使用的可能性，如此一來，方可凝聚對於使用研究評估指標的共識，甚至發展出具備在地特質的宣言。

## 參考文獻

- Alberts, B. (2013). Impact factor distortions. *Science*, 340(6134), 787. doi:10.1126/science.1240319
- David, D., & Frangopol, P. (2015). The lost paradise, the original sin, and the Dodo bird: A scientometrics Sapere Aude manifesto as a reply to the Leiden manifesto on scientometrics. *Scientometrics*, 105(3), 2255-2257. doi:10.1007/s11192-015-1634-2
- de Rijcke, S. (2014). The Leiden Manifesto in the making proposal of a set of principles on the use of assessment metrics in the S&T indicators conference. Retrieved from <https://citationculture.wordpress.com/2014/09/15/the-leiden-manifesto-in-the-making-proposal-of-a-set-of-principles-on-the-use-of-assessment-metrics-in-the-st-indicators-conference/>
- DORA. (2012). San Francisco Declaration on Research Assessment (DORA). Retrieved from <http://www.ascb.org/dora/>
- Halevi, G. (2014). Reporting back: STI 2014 Leiden, The Netherlands. *Research Trends*, 39. Retrieved from <http://www.researchtrends.com/issue-39-december-2014/reporting-back-sti-2014-leiden-the-netherlands/>
- Hicks, D., Wouters, P., Waltman, L., de Rijcke, S., & Rafols, I. (2015). Bibliometrics: The Leiden Manifesto for research metrics. *Nature*, 520(7548), 429-431. doi:10.1038/520429a
- Jacobs, H. (2013). Dear DORA. *EMBO Reports*, 14(11), 947-947.
- Leiden Manifesto for Research Metrics. (2015). Leiden Manifesto for research metrics. Retrieved from <http://www.leidenmanifesto.org>
- Loughborough University. (2016). Using bibliometrics responsibly. Retrieved from <http://www.lboro.ac.uk/research/scholcomms/assessment/respmetrics/>
- OECD. (2016). *OECD science, technology and innovation outlook 2016*. Paris, France: Author. doi:10.1787/sti\_in\_outlook-2016-en
- Pulverer, B. (2015). Dora the brave. *The EMBO Journal*, 34(12), 1601-1602. doi:10.15252/embj.201570010
- Reich, E. S. (2013). Science publishing: The golden club. *Nature*, 502(7471), 291-293. doi:10.1038/502291a
- Schmid, S. L. (2013). Beyond CVs and impact factors: An employer's manifesto. doi:10.1126/science.caredit.a1300186
- Verma, I. M. (2015). Impact, not impact factor. *Proceedings of the National Academy of Sciences*, 112(26), 7875-7876. doi:10.1073/pnas.1509912112

- Waltman, L., & Van Eck, N. J. (2016). The need for contextualized scientometric analysis: An opinion paper. In I. Ràfols, J. Molas-Gallart, E. Castro-Martínez, & R. Woolley (Eds.), *Proceedings of the 21st International Conference on Science and Technology Indicators* (pp. 541-549). Valencia, Spain: Editorial Universitat Politècnica de València. doi:10.4995/STI2016.2016.4543
- Wilsdon, J., Allen, L., Belfiore, E., Campbell, P., Curry, S., Hill, S., ... Johnson, B. (2015). *The metric tide: Report of the independent review of the role of metrics in research assessment and management*. doi:10.13140/RG.2.1.4929.1363
- Wouters, P., Glänzel, W., Gläser, J., & Rafols, I. (2013). The dilemmas of performance indicators of individual researchers: A urgent debate in bibliometrics. *ISSI Newsletter*, 9(3), 48-53.



# What We Have Learned from San Francisco Declaration on Research Assessment and Leiden Manifesto?

Carey Ming-Li Chen<sup>ab\*</sup> Wen-Yau Cathy Lin<sup>c</sup>

## Abstract

In recent years, the research performance evaluation of members of the academic community conducted by government or institutions has been applied with multiple indicators and peer review, however, there are many controversies about the design and application of research evaluation indicators. This article aims to introduce the development process of *San Francisco Declaration on Research Assessment (DORA)* and *Leiden Manifesto* and summarize their contents of guidelines and attempts to compare the differences between these two documents. It hopes that this article can arise the attention and reflection of research evaluation indicators and relevant issues from Taiwan academic community to reach consensus of utilization of research evaluation indicators. It will be beneficial to develop the version of declaration with local characteristics in the future.

**Keywords:** *San Francisco Declaration on Research Assessment (DORA), Leiden Manifesto, Informetrics, Research evaluation indicator*

## SUMMARY

## Introduction

In April of 2015, an article titled as *Bibliometrics: The Leiden Manifesto for research metrics* (Leiden Manifesto as the short form) was published on the journal *Nature*, in which several scholars from informetrics field discussed about some worrying phenomenon, such as abuse of indicators of research evaluation, conflicts between metrics analysis and results of research assessment, and over-emphasis on citation-based indicators; they also proposed ten principles for academic research assessment (Hicks, Wouters, Waltman, de Rijcke, & Rafols, 2015). In fact, this is not the first declaration about research evaluation indicators

<sup>a</sup> Assistant Researcher, Science & Technology Policy Research and Information Center, National Applied Research Laboratories, Taipei, Taiwan

<sup>b</sup> Doctoral student, Graduate Institute of Business Administration, National Taiwan University

<sup>c</sup> Associate Professor, Department of Information and Library Science, Tamkang University, Taipei, Taiwan

\* Principal author for all correspondence. E-mail: carey.mlchen@gmail.com

proposed by academic communities. In December of 2012, in the annual meeting of The American Society for Cell Biology (ASCB), a group of academic journal editors and publisher representatives issued together a suggestion about research assessment, titled as *San Francisco Declaration on Research Assessment (DORA)*. In this declaration, 18 principles about research assessment were proposed, including overall principles and specific principles for groups of different stakeholders. In this article, the authors focus on introduction and analysis of these two declarations, with an intention to increase awareness and reflections of Taiwan academic communities, for reaching a consensus and even developing a local version of Leiden Manifesto.

### **Reactions and Discussions after DORA was published**

After San Francisco Declaration on Research Assessment was published, many well-known academic journals discussed cases of over-relying on journal impact factor and their expectations about DORA in their editorials. Through an integrated search of articles ever citing or mentioning DORA, we can see 24% (45/191) are editorials in Scopus and 41% (13/32) in Web of Science, indicating that editors-in-chief had a higher awareness of this issue than other academic community members and declared the positions of the represented journals. With the increasing number of supporting institutes, some universities and research institutes changed their process of research assessment. However, some scholars stated that, even though the major appeal of DORA is calling for stopping using the journal impact factor as the evaluation indicator in any level, or scrutinizing content and quality of articles instead of using quantitative metrics as the major assessment approach, the problem of misconceptions and abuses of metrics would still exist. Instead of proposing which metrics should or should not be used, maybe it is more important to make a set of overall and comprehensive guidelines. In other words, we need a second version of DORA- a set of guidelines with people's consensus and for flexible applications (Jacobs, 2013).

### **Reactions and Discussions after Leiden Manifesto was published**

Different from DORA, Leiden Manifesto was a self-reflection of infometrics scholars, and an outcome of their brainstorming and debating. The proposed principles include stating concepts of informetrics and reminders about collecting data and computing metrics, as well as suggestions for general science assessment, indicating that their target groups include both the assessors and the assessed in the whole academic communities. From the search results in citation databases, we can see that the most citing journals include major journals in infometrics field, such as *Journal of Informetrics* and *Scientometrics*, as well as

the interdisciplinary mega journal *PLoSOne*. The Leiden Manifesto also raised the awareness of international institutes. For example, Higher Education Funding for England (HEFCE) published a report titled as *The metric tide: Report of the independent review of the role of metrics in research assessment and management*, stating the five features of responsible metrics, including robustness, humility, transparency, diversity and reflexivity (Wilsdon et al., 2015). This report, together with DORA and Leiden Manifesto, are considered the three important documents of research assessment. Although it is impossible that the research communities would stop using traditional bibliometric indicators at this point, it might be a turning point for the community to re-examine the reasonableness of using metrics. It would have a great impact on the research assessment model of the next decade (OECD, 2016).

### **Differences between Leiden Manifesto and DORA**

In terms of propaganda strategies, DORA is included in ASCB, while the Leiden Manifesto is not belonged to any societies of informetrics, such as ISSI Society. The Leiden Manifesto was co-authored by a group of informetrics scholars and submitted to the journal *Nature*. With a form of journal article, it could be cited and thus more well-known. The two documents are both posed on web pages, but the layout arrangements are somehow different. The web page of DORA has a signature page, while the Leiden Manifesto web page provides a more diverse set of propaganda channels, such as offering full-text translations of 16 languages at this point, including Traditional Chinese version. The spontaneous action of these international scholars indicates they not only agree on the content of Leiden Manifesto, but also would like to raise the awareness of academic community members in their countries and propose a correct attitude of using metrics.

### **Conclusions**

A highly debated issue in the Taiwan academic community is whether the relationship between academic competitiveness and utilization of research assessment metrics is “carrot and stick”; and thus it needs to re-examine the whole research evaluation system. The most criticized phenomenon is that the whole academic community over-emphasizes journal articles of SCI and SSCI, so scholars care only about accumulated points of journal impact factor and thus neglect the essence of research; while scholars are over-pursuing the number of citations, they forget that their researches should be with more societal impact. However, are these informetric indicators so unpardonably vicious? After re-examining the contents of DORA and Leiden Manifesto, as well as international

academic communities' reactions toward these proposed documents, we can see that the real problem is not in research assessment indicators themselves, but in the timing and approaches of using these indicators. When correcting the deformed phenomenon, maybe we can gain some insights from the declarations. Therefore in this article, we suggest that in the future the informetrics field and research policy field should cooperate to promote the declaration contents, for academic community members to have a correct view toward informetrics and understand the contexts of development, as well as raise their awareness of informetrics, including consistent terminology and definitions. Only when scholars have a correct knowledge about informetrics, they can use correctly, and thus even develop a declaration suitable for the local academic community in Taiwan.

#### **REFERENCE FOR ORIGINAL TEXT**

- Alberts, B. (2013). Impact factor distortions. *Science*, 340(6134), 787. doi:10.1126/science.1240319
- David, D., & Frangopol, P. (2015). The lost paradise, the original sin, and the Dodo bird: A scientometrics Sapere Aude manifesto as a reply to the Leiden manifesto on scientometrics. *Scientometrics*, 105(3), 2255-2257. doi:10.1007/s11192-015-1634-2
- de Rijcke, S. (2014). The Leiden Manifesto in the making proposal of a set of principles on the use of assessment metrics in the S&T indicators conference. Retrieved from <https://citationculture.wordpress.com/2014/09/15/the-leiden-manifesto-in-the-making-proposal-of-a-set-of-principles-on-the-use-of-assessment-metrics-in-the-st-indicators-conference/>
- DORA. (2012). San Francisco Declaration on Research Assessment (DORA). Retrieved from <http://www.ascb.org/dora/>
- Halevi, G. (2014). Reporting back: STI 2014 Leiden, The Netherlands. *Research Trends*, 39. Retrieved from <http://www.researchtrends.com/issue-39-december-2014/reporting-back-sti-2014-leiden-the-netherlands/>
- Hicks, D., Wouters, P., Waltman, L., de Rijcke, S., & Rafols, I. (2015). Bibliometrics: The Leiden Manifesto for research metrics. *Nature*, 520(7548), 429-431. doi:10.1038/520429a
- Jacobs, H. (2013). Dear DORA. *EMBO Reports*, 14(11), 947-947.
- Leiden Manifesto for Research Metrics. (2015). Leiden Manifesto for research metrics. Retrieved from <http://www.leidenmanifesto.org>
- Loughborough University. (2016). Using bibliometrics responsibly. Retrieved from <http://www.lboro.ac.uk/research/scholcomms/assessment/respmetrics/>
- OECD. (2016). *OECD science, technology and innovation outlook 2016*. Paris, France: Author. doi:10.1787/sti\_in\_outlook-2016-en
- Pulverer, B. (2015). Dora the brave. *The EMBO Journal*, 34(12), 1601-1602. doi:10.1525/embj.201570010
- Reich, E. S. (2013). Science publishing: The golden club. *Nature*, 502(7471), 291-293. doi:10.1038/502291a

- Schmid, S. L. (2013). Beyond CVs and impact factors: An employer's manifesto. doi:10.1126/science.caredit.a1300186
- Verma, I. M. (2015). Impact, not impact factor. *Proceedings of the National Academy of Sciences*, 112(26), 7875-7876. doi:10.1073/pnas.1509912112
- Waltman, L., & Van Eck, N. J. (2016). The need for contextualized scientometric analysis: An opinion paper. In I. Rafols, J. Molas-Gallart, E. Castro-Martínez, & R. Woolley (Eds.), *Proceedings of the 21st International Conference on Science and Technology Indicators* (pp. 541-549). Valencia, Spain: Editorial Universitat Politècnica de València. doi:10.4995/STI2016.2016.4543
- Wilsdon, J., Allen, L., Belfiore, E., Campbell, P., Curry, S., Hill, S., ... Johnson, B. (2015). *The metric tide: Report of the independent review of the role of metrics in research assessment and management*. doi:10.13140/RG.2.1.4929.1363
- Wouters, P., Glänzel, W., Gläser, J., & Rafols, I. (2013). The dilemmas of performance indicators of individual researchers: A urgent debate in bibliometrics. *ISSI Newsletter*, 9(3), 48-53.