第 3 卷第 4 期，2010 年 10 月

国际高等教育

Journal of International Higher Education, Vol. 3, No. 4, October, 2010
Graduate School of Education, Shanghai Jiao Tong University

排名专刊

上海交通大学高等教育研究院　主办

主编：刘念才
顾问：菲利普·G·阿特巴赫（美）
副主编：王琪
编辑：吴燕

刊号：沪（K）第 0621 号

编辑部地址：
上海市东川路 800 号上海交通大学高等教育研究院
邮编：200240
电话：021-34207986
电邮：jihe@sjtu.edu.cn
http://gse.sjtu.edu.cn/jihe/index.htm

Editor: Nian Cai Liu
Advisor: Philip G. Altbach (USA)
Associate Editor: Qi Wang
Assistant Editor: Yan Wu

Address:
Graduate School of Education
Shanghai Jiao Tong University
800, Dongchuan Road
Shanghai, 200240
China
Tel: 86-21-34207986
Email: jihe@sjtu.edu.cn
http://gse.sjtu.edu.cn/jihe/index.htm

《国际高等教育》简介

随着高等教育国际化趋势的日益加剧和中国高等教育的持续快速发展，中国与国外高等教育交流的需求与日俱增。《国际高等教育》正是在这样的背景下诞生的，并将努力发挥中外高等教育交流的桥梁作用。

本刊由两部分组成：美国波士顿学院国际高等教育中心（http://www.bc.edu/bc_org/avp/soe/cihe/）编辑出版的“International Higher Education”的中文版，每年 4 期；由上海交通大学高等教育研究院围绕中国与国外高等教育热点问题编辑出版的专辑，每期一个主题，每年最多出 6 期。

波士顿学院国际高等教育中心编辑的“International Higher Education”每期包含十余篇有关国际高等教育现状和发展趋势的文章以及新书简介。上海交通大学高等教育研究院编辑的专辑则包含世界一流大学、大学评价与排名、研究生教育、科学英才迁移等热点问题的国内外研究论文、研究报告、综述文章、书评以及有关的会议和活动信息。

Journal of International Higher Education (Guoji Gaodeng Jiaoyu) is an online journal with an aim of playing the role of bridge between Chinese and international higher education communities. It consists of 4 issues each year translated from the “International Higher Education” published by the Boston College Center for International Higher Education (http://www.bc.edu/bc_org/avp/soe/cihe/) in USA and up to 6 issues each year focusing on selected topics of both Chinese and international interests, such as world-class universities, university ranking, graduate education, and migration of academic talents.

Each issue translated from the International Higher Education by Boston College contains more than a dozen of short articles covering major aspects and trends of international higher education and new publications. Each issue focusing on selected topics of higher education contains original studies, preliminary reports, review papers, letter to the editor, book reviews, and up-to-date information on activities and opportunities on the selected topics in China and around the world. The issues focusing on selected topics will be translated into English and published online at http://gse.sjtu.edu.cn/en/.

世界大学学术排名

刘念才

上海交通大学高等教育研究院院长、教授

电子邮箱：ncliu@sjtu.edu.cn

建立世界一流大学是几代中国人的梦想。1998 年 5 月，时任国家主席的江泽民同志在庆祝北京大学建校 100 周年大会上向全社会宣告：“为了实现现代化，我国要有若干所具有世界先进水平的一流大学”。随后，中国政府开始实施“985 工程”，重点建设若干所世界先进水平的一流大学和一批一流学科。1998 年，上海交通大学成为“985 工程”第一批建设的 9 所大学之一。当时有不少大学都提出了创建世界一流大学的奋斗目标和时间表，上海交通大学也提出了自己的建设目标。作为化学化工学院的一名教授、学院副院长，我有幸参与了学校创建世界一流大学的战略规划工作。

在起草战略规划的过程中，我的脑海中始终徘徊着若干问题。世界一流大学的特征和标准是什么？全世界应该有多少所世界一流大学？中国的大学在世界大学体系中的位置如何？中国的大学如何能够缩小与世界一流大学的差距？为了回答这些问题，我们开始对中国的大学与世界一流大学进行对比分析，并最终对全世界的大学进行了排名。

从 1999 到 2001 年，我和程莹博士等 3 位同事一起，根据学校的学术或科研表现、选取国际可比的客观数据，对中国的大学与 4 组美国的大学进行了对比分析，包括美国最顶尖的大学和普通的研究型大学。该项研究的主要结论是：中国的顶尖大学大概处在全世界 200～300 名的位置。在上海交通大学的战略规划制定过程中，采用了该项研究的对比与分析结果。我们最终还形成了名为《我国名牌大学离世界一流大学有多远》的咨询报告，提交给了教育部。

这份报告得到了大量正面评价，有许多评价都希望我们做一个真正的世界大学排名。许多为了其他目的来访的外国友人在了解了我们的研究之后，也鼓励我们做世界大学排名。他们说，不仅在中国，而且在其他国家，大学、政府和其他的利益相关者都对世界大学排名有兴趣。在众多的鼓励下，我决定制作一个真正的世界大学排名。我与 3 位同事又用了 2 年时间进行了更多的研究，终于在 2003 年初完成了“世界大学学术排名（Academic Ranking of World Universities，ARWU）”，并于 2003 年 6 月在网站（http://www.arwu.org）上公布了排名。

ARWU 排名方法

ARWU 采用 6 个客观指标对世界大学的学术表现进行排名。这 6 个指标及其权重分别为：获诺贝尔奖和菲尔兹奖的校友折合数（10%）、获诺贝尔奖和菲尔兹奖的教师折合数（20%）、各学科领域被引用次数最高的科学家数（20%）、在《Nature》和《Science》上发表论文的折合数（20%）、被科学引文索引（SCIE）和社会科学引文索引（SSCI）收录的论文数（20%）、上述五项指标得分的师均值（10%）等六个指标。

ARWU 的排名对象包括：所有曾经有教师或校友获得过诺贝尔奖或菲尔兹奖的大学；所有有高被引科学家的大学；过去 10 年中所有在《Nature》或《Science》杂志上作为通讯作者单位发表过论文的大学；以及各个国家被科学引文索引（SCIE）和社会科学引文索引（SSCI）收录论文数较多的大学。被扫描的大学有 2,000 余所，每年实际被排名的大学有 1,200 余所，网上公布的是处于世界前 500 名的大学。考虑到第 50 名之后大学的总体得分差异明显减缓，因此在公布排名结果时，前 100 名大学按照简单数字排序，100～200 名的大学按 50 所一组公布，而 200～500 名的大学则按 100 所一组公布，同一组内的大学按字母顺序排列。

ARWU-FIELD 和 ARWU-SUBJECT

自 ARWU 发布之后，受到了全世界的关注，收到很多希望我们按照学科领域或大学内部的学院进行排名的提议。在经过了大量研究和探索之后，我们于 2007 年 2 月和 2009 年 10 月，分别又推出了“世界大学学科领域排名（Academic Ranking of World Universities by Broad Subject Fields，ARWU-FIELD）”和“世界大学学科排名（Academic Ranking of World Universities by Subject Fields，ARWU-SUBJECT）”。

ARWU-FIELD 进行排名的五个学科领域分别为理学、工学与计算机科学、生命科学与农学、临床医学与药学以及社会科学。ARWU-FIELD 中没有对人文与艺术领域的排名，是因为找到可靠且具有全世界可比性的数据存在很多技术难题。而心理学和其他跨学科领域则因其跨学科的特点而没有被排名。在学科领域排名中，我们引入了两个新的指标：第一个是“高质量论文比例”指标，即发表在每个学科领域前 20%期刊上的论文比例；第二个是在工科排名中增加了“年度科研经费”指标，替代了获奖校友和获奖教师指标。在排名中公布的是每个学科领域中排名前 100 名的大学。

ARWU-SUBJECT 对五个学科进行了排名，分别是：数学、物理学、化学、计算机、经济学/商学。在排名中公布的是每个学科中排名前 100 名的大学。

ARWU 的影响

尽管 ARWU 的初衷是分析中国名牌大学在世界大学体系中的地位，但是作为世界首家多指标的全球性大学排名，ARWU 的发表获得了全世界的关注，产生了广泛的影响。《经济学家》杂志 2005 年的一份长篇报告提到 ARWU 是“被最为广泛使用的世界大学排名”；美国的《高等教育记事》的一篇文章中称赞 ARWU 是“最有影响力的国际排名”。

排名方法的客观性和透明性是 ARWU 的显著特点。欧盟委员会 2003 年 12 月 31 日在其科技网站上将 ARWU 作为头条新闻进行了报道并给予积极评价。牛津大学校长（Chancellor）Chris Patten 在英国议会科学委员会的演讲中提到，“ARWU 的排名方法看起来相当可靠，……是一个公正的比较”。ARWU 的结果被广泛引用作为政策分析和推进改革的依据之一。美国罗切斯特理工大学的校长 Bill Destler 在《自然》上撰文，以 ARWU 的结果为证据比较西欧和美国的创新能力。《科学》刊载的一篇文章中指出，“在 ARWU 排名中法国大学表现不佳，因此引发了法国有关高等教育的全国性辩论，并于上月（2007 年 7 月）出台了新的教育法，赋予了大学更多的自由”。

完善排名

ARWU 主要是根据有关学术或科研表现的第三方国际可比数据对全世界的研究型大学进行的排名，用于排名的数据每个人都可以核查。这项工作是出于我们的学术兴趣而独立完成的，对中国大学的战略规划产生了一定的影响。不过，ARWU 也存在很多的方法论和技术上的问题。

从方法论上讲，包括科研、教学和社会服务这三者在排名指标与权重中的合理体现，非英语发表论文的引入，国际奖项的选择，以及获奖者的经历梳理等。从技术角度讲，包括大学的界定与名称问题，数据的收集与数据库的整理，以及论文如何归属到正确的大学和学科等。目前，我们的团队正在认真研究排名中存在的所有问题，并努力完善我们的排名。

除了学科领域排名之外，我们还在探索其他类型排名的可能性，特别是根据功能、学科特征、历史、规模、经费等对不同类型的大学进行排名。我们也在不断进行各种有关排名的理论研究，为让公众更好地理解排名而努力。此外，我们还积极地参与有关排名的国际组织的活动，比如，“IREG－学术排名与卓越国际协会”（IREG—International Observatory on Academic Ranking and Excellence，http://www.ireg-observatory.org）。

THE“世界大学排名”的新方法

Phil Baty

《泰晤士报高等教育》“世界大学排名”主编、《泰晤士报高等教育》杂志副主编

电子邮箱：phil.baty@tsleducation.com

《泰晤士报高等教育》（Times Higher Education，THE）杂志自 2004 年发布“世界大学排名”以来，每年新排名的发布已成为国际高等教育日程上的重要事件。

最初，这个排名被用来帮助学生选择大学，现在则经常被用来帮助学术人员挑选合作伙伴或新雇员，以及帮助大学管理层制定学校的战略、分析学校的表现。由于国家的经济正向知识导向和创新导向型转变，大学排名也被政府和政策制定者用来制定或改革国家的高等教育政策。

2009 年 11 月，THE 公开承认，虽然已发布的排名得到了国际上的广泛关注，但却并不能很好地实现上述目标。由于该排名已经具有相当的影响力，杂志的编辑认为自己有义务使排名更严谨、准确、平衡和透明。同时，国际学术界对这个排名的重视程度远远超过预期，所以杂志需要更加负责地从事排名活动。

新的数据提供伙伴

2009 年 11 月，杂志终止了与 QS 公司长达六年的合作，该公司在 2004～2009 年期间为“世界大学排名”提供数据支持。从 2010 年起，THE 采用了新的排名方法发布世界大学排名，所有数据由国际科研信息专业机构汤森路透公司（Thomson Reuters）收集和提供。

原 THE-QS 排名方法中存在的问题已被全面深入地研究整理过。其中最有争议的是所谓的“同行评价”指标。一所大学排名总得分的 40%取决于“同行评价”的结果，而这实际上只是一项简单的意见调查，询问受访者认为哪所大学应排在世界前列。许多人原则上都反对在排名中使用任何主观评价，他们认为“同行评价”是在反映大学过去的表现而非现在的表现，评价结果往往基于传统印象，有时甚至是根本不了解。

但是，我们认为声誉指标能提供简单定量数据所不能提供的有用信息。并且，声誉在如今高度竞争的全球高等教育市场中的作用同样至关重要。而为了使声誉调查变得真正有价值，必须采用非常严谨的方法来实施。

我们之前的排名数据提供商 QS 进行的声誉调查回收的样本数量非常少，比如 2009 年大约有 3,500 人参与了调查，这只是全世界数百万学者中的很小一部分。2009 年的 THE-QS 排名中声誉调查指标得分的计算是基于过去三年调查结果的累计，共 9,386 人。从 THE 的角度来看，样本的数量太少且赋予这类指标的权重过高。

改进后的声誉调查

2010 年，我们对声誉调查做了重大改进。汤森路透公司请了专业的民意调查公司益普索媒介研究（Ipsos MediaCT）实施新的声誉调查，计划调查更多的学术人员，并且这些学术人员从学科分布和地域分布上都能够真

正代表世界大学群体的实际情况。

可能最值得注意的是，这份调查只发送给受邀参与的对象，他们对本领域都非常了解。大部分受访者是通过汤森路透公司的文献数据库筛选出的，这样可以确保他们曾经发表过研究成果。我们拒绝采取“散弹枪”的方式，即群发邮件来收集反馈，也不采取让所有学者都能自愿参与声誉调查的方式。

汤森路透公司的学术声誉调查很成功，仅仅三个多月就收集了 13,388 份反馈，是原先 THE-QS 排名时一年收集总数的四倍左右。

而重要的不只是样本的数量，调查也满足了我们对样本更具代表性的需求。约 30.2%的受访者来自亚太地区，38%来自美洲，28.3%来自欧洲，3.5%来自非洲。

反馈覆盖了所有学科：大部分受访者（22.8%）来自工程和技术领域，20.5%来自自然科学领域，17.8%来自生命科学领域，17.8%来自社会科学领域，15.3%来自医疗卫生领域，5.8%来自艺术人文领域。同时，我们还采用统计分析方法消除反馈偏差。

约 6.5%的受访者表示自己是“机构的高级管理人员”，大部分的受访者（68.9%）表示自己是“学术人员”，14.9%表示自己是“研究人员”。受访者普遍有较长的高等教育领域工作经历，平均工作年限为 17 年。当被问到主要工作内容时，54%的受访者表示是科研，31%表示是教学，12.6%表示是管理服务，还有 1.8%的人现在不工作。

2010 年，受访者首次被要求对教学和研究两方面的声誉进行分别评价，因此，2010 的“世界大学排名”中首次出现教学声誉指标。

为了明确评价内容、获得有价值的反馈，调查中的问题都是经过精心准备的。比起简单笼统的问“哪所大学最好”，调查会询问得更详细，以此来获得更有信息量和前后一致的答案，比如问受访者申请攻读研究生学位的本科毕业生中来自哪所大学的最优秀，或者他们会推荐他们最好的本科毕业生去哪里攻读研究生。

在新的排名，声誉调查指标所占权重总共为 34.5%，与原来的 THE-QS 排名相比，虽然声誉调查的方法本身得到了显著改善，但是其权重仍然被降低。

学科标准化后的引用数据

另一个关于 THE-QS 排名方法的争议即是研究质量的测量方式。在原来的方法中，QS 计算一名学者的学术成果被同行引用的次数。QS 用一所大学所有论文的被引用总次数除以大学员工的数量（全时折合）作为一项指标，权重设为 20%。

但这一方法没有把学科间不同的引用习惯考虑进去，不同学科的平均引用率各不相同。举例来说，学科领域的平均引用率（数据来自汤森路透公司的 Essential Science Indicators 数据库，1998 年 1 月 1 日～2008 年 10 月 31 日）从计算机科学的 3.06、工程领域的 3.83 到免疫学的 20.58、分子生物学的 24.75 不等。

QS 未能对引用数据进行标准化处理，从而无法反映不同学科在被引用数量上的差异，这样就对那些在平均引用率较低学科有优势的大学不利，而偏向那些在平均引用率较高的学科有优势的大学。这是不公平的，并且会产生误导。

汤森路透公司拥有涵盖 12,000 份最有影响力的期刊和超过 110,000 份会议论文集的数据库，THE 将会利用如此庞大的引文数据库制作 2010 年和以后的排名。

重要的不仅仅是数据的质量。通过与引文数据库所有者的直接合作而不是简单地从第三方购买数据，THE 还可以利用世界领先的专业知识来对引文数据进行整理和分析。

因此，从 2010 年起我们对引用数据进行了调整，把学科差异考虑进去。THE 的年度 200 强排名是在学科领域的基础上制作完成。数据将按六个大学科领域进行收集与分析，分别是：艺术和人文、医疗卫生、生命科学、自然科学、工程和技术、社会科学。2010 年的排名包括以上六个学科领域的独立排名。总的 200 强排名则根据学科领域排名的结果来制作。

上述方法的采用，体现了对原 THE-QS 排名的另一重大改进。原先的五个学科领域排名只是基于一项指标，即 QS 声誉调查的结果。因此原 THE-QS 的学科领域排名实际上仅仅是对声誉的主观测量。

排名方法

在对排名方法进行修订时，THE 曾经就排名方法草案公开征询意见。根据排名应当减少对任何单个指标的过分依赖而使用一系列指标，并对这些指标按高等院校的功能和使命进行归类的原则，THE 在 2010 年的排名中使用 13 个独立的表现指标，这超出原 THE-QS 方法中使用的 6 个指标。这些涵盖面广泛的独立指标被分为 4 类，进而组成排名的指标体系。我们对大学表现进行评估的四个关键方面是：研究；技术创新和对经济的贡献；国际化程度；以及一个广义的“机构指标”，它包括教学声誉、收入和师生人数。

针对“研究”，我们采用 5 个独立的指标。其中包括“引用影响力”，即一所机构发表论文的篇均引用率；包括权重较低的对人均科研产出规模的统计，即每名研究人员发表的论文数量；还包括机构的人均科研收入；以及声誉调查中关于科研声誉的调查结果。

“机构指标”包括声誉调查有关教学声誉的调查结果；人均总经费；本科生人数；以及博士学位授予数与本科学位授予数的比值。

2010 年，“技术创新和对经济的贡献”采用来自企业的研究经费的人均值这一指标。今后几年，这一类指标还可以包括与来自企业的研究人员共同发表的论文数量，和雇主对毕业生表现的主观评价调查。

“国际化程度”则继续采用留学生比例和外籍教师比例。将来可能还会包括与国际伙伴合作论文的数量。

面向二十一世纪

我们相信由汤森路透公司支持的 THE “世界大学排名”新方案体现了对原 THE-QS 排名的重大改进。相信这份排名是更准确、平衡、严谨和透明，并反映全球范围内现代研究密集型高校在各方面的优势。我们希望排名能够成为帮助全球高校在不断变化的环境中准确定位的重要工具，而不仅仅只是为了满足好奇心。

世界大学网络计量学排名

Isidro F. Aguillo 和 Noelia Gutiérrez Labajos

西班牙 IPP-CCHS-CSIC 网络计量学实验室

电子邮箱：Isidro.aguillo@cchs.csic.es

背景

20 世纪 90 年代中期，网络计量学实验室（Cybermetrics Lab）作为西班牙最大的公共研究机构国家研究委员会（Consejo Superior de Investigaciones Cientificas）下属的研究小组，开始用文献计量学和科学计量学的方法开发网络定量指标。网络计量学实验室的研究对象主要是学术机构，由于这些机构信息的网络化程度还比较低，并且在那时比较常用的网络计量学指标——网络影响因子（web impact factor）存在一定的缺陷，因此研究进展不大。1999 年，网络计量学实验室开始使用检索器收集数据，但考虑到用检索器需要耗费大量的人力和计算机资源，最后决定采用商业搜索引擎来进行数据收集。

2003 年，上海交通大学发布的世界大学多指标综合排名为开发基于网络计量学指标的排名提供了范式。2004 年，网络计量学排名（www.webometrics.info）诞生，它把促进科学知识的公共传播作为明确目标，即所谓的“开放获取（Open Access）”计划。网络计量学排名作为一项研究，过去几年方法一直在不断改进，并且以后也将继续改进。尽管方法的稳定性并不是追求的目标，但 2006 年以来排名方法已基本稳定。

排名假设和模型

互联网正成为最重要的学术交流工具，人们通过互联网能获得越来越多的科学资讯。电子出版物价格便宜、获取便捷，并能在全世界拥有庞大的受众。互联网不仅可以用来发布正式论文，同时也是展示学术机构的组织结构、活动和成果的舞台。网络计量学排名的假设是：大学的网络显示度能够反映它们的学术和服务质量、成果的影响力和国际声誉。如今，大学在“开放获取”方面的努力已成为大学表现的重要内容，也是设计网络计量学排名时的考虑因素。

若假设成立，我们就可以用网络计量学指标对大学进行描述与评估，为世界范围内的科研与高等教育政策制定提供一个新的工具。尽管网络显示度还未成为一些大学优先考虑的事，该排名也在为推动这些大学提高网络显示度发挥积极作用。

网络计量学排名的编制遵循以下原则：（1）网络计量学排名是一个名次表。排名使用多个定量变量加权获得的综合指标对大学的表现提供简单、实时和整体性的描述。（2）网络计量学排名只采用一个综合指标进行排名。我们没有找到其他更好的可行方案，今后需要继续研究各个变量对综合指标的影响与贡献。（3）网络计量学排名是一个全球性排名，目标是对包括发展中国家在内的全世界近 20,000 所大学进行排名。（4）网络计量学排名考虑大学的所有使命。“研究表现”对世界一流大学来说是一类重要指标，而对于多数大学，其他使命也很重要。“教学”，特别是课外学习和远程教育，可以通过对网络

行为的测量间接得出，而所谓的第三使命“知识转移”，也可以通过链接分析得出。

定性分析适用于小样本情况，但即使在这种情况下，对比分析（比如调查）也带有倾向性。定量分析对于分析封闭大环境，比如以研究为基础的学术环境非常有效（文献计量学），但显然网络空间是一个更为复杂的环境。举例来说，链接的数量更大、动机更加复杂，有时还会出现意料之外的使用者。实际上，我们面临的数据还要多得多（以百万计），但是根据大数定理，即使在这样复杂的情况下，也有可能出现有价值的规律。这对排名来说就显得特别有意义。

排名构建

世界大学网络域名目录由 20,000 多条记录构成，对它的编译、整理和更新花费了好几年时间。并不是所有大学只有一个域名，一些大学有几个主域名，有些甚至在使用一个新域名后，原来的域名仍在使用并且比新域名具有更高的显示度。由域名所反映出的独立的高等院校数量共有约 18,000 所，每年还会新增几十所，并且有几百所学校会更换网址。

由于样本数量巨大、涉及的变量众多，我们选择了一种间接的数据获取方法：使用主流商业搜索引擎提取数据。商业搜索引擎通常具有庞大且更新及时的数据库，能够方便地获取网络计量学指标。为避免各个搜索引擎由于覆盖面不同所带来的结果不一致和其他差异，我们同时使用了几个搜索引擎，并在很短的时间内从每个搜索引擎提取两次数据从而降低误差。

网络计量学排名中使用的综合指标借鉴了期刊影响因子（journal impact factor）的计算方式，即被引用次数（显示度）除以论文数（规模）。在互联网中，“显示度”可以由外部链接（被其他网站引用）数量来反映，而“规模”是指网页的数量。由于这些变量呈现负指数分布，我们选择了一个规模与显示度各占 50%的权重分配模式。现行的网络计量学排名的计算方法如下：（1）规模：由以下指标加权组合得出：网页数（20%），即通过 Google、Yahoo、Bing 和 Exalead 搜得的网页数量（进行对数标准化后取中值）；丰富文件数（15%），即用同样方法搜得的丰富文件（包括 pdf、doc、ps 和 ppt 格式）数量；论文数（15%），即被 Google Scholar 收录的论文数量。（2）显示度：占 50%权重，统计 Yahoo 和 Exalead 显示的外部链接数量。现行方法对被学术网站引用的链接给予额外分值。

网络计量学排名一直以来都公布每所大学在每个指标上的名次，并据此计算每所学校的综合指标得分。2009 年我们开始考虑一种新的方法，即每所学校综合指标得分的计算是基于该校在各个指标上的名次被标准化后的值，该方法到目前为止尚未对外公布。为避免相同分值的情况大幅增加，我们正在研究比较几种可能的处理方式。

我们主观决定对指标赋予相同的权重，但同时也尝试把非正式和正式流量都考虑进去。如此一来至少有两个指标有待改进：显示度应既考虑来自非学术网站的外部链接数，又考虑来自学术网站的外部链接数，并且后者应被赋予更大权重，但这种做法从技术上来说很困难，因为牵涉到庞大矩阵的计算。Google Scholar 的 beta 版本现仍有许多问题，比如覆盖的偏向性、无法还原全文本或特殊格式，以及按“年”或“学科”进行搜索得到的结果偏差。如果这些问题在近期能得到解决，那么原定的权重分配模式就应被

调整。

这一排名方法旨在从整体上反映大学的三个使命，尽管这里面存在一些显而易见的缺陷：确定每项使命的分别贡献并不容易；网络命名中的“坏习惯”会对结果产生巨大的负面影响；此外，这种排名方法存在明显的学科偏向，尽管它对被传统文献计量学低估的学科（如工程、计算机、社会科学和人文学科）还算有利。

排名结果

对排名结果的初步分析显示，该排名与其他排名具有宏观相似性。导致单个学校在不同排名中名次差异的原因有：排名对象的界定（是否包括大学医院）、技术误差以及较为严重的方法缺陷（如问卷调查和权重分配）；网络计量学方面的原因有：大学信息网络化程度不一和一些有问题的网络域名使用方法（有多个不同的域名、废弃域名仍在使用、共享域名）。

然而，排名结果最重要的意义是反映了学术数字鸿沟的存在。在 2010 年 1 月的排名中，前 100 名大学在各地区的分布为：北美洲 76、欧洲 17、亚洲 3、大洋洲 2、拉丁美洲 2；前 200 名的情况是：北美洲 114、欧洲 60、亚洲 15、大洋洲 6、拉丁美洲 4、阿拉伯地区 1；前 500 名的情况为：北美洲 200、欧洲 223、亚洲 45、大洋洲 14、拉丁美洲 14、阿拉伯地区 3、非洲 1；前 1000 名：北美洲 370、欧洲 408、亚洲 134、大洋洲 35、拉丁美洲 44、阿拉伯地区 4、非洲 5；前 8000 名：北美洲 2,577、欧洲 2,424、亚洲 2,009、大洋洲 80、拉丁美洲 744、阿拉伯地区 115、非洲 51。前 200 名大学中，北美（美国和加拿大）大学的数量远远超过欧洲大学的数量，这种情况在其他排名中都没有出现过，这显然和大学的网络建设政策有关。欧洲和亚洲大学应认识到必须加强对网络出版物的支持、开发大型“开放获取”资源库和促进信息网络化。

研究密集型大学，即所谓的世界一流大学，在本排名中名次最高，但也有一些具有区域影响力的拉丁美洲和亚洲的大学取得了好名次。

结束语

网络指标容易获取，并且能为人们提供认识全世界学术机构（包括来自发展中国家和金砖四国的学术机构）的国际视角。因此网络计量学排名是学生、学者、研究人员和政策制定者的一个很有用的工具。但是，每个指标的作用和意义等相关问题还需要进一步研究。

排名使用的网络计量学指标虽然与规模有关，但是与网站的设计或访问量无关。排名提供了全世界众多大学的网络表现的排名。

下一步研究计划包括：对机构进行分类（按公/私立、规模、主要学科或领域），调整指标权重，研究丰富文件的格式（重新评估 ps 和 xls 格式），评估多媒体文件（视频），考虑新的搜索引擎（如本地搜索引擎 Baidu、Yandex 和 Yahoo Japan），以及提高数据的可视化程度。其他计划包括：合并大学排名与研究机构排名，开发学科与研究团队排名，以及结合文献计量学指标的排名。

网络计量学排名由以下网站组成：大学排名（www.webometrics.info），研究机构排名（research.webometrics.info），医院排名（hospitals.webometrics.info），资源库排名（repositories.webometrics.info）和商学院排名（business-schools.webometrics.info）。

台湾高等教育评鉴中心之大学排名对台湾高等教育质量提升的影响

侯永琪

辅仁大学教师发展与教学资源中心主任、财团法人高等教育评鉴中心基金会国际交流组组长

电子邮箱：035440@mail.fju.edu.tw

高等教育质量保证与追求卓越—作为排名者的高等教育评鉴中心

以往大部分的大学排名多由媒体发布，例如《美国新闻与世界报导》的排名，但如今，进行大学排名的机构已经相当多元化了。有些是由大学自身产出的排名，例如上海交通大学的“世界大学学术排名”；有些则是政府的高等教育认证机构所发展出的大学排名。在某些国家中，排名本身就是认证过程的一部分，由认证机构或者主管高等教育的政府部门来进行排名。台湾财团法人高等教育评鉴中心基金会（以下简称评鉴中心，HEEACT），即属于后一类中最为成功的例子之一。

2005 年修正并实施的《大学法》规定，所有台湾的大学校院都必须定期地接受评鉴，由教育部委托一家独立的评鉴机构来实施具有一定程序和标准的认证。根据此法，评鉴中心于 2005 年成立，并开始对 76 所四年制综合大学进行评鉴实践。在此同时，作为台湾高等教育的质量保证机构，政府还希望评鉴中心能够提供国际竞争相关的数据与信息，作为政府制订高等教育政策的参考。因此，从 2007 年起，评鉴中心开始进行各类大学排名系统的研发，包含“台湾大学校院科学期刊论文质量统计分析”、“台湾大学校院 ESI 论文统计”、“世界大学科研论文质量评比”、“专利计量分析”、“大学产学合作绩效评量”，以及“大学选校互动系统”等。评鉴中心研发这些大学校院排名，是为了在当今激烈的国际中，协助各校发展自身的竞争力，找到自己的定位。

评鉴中心的全球排名—“世界大学科研论文质量评比”

评鉴中心于 2007 年推出的“世界大学科研论文质量评比”，旨在通过分析近年来世界前 500 名的大学在科研论文方面的表现，了解台湾的大学与其他国家和地区大学之间的差距所在。此外，评鉴中心还希望这个全球大学排名能够在如何建设有影响的研究型大学方面，为规模较小的新发展国家提供洞见。但评鉴中心清楚地指出，这个排名仅仅呈现了大学在某些方面的研究表现，绝不代表大学全面的行政、教学与服务的质量。

“世界大学科研论文质量评比”采用了 SCI 与 SSCI 数据库的数据来评估大学的研究绩效，因为评鉴中心认为在同行评议的国际期刊上发表论文是最主要的科研产出方式，因此，发表在此类期刊上的论文数量是反映大学科研表现的重要指标。2007 年，首度公布的排名共采用 9 个指标，分别代表了科研论文表现的三大方面：学术生产力、学术影响力、学术卓越性。2008 年，评鉴中心研发了新的领域排名，因此将学术卓越性方面的“卓越领域数”指标删除。目前，评鉴中心共采用 8 个指标来评估大学的科研产出表

现，分别是近 11 年论文数、当年论文数、近 11 年论文被引次数、近 2 年论文被引次数、近 11 年论文平均被引次数、近 2 年 H 指数、高被引文数，以及当年发表在高影响期刊上的论文数。

评鉴中心陈振远执行长就指出，“世界大学科研论文质量评比”所采用的客观指标，充分考虑了每所大学长期与短期科研表现的差异性。定量与定性并重的 SCI 与 SSCI 的数据分析，可使全球大学排名的结果更为公平。而这个排名也将历史因素考虑进去，以使大学间的比较能更为客观。此外，评鉴中心在计算得分时还采用了平均值，以避免学校规模影响排名结果。

2008 年，评鉴中心又发表了另一份考虑大学规模的排名，以更为缩小大学规模对于排名结果的影响程度，提供更为公平的排名结果。同年，评鉴中心还发布了全球大学领域排名，包括各领域前 300 名的大学名单。六大领域分别是农学、临床医学、工学、生命科学、理学与社会科学。

根据评鉴中心于 2009 年发表的排名结果，美国与欧洲的大学在国际高等教育界中，依然居于龙头的地位。尤其是美国的大学，占居了前十名的位置；而亚太地区的大学仅有 2 所进入前 30 名，且都是日本的大学。

事实上，由于“世界大学科研论文质量评比”采用的方式十分科学，排名的对象也为全球大学，其在全球已渐渐成为颇具影响力的全球大学排名之一。现今在每年发布日，其网站皆有 3,000 人次以上的浏览量。该排名系统，不但吸引了国际的目光，也给想发展世界一流研究型大学的国家以重要参考。

大学选校互动系统

由评鉴中心开发的“大学选校互动系统”，是另一项非常创新的排名研究计划。虽然评鉴中心每年都将认证结果与认证报告上网公告，但因信息过多，而导致社会大众不易充分地了解。因此，作为大学校院的市场代理人（marketing proxy），又肩负台湾高等教育国际化功能的评鉴中心，于 2009 年 10 月正式启用了“大学选校互动系统”。

与传统的大学排名相比，“大学选校互动系统”给予用户指标选择与权重决定的自主权力。使用者根据自己的判断，可以选择 3 至 10 个指标，并且自由分配每个指标的权重。此外，使用者还可以进一步设定比较的范围，比如区域、类型、规模与学术领域。而用户关心但却没有列在指标部分的信息，则在最后的结果部分呈现出来。这些信息包含创校年份、学校使命、注册人数、系所数目、学校网站、评鉴结果、经费补助、入学申请、校友表现、学费、学生社团等。

“大学选校互动系统”以用户为导向进行建置，社会大众不该以传统大学排名的观点来检视此系统。此系统最重要的特色在于为使用者提供指标与权重的自主权，使得每位使用者能根据自己的选择，建立属于个性化的大学排名清单。另外，此网站也呈现了相当丰富的信息（以上、中、下的分群方式呈现结果）。此方式与德国高等教育发展中心排名采取的呈现方式相同，不将大学各项指标进行尖锐又精确的名次排序。

“大学选校互动系统”现今已经成为台湾高中毕业生及国际学生了解台湾大学校院的便利工具。然而，这些使用者都希望在可选择指标的数目上能有更多的弹性，并且提供更透明的大学校院信息。评鉴中心也在 2010 年，持续更新原始资料，并将近 80 所的技职校院也纳入此系统。作为长远目标，评鉴中心将会邀请亚太地区的大学也加入这

个系统。可以预见，此系统所提供客制化且丰富的信息，将使得越来越多本地与国际的学生受惠。

评鉴中心的排名与世界一流大学建设

21 世纪的全球化潮流，使得各国和地区、大学面临了各种挑战，也带来了各种机会。近年来，各国和地区、大学都十分关注高等教育质量保证这一议题，并实施各种政策，以强化全球的竞争力。因此，许多国家和地区都通过大学排名作为建立世界一流大学的基础，台湾也不例外。

为了回应这个潮流，台湾政府也挹注 5 年 500 亿，实施了建设世界一流大学及顶尖研究中心的计划。这个计划的目的在于协助大学发展，在 5 年内让至少一所大学进入世界前 100 名，并至少有 15 个重点系所或是跨校研究中心跃升到亚洲顶尖地位。这项卓越计划也与其他政策有所关连，例如为提升台湾高等教育的国际化水平，改革整体高等教育系统结构的政策。更为重要的是，这项计划展示了台湾大学希望同各国高等教育机构竞争，晋升为世界一流大学的雄心。为了达到这个理想，第一期卓越计划一共有 11 所研究型大学接受资助，它们必须在计划执行过程中，完成五阶段的评估，才能再继续下一期的计划资助。同样地，教育部也持续关注这些接受资助的大学在评鉴中心全球大学排名中的表现。

在评鉴中心发布的全球大学排名中，2008 年仅有 5 所台湾大学进入前 500 名。与此相比，2009 年共有 11 所台湾研究型大学进入前 500 名，包括国立台湾大学（102 名）、国立成功大学（307 名）、国立清华大学（347 名）、国立交通大学（456 名）、长庚大学（479 名）、国立中央大学（483 名）与国立阳明大学（493 名），占亚太地区前 500 名的 8%，且这几所大学皆是接受卓越计划资助的大学。

虽然每年评鉴中心的全球大学排名结果显示出，台湾的一些大学已具有世界一流大学的潜质，学术上的表现也越来越亮眼，并受到世界的肯定。但评鉴中心的全球大学排名也引发了台湾部分大学校长，甚至是评鉴中心董事针对其方法与目的的严厉批评。而一些没有进入前 500 名的大学也担心，评鉴中心的排名可能会成为审核进入第二阶段卓越计划大学的重要遴选标准。另一方面，也有人质疑这些卓越计划所资助的研究型大学的教学质量，因为这些大学当中有一些系所在评鉴中心第一轮的评鉴结果是属于“待观察”，需来年再接受评鉴中心的再评鉴。

结束语

评鉴中心的陈振远执行长解释了认证与排名的不同功能：“有别于仅考虑科研产出的全球大学排名，评鉴中心的认证十分重视教学。我希望大学不要错用或混用这两个工具，以免受到误导”。然而，相当吊诡的是，某些研究型大学，在提供学生良好的学习环境方面已承受颇高的压力，被要求增进教学质量，但有些教学卓越大学却已将未来的愿景设定为“进入前 500 名”。不可避免地，评鉴中心的认证标准与排名指标将被越来越多的台湾高等教育机构视为自我提升的工具，并相当程度地影响各大学中长期政策的制定。现今，高等教育的质量与卓越，在台湾是个重要的议题。全球化的潮流已深刻影响了台湾高等教育生态，而来自其他国家大学的竞争与绩效导向文化，也将会更为强化高等教育排名与评鉴对于大学的重要性。

德国 CHE 的大学排名

Gero Federkeil

德国高等教育发展中心项目主管

电子邮箱：gero.federkeil@che-ranking.de

引言

在过去的二十年中，世界上许多国家都出现了大学排名。尽管历史已不算短（1983 年《美国新闻与世界报道》发布首份排名），大学排名仍备受争议，尤其在高等院校内部。亚历•亚瑟（Alex Usher）和马西莫•萨维诺（Massimo Savino）指出，“无论排名在哪里出现，都受到了公众的密切关注，并使高校感到不安”。大学排名在竞争激烈的环境中增加了高等教育体系的透明性，为有意接受高等教育的学生、家长，以及雇主提供了参考。它既是竞争的媒介又呈现了竞争的结果，可以说是知识经济社会的内在要求。这意味着排名在测量的同时也在重塑竞争结构。由于大学排名构建了关于高等院校的质量等级，并且拥有较高的社会显示度，这可能会对个别院校的市场行情产生影响，比如申请者的态度，所以会受到高等院校的密切关注甚至是质疑。

德国的第一份大学排名是在 1989 年由一份周刊发布的。随后几年也出现了一些排名，大部分由杂志发布。1998 年，高等教育发展中心（简称 CHE）发布首份大学排名。CHE 是一家私营、独立的非营利性机构，由德国大学校长协会和一家私营企业基金会共同创办。发展基于科学方法的大学排名是 CHE 创办的目标之一。大学校长协会的一个研究项目曾经开发了一系列描述大学情况的指标，CHE 排名就是在这个基础上做出来的。

CHE 排名：目的与发展历程

CHE 排名的准备经过了为期两年的与评估、方法论专家以及学生的反复讨论，后者主要是就期望从排名中获得怎样的信息给出意见。排名的主要目的是为潜在的学生提供有关大学和专业的信息，帮助他们做出合理的选择。此外，排名还为大学提供关于学校表现的比较信息，包括的指标有：教学、资源和设施、科研活动，以及一些重要但与学校表现无关的信息（例如所在地房租水平、大学规模等）。指标中包括科研活动的原因有两个：首先，对小部分潜在学生来说，关于科研活动和表现的信息对他们选择大学至关重要；其次，高校本身不太会接受不考虑科研表现的大学排名。由于科研不是 CHE 大学排名（主要针对潜在学生）的重点，CHE 决定发布一份专门的“科研排名”，为想要了解科研表现的人群提供详细信息。在这份排名中，科研数据被更为细致地分析和公布。除了公布出版物数量、被引次数、博士人数、科研经费和专利等单个指标外，排名还包括了对指标之间相互关系的统计分析。

排名覆盖的学科数量一直在增加，到目前为止共 32 个。学科的选择主要基于学生的数量，通常是那些在德国拥有最多学生的学科。现在，在被排名学科学习的学生占德国学生总数的 80%。对更多的学科进行排名，

尤其是对小众学科进行排名有诸多局限：首先，学科的扩展无论是对 CHE 还是对需要为排名提供数据的大学来说，都需要投入大量的精力和资源。其次，开设小众学科的学校数量很少（少于 15 所），因此对它们进行排名意义不大。在博洛尼亚进程的大背景下，人们越来越想知道有关高校和专业的国际可比信息，CHE 自 2004 年开始，在大学排名中逐步增加奥地利、瑞士和荷兰的大学。

排名的理念

CHE 排名有三方面的主要特征：

排名对象是专业/学科而非大学整体：与大多数排名不同的是，CHE 排名是基于学科的排名。来自 CHE 排名的数据显示，大学各科系的表现差异可能会非常大。一所大学可能在物理学科排名中名次较高，但是在历史学科排名中名次较低。作为 CHE 排名主要目标群体的潜在学生想要了解的是他们准备学习的学科的信息，而大学各个学科的平均表现参考价值不大。举例来说，如果一所大学在整体排名中处于中游水平但是在物理学科排名中名次靠后，那么对想要学习物理的潜在学生来说，实际上整体排名并没有什么意义。

多维排名而非综合得分：各个大学排名的指标数不尽相同，但大多数排名会对各个指标赋予具体的权重，进而计算总得分。通过选择一组特定的指标并赋予特定的权重，排名实际上强行定义了大学的质量是什么。但是根据美国国家民意研究中心（U.S National Opinion Research Center）的研究，人们在赋予指标权重时既没有理论上的依据也没有经验支持。在制作排名的过程中，我们认为必须要考虑排名用户/目标群体的不同决策偏好。有的学生在寻找科研活跃的大学（通过科研经费、出版物数量等来测量），还有的学生可能在寻找师生关系密切或学时短的大学。因此，计算总得分并不能满足不同的需求。

此外，计算大学整体得分模糊了大学或专业在某一具体方面的表现差异。这在同时包括教学指标和科研指标的大学排名中显得尤为明显。一所科研表现良好的大学未必能为学生提供良好的教学，反之亦然。多维排名可以为分析大学的优势和弱势提供不同的视角，这也是体现大学质量多元性的唯一途径。亚瑟和萨维诺认为：“大学对排名感到不安的主要原因之一是对指标赋予权重的做法越来越流行，最后只给出一个涵盖一切的质量总得分”。

分组而非具体的名次表：自《美国新闻与世界报道》发布排名以来，大多数排名都给出了大学的具体名次。这意味着指标数值的任何差别都代表了大学在质量上的差异。名次表不可避免地会带来对指标数值微小差异的误读。在许多情况下，数据的精准度不足以给出清晰明确的大学名次。用统计学的说法，这忽略了数据的标准误差。因此，CHE 排名针对每个不同的指标，将大学归入三个质量组别：前列、中等和靠后。同一组别内的大学不再加以区别，而是按字母顺序排列，所以不存在名次表。

指标

指标选择对大学排名至关重要。我们可以根据数据来源和指标质量（相关性、有效性）来辨别大学排名。指标应该与排名的目标群体有关。在近两年的准备期间，CHE 得到了咨询委员会（包括评估专家和专业机构、大学协会的成员）的帮助，并通过与高中毕业生和大学生的讨论，从而确定相关的指标。

经过这一过程，我们获得了一个含有 9 大类决策相关指标的排名模型。每一大类指标都包含若干个二级指标，总共约有 20～30 个二级指标（各学科有一定差异）。大类指标涉及城镇概况（如所在地平均租金）、学校概况（规模、创建年份、类型）、学生特征、课程与教学情况、就业情况、科研情况，以及教授和学生的综合评价。9 大类指标及其二级指标分别为：

学生：申请录取率。

学习成果：在规定时间内取得学位的学生比例；取得学位的平均时间；毕业/退学比。

国际化：专业国际化程度、留学生（学位生和交换生）、教师国际化程度、用外语教学的课程等指标；对海外学习的支持（学生调查数据）。

科研：人均外部科研经费；师均博士生数；人均论文发表数；论文篇均被引次数；科研声誉（教授调查数据）。

教学：平均取得学位的时间/规定时间内取得学位的学生比例；生师比；教学质量（学生调查数据）；专业的组织形式（学生调查数据）；教师与学生的联系（学生调查数据）；教师对学生的支持（学生调查数据）。

设施：校舍情况等（学生调查数据）；图书馆（学生调查数据）；实验室（学生调查数据）。

就业：就业指导以及实践课程（学生调查数据、毕业生调查数据）；是否设立就业服务机构；就业指导（学生调查数据、毕业生调查数据）；问题解决能力的传授（毕业生调查数据）；知识迁移能力的传授（毕业生调查数据）；学习方法的传授（毕业生调查数据）；团队协作能力的传授（毕业生调查数据）。

总体评价：学生总体满意度（学生调查数据）；学校声誉（教授调查数据）。

学校和城市：学校规模；学生在当地的膳宿费用；学生宿舍的床位数。

除了以上这些指标，CHE 排名还包括大量关于大学的描述性信息。

结果发布

为了广泛宣传排名结果，CHE 与一家媒体合作共同发布排名。排名结果既公布在它们的定期出版物上，又出现在一份独立的学生择校指南和互联网上。

网络版大学排名最重要的特征是提供了一个交互式、个性化的排名（称为“我的排名”），即根据多维排名的基本方法，不赋予指标权重，使用者可以根据自己的偏好选择若干个指标（最多五个），并在每个指标上分别限定他希望了解的学校范围（前列组，中等以上组，或是全部），最终得到符合个人需求的大学排名。由于这样的个性化排名因指标选择的不同而结果差异巨大，因此反映了大学各个学科的优势、劣势和特色。

结论

CHE 排名开发的这种独特排名方法在许多有关排名的比较研究中得到了很高的评价。这个方法不同于其他主流排名，它以学科/专业而不是大学整体为对象进行排名，它是一个多维排名从而避免了简单计算综合得分，它把大学分成三组而不是给出具体名次因此避免了不切实际地夸大大学表现的差异。这样的方法同网络交互的结合而得到的个性化排名提供了关于大学优势和弱势的详细信息，因此既能帮助潜在学生找到最适合自己的大学，也有助于被排名单位以及研究人员进行学校之间的对比。

波兰《远景》排名：改进的艺术

Kazimierz Bilanow 和 Waldemar Siwinski

Kazimierz Bilanow：波兰远景基金会国际事务主管；

电子邮箱：k.bilanow@perspektywy.pl

Waldemar Siwinski：波兰远景基金会创始人、主席

电子邮箱：w.siwinski@perspektywy.pl

汽车自发明以来就都配备了底盘、发动机和车轮，然而今天已没有一个部件和最初的一模一样。尽管他们的基本功能完全相同，但都经历了巨大的改进。1950 年最棒的汽车肯定不会是 2010 年最棒的。同样地，波兰《远景》（Perspektywy）排名创始人、具有工程教育背景的斯文斯基先生（Waldemar Siwinski）明白，任何事物都必须因科技进步和人们预期的改变而不断更新。《远景》排名在波兰高等教育市场的优势地位就归功于长期坚持良好的操作和不断进行改进。

源起

波兰教育排名的起源可以追溯到 1992 年《远景》杂志发布的首份华沙中学排名。这个排名一经发布就获得了成功，所有杂志迅速销售一空。随后，《远景》的中学排名从华沙地区扩展到整个波兰范围，几年后《远景》决定开始从事波兰大学的排名。

20 世纪 90 年代初期，波兰高等教育的发展状况使得人们对这种排名有了需求。随着政治体制的变化和人口的增长，波兰的教育特别是高等教育经历了前所未有的大发展。出生人口数从 1982 年的 350,000 人增长到 1984 年的 720,000 人。中学生数量的第一次高峰出现在 90 年代后期，强烈地冲击着原有的高等院校。波兰的高等教育经历了快速的扩张，这种扩张从世界范围来看都是罕见的。15 年内波兰学生的数量从 38 万增长到了 200 万。为了满足需求，波兰新建了 35 所公立大学和 380 所私立大学。

尽管此前的中学排名被人们普遍接受和欢迎，但众所周知学术排名是一个更具挑战性的领域，需要很多前期准备工作。《远景》花了整整两年进行准备。首先我们与主要大学的校长们和大学校长协会的领导者们举行了一系列会议，详细讨论排名的基本原则和开展方式。这些会议使我们意识到大学希望《远景》严格遵守各方共同议定的基本原则来从事排名活动。为保证排名的公正、独立和透明，我们成立了非营利性的《远景》教育基金会，负责选定排名标准、指标和所有其他排名相关事宜，《远景》出版社负责排名的发布以及和市场有关的工作。

排名指标与方法

虽然排名需要尽可能做到中立和客观，但过程中仍会涉及一些主观因素，例如评价标准与权重的选择。《远景》选择了一批专家学者来决定评价标准，一个名为 Kapitula（来自拉丁语 Capitulum）的委员会在波兰成立，由声誉卓著的学者（前大学校长）等社会人士组成。他们的知识和经验对选择合适的排名方法很有价值，因此提高了排名结果的可

信度。

创始人在工程领域的经验表明，作为一种测量工具的排名必须被进行正确的定标。为了确保第一份大学排名不偏离当前民意太远，我们决定采用比较保险的方式，对“声誉”指标赋予较高的权重（50%）。排名进行了两项调查，其中一项的调查对象为学术人员，另一项的调查对象为雇主。但我们从排名开始时已做出决定，此后几年“声誉调查”和“同行评价”的权重会逐步减少。

由于《远景》认为透明性对排名很重要，所以杂志的首份大学排名不仅发布了结果，而且详细地公布了整个排名制作的过程，其中包括一份参与同行评价调查的约 1,000 名教授的名单。

《远景》的首份大学排名使用了 15 个指标，涉及三大类，分别是：声誉（50%），科研实力（30%）和学习环境（20%）（见“Perspektywy – Ten Years of Rankings”，Higher Education in Europe, Volume 27, No. 4, 2002）。自 2000 年发布首份排名以来，十年中《远景》对排名进行了缓慢、持续的改进，对指标、数据来源等进行了调整，以便提供更准确的波兰高等教育图景。2005 年，《远景》排名引入了一类新指标，即“国际化”，占 10%的权重。同时，“声誉”的权重下降到 30%，“科研实力”上升到 40%。在 2010 年的大学排名中，另一类指标“创新”被引入指标体系，占 5%的权重，而“声誉”的权重减少到 25%。

“软”数据和“硬”数据

对各大类指标的权重进行调整以及引入新的指标类别导致了指标数量的增加，指标总数从首份排名中使用的 15 个增加到 2010 年排名中使用的 32 个。这一方面反映了波兰高等教育的变化；另一方面，也是因为可获取的并且可靠的数据越来越多，比如 2005 年引入的“国际化”指标和 2009 年引入的基于 SCOPUS 数据库的文献计量学指标。

对《远景》排名的改进，呈现出基于“软”数据（通过问卷调查获得的数据）的指标比重下降，基于“硬”数据（来源于电子文献与引文数据库）的指标比重上升的趋势。即使是像“雇主调查”这样的“软”数据也由独立公司潘道尔调查机构（Pentor Research International）通过开展专业的调查获得。在引入新指标和采用新数据的同时，《远景》也一直注重保持排名方法的连续性，使得不同年份的排名结果之间可以进行比较。具体排名指标及权重的变化，请参见《远景》排名的网站（www.perspektywy.org）。

排名涉及的数学计算非常简单，大多基于多指标分析法（MCA），而收集计算所需的数据则困难得多。《远景》排名主要借鉴的是由鲍勃•摩斯（Bob Morse）及其团队开发的《美国新闻与世界报道》的排名。

分类排名

鉴于波兰公、私立高等院校的多样性和不同的使命、目标，除了对所有大学进行的综合排名外，《远景》同时提供了不同类别高等院校的分类排名。事实上，2010 年的《远景》大学排名就不只是一个综合排名，而是由 5 个子排名组成，分别是：学术型院校（包括所有有博士学位授予权的高等院校，艺术类院校除外）排名；有硕士学位授予权的私立高等院校排名；有学士学位授予权的私立高等院校排名；公立高等职业学校排名和学科排名。

以上所有子排名采用相同的 5 类指标，分别是：“声誉”、“科研实力”、“学习环境”、

“创新”和“国际化”。每类指标在不同的子排名中被赋予不同的权重。2010 年的《远景》大学排名中，华沙大学（University of Warsaw）获得了波兰最佳大学的称号，科兹明斯基学院（Kozminski Academy）成为波兰私立大学中的第一名。

《远景》排名不仅是在报告波兰高等院校的客观情况，也为提高波兰高等教育的质量做出了贡献。目前波兰的高等教育在国际化方面远远落后于欧洲其他国家。自 2004 年“国际化”（开始时被赋予 10%的权重，随后增加到 15%）指标被首次引入《远景》排名以来，英语授课专业的数量就一直在增加，在波兰学习的外国留学生人数也出现了显著增长。由于高等院校很关注自己在《远景》排名中的位置，于是紧跟排名指标的变化并尽可能瞄准这些指标，从而提高它们在排名中的名次。由于“创新”指标最近才刚被引入排名体系，要看到它的影响还需要一些时间。

良好的操作获得认可

《远景》大学排名连续发布了 11 年，在这期间波兰也出现过其他排名，但没有一个能维持很长时间。为何只有《远景》取得了成功而其他都失败了呢？

我们认为《远景》排名的成功源于对良好操作原则的坚持，这些原则可以在《高等院校排名的柏林原则》中找到根据：保持排名方法的透明性和经常向利益相关者（包括大学校长和学生）咨询；由知名学者和杰出社会人士组成独立的排名委员会（Kapitula），对排名过程进行监督；持续改进排名方法，以确保排名能够反映高等教育形势的变化；关注高等院校在使命上的差异，根据不同类别高校的特征发布多个子排名；公布排名的详细数据，提供对排名指标及其优缺点的分析和评论；积极发展那些可以把排名结果传递给公众的合作伙伴（例如“Rzeczpospolita”是波兰一家主要日报，该日报被授权在国内发布排名结果）。由于坚持了这些原则，《远景》排名获得了公众和学术界的广泛认可。

《远景》团队认识到，波兰的高等教育是欧洲和全球高等教育的一部分。为了与世界各国的排名保持同步，《远景》教育基金会积极参与了国际排名专家组，作为秘书机构服务于“IREG－学术排名与卓越国际协会”。

日本《朝日新闻》的大学排名

Testuo Kobayashi
日本《朝日新闻》出版公司大学排名主管
电子邮箱：kobayashi-t7@asahi.com

引言

1994 年，《朝日新闻》出版公司首次发布“大学排名”，至 2010 年共发布了 17 期。

当初决定制作“大学排名”是因为我们想为正在准备大学入学考试的学生提供更多关于日本大学的信息。在日本，高中生一般根据大学的最低录取分数来决定上哪所大学。事实上，日本所有的大学都是按招收学生在标准测试中的得分被排名的，排在前 20 名的大学被认为是最有声望的。申请者也会考虑大学的其他方面，如教育和研究、传统和历史，但一直以来都会把测试得分作为最重要的因素。在这种情况下，我们决定换一个角度制作一份不同的“排名”，为高中生提供更多方面的信息。例如，按每所大学教员发表的学术论文数量、被引用次数、日本政府资助的研究经费数量对大学进行排名。

然而，在 20 世纪 90 年代初期，我们的想法并没有被公众接受。那时人们不赞同排名的想法和做法。所以当我们在刚起步的时候，大学里的人不愿意为我们提供信息。但是形势发生了变化。在过去的十年中，日本的大学面临着一项巨大挑战，即人们期望它们在社会中发挥更大的作用。针对大学究竟为社会做了多少贡献的评估变得越来越多。人们批评大学的教育和研究未能跟上社会的发展，比如与企业界联系不够、基础研究薄弱、诺贝尔奖级的顶尖科学家比较少，以及毕业生与社会需求存在差距。为了解决这些问题，政府、大学、企业和所有关注日本大学未来的机构和个人都认为应当实施一项新政策，即引入竞争机制从而振兴大学的教育和研究。

在这样的背景下，“大学排名”的作用也发生了巨大变化。一开始我们主要为高中生提供各种信息，而现在我们把更多的基础数据提供给大学进行自我评估。这些基础数据包括：大学产出的论文数、被引用次数、日本政府资助的研究经费数、对社会的贡献。“大学排名”提供了许多个角度，比如，可以比较 A 大学和 B 大学在“物理学科的被引次数”。

排名的内容

我们的“大学排名”基于以下三方面：教育、研究、对社会的贡献。它们都是大学的使命。

教育。就教师和教学设施而言，排名涉及的指标有：生师比，生均电脑数，生均图书，食堂和餐厅的质量，以及针对残障学生的设施和服务，如电梯、斜坡、盲文书、文件放大器、助听器等。

就教育本身而言，我们针对学生满意度进行问卷调查。同时，我们还调查各所高中会推荐学生就读哪些大学。此外，我们还问卷调查了 300 家公司的人力资源部门，了解哪所大学的毕业生拥有良好的专业知识、销售管理技能和战略计划的能力。

研究。我们选择论文数、被引次数、科

研经费数，以及大学和公司合作研究的项目数等指标来进行排名。我们特别关注合作研究项目，因为日本政府大力倡导此类合作研究。

对社会的贡献。什么是学者对社会的贡献？我们认为学者应以他们的专业知识和眼光，公开提出自己的观点以启发和引导人们。为了体现这一点，我们在排名中使用了媒体曝光率、对社会开放的研讨会的频率、取得的专利数量和政府委员会成员数。我们还收集了关于学者多长时间为报纸、杂志和各种出版物撰稿的数据。

我们认为媒体曝光率和开放性研讨会的数量很重要，因为这显示了学者对启发社会所作的贡献。举例来说，通过对美国“9.11”恐怖袭击和轰炸阿富汗等热点事件的历史与政治背景进行解读，可以使人们对这些事件有更深刻的认识。

政府委员会中的学术成员数也是衡量社会贡献的一个有用指标。日本有许多不同类型的政府委员会在讨论新的政策方向。统计政府委员会中的学术成员数有利于评估大学的学者凭专业知识为构建美好社会所作的贡献。然而，我们也意识到一些委员会成员并没有履行职责，而仅仅是官僚化运作的摆设。

其他排名的角度包括学费、学术人员平均年龄、总经费、教员工资、著名校友（如国会议员、著名运动员和公司总裁）。此外，我们还对入学考试的成绩、大学网站和大学宣传册进行排名。

问题与展望

我们在制作排名的过程中遇到的最大挑战是建立标准的排名方法，以此评估日本的大学。日本大学的规模差异很大，一些规模很大的学校几乎涵盖了所有学术领域，而也有一些是小型的专业学院。由于各个大学的学生数、教职人员数和办学条件差异相当大，因此很难按同一个标准进行评估，这就是我们不做综合排名的原因。

实际上，对高等教育质量的评估还有待改进。我们一直在寻找能够评价学生培养质量的公平而又客观的方法，当然这是一个很大的挑战。

加拿大《环球邮报》的“大学评级”和“大学指南”

Alex Usher

加拿大高等教育战略联盟主席

电子邮箱：ausher@higheredstrategy.com

2002 年，加拿大全国发行的《环球邮报》（the Globe and Mail）首次发布“大学评级”。近十年来，该评级及其相关网络出版物经历了相当大的改进，本文将回顾这一历程。

评级每年发布在《环球邮报》的一份彩色增刊上，最初它被视为对《麦克林》大学排名的回应。《麦克林》大学排名是由周刊《麦克林》杂志发布的首个加拿大大学排名，它的方法和形式与《美国新闻与世界报道》的大学排名十分相似，两者都基于对大学的调查数据（与此相对的是第三方数据），辅以对大学管理层和雇主进行的声誉调查。

相比之下，《环球邮报》的“大学评级”有所不同。《环球邮报》没有采用来自于学校的调查数据以及现成的第三方数据，完全依靠对学生进行的问卷调查。在这些调查中，学生被要求在一系列指标上对大学做出评价，包括教学质量、学生服务质量、课程丰富性、基础设施等。并且，《环球邮报》“大学评级”为了以示区别，称自己是一份“评级（report card）”而不是传统意义上的排名（ranking），学生对各项指标的评价的平均得分被转换为字母等级（如 A，B+等）。和许多欧洲的大学排名一样，评级对各项指标不设权重。然而，在问卷调查中学生会被问到一个关于大学体验的总体满意度的总结性问题，由于这个问题在结果中最先显示，因此它的得分在功能上等同于传统大学排名中的最终加权得分。

《环球邮报》基于学生调查所做的“大学评级”把本科生（和他们的父母）作为直接的市场目标。评级不对大学声誉进行测量，也不采用来自第三方关于科研收入和文献计量学的数据。这一方面反映了他们不愿意从事自己不擅长的数据处理，另一方面也反映了大学的科研活动与学生的在校体验关系不大。显然，使用满意度作为测量指标会对科研密集型大学不利，因为一般情况下各种满意度和大学的规模都是成反比的。因此，尽管所有大学按同一标准被划分了等级，针对一个问题的所有答案也被一同显示，但为了使比较更为公平，最终的结果却是按大学规模进行分类后显示的。

一开始，“大学评级”是《环球邮报》和一家名为 Strategic Counsel 的市场调研公司的合作项目。2005 年，为创建类似德国高等教育发展中心大学排名那样基于网络的排名，高等教育战略联盟（Higher Education Strategy Associates，其前身为美国的教育政策研究所加拿大分部）也加入进来，于 2006 年秋天启动了大学指南（Globe Navigator）项目。这一基于网络的排名有 25 个指标（后来缩减到 17 个），用户可以选择 1～5 个指标对大学进行排名，还能按地域筛选结果。在德国高等教育发展中心排名中，大学并非根据每个指标排列名次，而是被粗略地分成几组（即前 1/3，中间 1/3 和后 1/3）。《环球邮报》大学评级中的大部分指标基于 Strategic

Counsel 公司所做的学生调查，而高等教育战略联盟在大学指南中还加入了一些第三方数据（包括图书馆藏书、文献计量学指标、学费、运动队成绩等数据）。因此，大学指南不完全是《环球邮报》“大学评级”的网络版本，它还提供了额外的数据与分析。

在 2006 年以前，Strategic Counsel 公司依靠一家名为 UThink 的公司开展学生调查。UThink 是 StudentAwards.com 的咨询机构，后者为学生提供在加拿大申请奖学金的信息，以换取学生的电子邮箱地址和基本资料。UThink 向数据库中的学生发送邮件，代表 Strategic Counsel 邀请他们参与网络调查，再由 Strategic Counsel 回收调查结果。通常每年有 1～1.5 万学生参与调查，这意味着有些大学的学生反馈非常少。

随着高等教育战略联盟的加入，因其与大学关系密切，我们采用了一个与以往不同的数据收集方法。和德国高等教育发展中心排名的做法相类似，大学本身也参与进来。原本由 UThink 向学生发送调查邀请，现在则要求大学自己进行，几乎所有大学都同意了。大学同意加入这一计划的原因有很多：第一，大学可以免费获得数据（每个参与其中的大学会收到一份关于本校学生的调查结果）；第二，他们希望能够提高调查的反馈数量和数据质量；第三，越来越多的大学对《麦克林》排名的数据收集方法感到失望（2006 年开始《麦克林》排名遭到一大批学校的抵制，在公开信特别提到了该年 6 月在大学排名国际专家组会议上通过的柏林原则），从而转向其他排名，特别是高等教育战略联盟提倡的德国高等教育发展中心排名方法，这个方法看起来更为准确、合理。数据收集方法的改变导致每年参与调查的学生数量显著上升，现在一年总的反馈数在 3.5～4 万之间。

2006～2009 年期间，《环球邮报》“大学评级”和“大学指南”对名称做了略微调整。《环球邮报》“大学评级”（University Report Card）更名为加拿大“大学评级”（Canadian University Report）；“大学指南”的名称也进行了调整，从 Navigator 更名为 Campus Navigator，成为《环球邮报》在 2008 年秋天启用的专门网站“Globe Campus”的一部分。

2009 年末，Strategic Counsel 公司退出合作项目，高等教育战略联盟接管了学生调查。此时，学生调查出现了一些小变化：最显著的是关于满意度的问题减少了，而关于个体特征（包括学生在大学和中学的学术表现，以及关于个性的若干指标）的问题增加了，还增加了学生对自己大学的看法（例如，“更关注本科生还是更关注研究生”，“课程更偏向理论还是更偏向应用”等）。增加这些问题的原因不是为了制作排名或评级，而是为了加深对满意度决定因素的理解。在此，满意度是指提供的服务满足期望的程度。如果成绩好的学生期望值也较高，那么选拔度更高的大学（在加拿大往往是规模较大的大学）的满意度往往较低。关于图书馆满意度的数据最能证明这一点。就读于小规模大学的学生对于学校小型图书馆的满意度高于多伦多大学的学生，而实际上后者拥有的图书馆位列北美图书馆藏书量的前 20 名。新的数据收集方法有助于研究在控制某些因素（例如选拔度）的基础上分析满意度数据。

2010 年，加拿大“大学评级”和“大学指南”做出了一个更加重要的改进，即分领域提供数据，与教学有关的数据（如对教师的满意度、课程丰富性等）将按七大类学科提供，分别是人文与社科、自然科学、工程和建筑学、卫生科学、教育学、商学、音乐和美术。

哈萨克斯坦大学排名

Sholpan Kalanova
哈萨克斯坦独立教育质量保障署署长
电子邮箱：skalanova@mail.ru

概况

哈萨克斯坦大学排名的主要目的是：为学生、家长、雇主、政府机构、公共和国际组织提供关于高校质量的恰当、客观的信息；促进高校之间的竞争；推动高校的发展。

哈萨克斯坦大学排名是由哈萨克斯坦教育与科技部国家认证中心（National Accreditation Center of the Ministry of Education and Science）于 2006 年首次开发和公布的。2008 年，哈萨克斯坦独立教育质量保障署（Independent Kazakhstan Quality Assurance Agency for Education，IQAA）成立。它的成立是为了依照柏林原则的精神来从事排名活动：高校的质量保障必须由独立的非政府机构进行，这样可以不受教育部、高校和其他机构的影响。哈萨克斯坦独立教育质量保障署于 2008 年开发并制作了首份独立的哈萨克斯坦大学排名，目前正在制作 2010 版的排名。

2004～2005 年，大学排名在哈萨克斯坦开始受到教育与科技部的重视，随后排名流行起来。像上海交通大学“世界大学学术排名”这样的全球性排名变得越来越有影响力。“欧洲高等教育”2002 年第 27 卷第 4 期刊登了关于联合国教科文组织欧洲高等教育中心（UNESCO-CEPES）于 2002 年在华沙召开的排名研讨会的会议论文。这些文章对我们的排名方法有着重要的影响，因为这些文章强调排名方法不仅应关注大学的表现，还应考虑到大学所处的教育体系。与此同时，我们需要通过指标来引导哈萨克斯坦的大学达到国际水准。时任欧洲高等教育中心主任的 Jan Sadlak 博士（现任“IREG－学术排名与卓越国际协会”主席）对我们工作的影响很大并且支持我们的排名方法。在 2007 年的第 3 届 IREG 上海会议上，新开发的排名方法被用英文首次介绍。

排名指标与方法

为得到一份平衡的排名结果，我们的排名采用多个渠道的数据，主观与客观指标相结合的方法。排名指标体系由三部分构成：（1）高校学术资源评价（60%）；（2）专家问卷调查（10%）；（3）雇主、政府机构和企业问卷调查（30%）。

对高校学术资源的评价信息是从大学收集的。高校必须以电子或书面文档的形式提供有关学术指标的客观信息，并有证明文件支持，以确保信息的可信度。信息的收集期为 2～4 个月，之后对信息进行分析。从大学取得的数据还要和来自政府机构、高校网站以及其它信息源的数据进行核对。

在高校学术资源评估中，我们对高校应当如何参与和操作有明确的说明。值得一提的是，在过去的几年中我们不得不减少指标的数量。这是因为先前的排名是由教育与科技部管理及提供经费，现在则由独立教育质量保障署自己负担。因此我们需要改进排名的方法以减少指标的数量，这一过程需要走

访多所大学、进行相互交流。但我们意识到排名方法还必须符合《柏林原则》关于确保指标透明性、相关性和有效性的要求。先前的排名使用了 7 大类 42 个指标，现在我们使用 6 大类 36 个指标。为了确保排名结果的客观和透明，一些难以核实的指标被排除出指标体系。

高校学术资源评价由 6 大类指标构成：

第 1 类指标评价学生构成，含 5 个二级指标。其中一个二级指标测量全日制学生占所有学生的比例。这是因为对业余教育的质量是否能达到全日制教育的水平还存在质疑。此外，还有一个二级指标测量受政府奖学金资助的学生数量。哈萨克斯坦设有特别的政府教育奖学金，由教育与科技部根据国家统一招生考试的成绩在入学前颁发给即将入学的学生。获得政府奖学金的学生可以选择国立或非国立大学。因此，私立大学也有机会吸引政府经费。

第 2 类指标评价学生选拔度和学生学习成果，含 5 个二级指标。其中很重要的是一项由教育与科技部举办的学业中期测试（Intermediate Governmental Control Test），学生必须在完成两年的本科学习后通过这项测试，我们的排名中共有 3 个二级指标对此进行测量。此外，还设有一个二级指标考察高校的学生在国内外竞赛中取得的成果。

第 3 类指标评价教师构成，含 5 个二级指标。这类指标包括拥有学位的教师比例，获得政府颁发的“年度最佳教师”称号的教师比例。这一奖项是 2005 年由政府设立的，奖项的评审基于竞赛机制，胜出者将获得科研与海外进修的经费。其他指标还包括高校全职教师的比例和获得政府其他奖项的教师比例。

第 4 类指标评价高校的科研和创新潜力，含 8 个二级指标。我们重点关注的是科研经费，其中一个二级指标是师均科研经费数。其他重要的指标有：师均在哈萨克斯坦和其他独联体国家（前苏联国家）的期刊上发表的论文数，师均在国外期刊上发表的论文数。被引数据目前还没有被作为指标使用。其它指标包括：专利数量，博士论文评审委员会的数量，是否有国家实验室和高科技园区。

第 5 类指标评价国际交流水平。二级指标包括：参与国际交流的学生比例、教师比例，获得政府颁发的“Bolashak”奖学金（到全球顶尖大学进修）的毕业生人数，与国外大学的合作培养项目数量，新增指标包括外籍教师比例等。

第 6 类指标评价大学信息资源水平，含 5 个一级指标和 5 个二级指标。其中最重要的是大学的网站和内容，包含的指标有：网站信息量，更新频率，网站链接数和访问量。这类指标过去属于第 4 类指标，但由于网络的普及，大学的网络信息资源变得越来越重要，我们有必要对这类指标引起重视。

排名的第二和第三部分基于问卷调查，由专家和雇主对所有高校进行评价。

排名的第二部分基于专家对高校的评价。我们根据资历挑选出符合条件的专家，在问卷中给出对大学从“优秀”到“不满意”的 5 个等级以供选择。之前问卷被发放给学生和教师，通过 10 个问题来确定他们对高校的满意度水平。但后来问卷转而发放给高等教育领域的专家，因为对学生和教师进行问卷调查需要走访每所大学，成本太高。

排名的第三部分基于雇主就雇员表现的问卷调查。问卷中也列出了从“优秀”到“不满意”的 5 个选项。对雇主的调查问卷被发放到国家政府部门、地方政府、国有企业和

部分小型企业。此外，我们还采用了电话和邮件的形式进行问卷调查。

每所大学的排名顺序根据各分项指标汇总后的结果确定。排名包括一个涵盖所有大学的总排名和一个把大学按组划分的学科排名。此外，除了总排名和学科排名，基于学术指标、专家和雇主评价的独立排名也被公布。哈萨克斯坦现有 132 所大学，其中国立大学 42 所，非国立大学 90 所。国立大学和实力强的非国立大学更愿意参与排名，而实力弱的大学不愿意参与，因为他们担心在排名中名次很差。结果导致实力强的大学吸引了更多学生和国际经费，而实力弱的大学始终举步维艰。

我们的排名方法一直在不断地变化和改进，尽管不是大的改变。影响排名方法变化的原因之一是以 IREG 系列会议为平台的与其他排名机构的交流与合作。例如，我们最新的排名方法通过新增若干指标提高了网站质量的比重，还增加了国际合作有关的指标，以促进高校在这些方面更加完善。排名结果刊登在全国发行的报纸上，如《Liter》和《Express K》，还公布在 www.iqaa.kz 网站上，因此排名结果对所有人开放。目前，排名受到了即将入学的学生、政府机构和各种社会组织的欢迎。

结束语

当前哈萨克斯坦的教育发展战略计划包括使若干所哈萨克斯坦大学进入世界一流大学的排名。这使得在哈萨克斯坦，人们对大学排名的兴趣更大。实际上，国内大学排名的发布和发展在其中也发挥了作用。

The Academic Ranking of World Universities

Nian Cai Liu

Professor and Dean, Graduate School of Education, Shanghai Jiao Tong University

E-mail: ncliu@sjtu.edu.cn

Introduction

The development of world-class universities has been a dream of the Chinese people for generations. At the 100th anniversary of Peking University in May 1998, the then president of China declared that the country should have several world-class universities—resulting in the 985 Project, which is especially for building world-class universities in China. In 1998, Shanghai Jiao Tong University was selected by the Chinese government to be among the first group of nine universities in the 985 Project. At that time, many top Chinese universities drew up their strategic goals as to become world-class universities, and most of them set up a timetable. Shanghai Jiao Tong University was no exception. As a professor and vice-dean of the School of Chemistry and Chemical Engineering of the University, I was accidentally involved in the strategic planning process of building Shanghai Jiao Tong University into a world-class university.

During the process, I asked myself many questions. What is the definition of a world-class university? How many world-class universities should there be globally? What are the positions of top Chinese universities in the world higher education system? How can top Chinese universities reduce their gap with world-class universities? In order to answer these questions, we started to benchmark top Chinese universities with world-class universities and eventually to rank the world universities.

From 1999 to 2001, Dr. Ying Cheng, two other colleagues and I worked on the project of benchmarking top Chinese universities with four groups of US universities, from the very top to ordinary research universities, according to academic or research performance and based on internationally comparable data. The main conclusions included that top Chinese universities were estimated to be in the position of 200 to 300 in the world. The results of these comparisons and analyses were used in the strategic planning process of Shanghai Jiao Tong University. Eventually, a consultation report was written and provided to the Ministry of Education of China.

The publication of the report resulted in numerous positive comments, many of which involved the possibility of making a real ranking of world universities. During the time, many friends from different parts of the world, who visited us for other purposes, learned about our study and encouraged us to do world rankings. They reminded us that not only in China but also universities, governments, and other stakeholders in the rest of the world are interested in the ranking of world universities. Therefore, I decided to undertake this project and we spent another two years until the Academic Ranking of World Universities (ARWU) was first completed in early 2003. In June 2003, the ARWU was published on our web site (http://www.arwu.org).

Methodologies of ARWU

The ARWU uses six objective indicators to rank world universities. The indicators and its weights are the number of alumni winning Nobel Prizes and Fields Medals (10%), number of staff winning Nobel Prizes and Fields Medals (20%), number of highly cited researchers selected by Thomson Scientific (20%), number of

articles published in journals of *Nature* and *Science* (20%), number of articles indexed in Science Citation Index—Expanded and Social Sciences Citation Index (20%), and per capita performance with respect to the size of an institution (10%).

We have scanned every institution that has any Nobel Laureates, Fields Medals, Highly Cited Researchers, or articles published in *Nature* or *Science*. In addition, major universities of every country with significant amount of articles indexed by Science Citation Index-Expanded (SCIE) and Social Science Citation Index (SSCI) are also included. In total, more than 2,000 institutions have been scanned, and about 1,200 institutions have actually been ranked. A list of the top 500 institutions has been published on the Web. Considering the significance of differences in the total scale, the ARWU results include groups of 50 institutions in the range of 100 to 200 and groups of 100 institutions in the range of 200 to 500.

ARWU-FIELD & ARWU-SUBJECT

Ever since its publication, the ARWU has attracted attention from all over the world. Numerous requests have been received, asking us to provide a ranking of world universities by broad subject fields/schools/colleges and by subject fields/programs/departments. We have tried to respond to these requests.

In February 2007, the Academic Ranking of World Universities by Broad Subject Fields (ARWU-FIELD) was published on our web site. The five broad subject fields include the natural sciences and mathematics, life and agriculture sciences, engineering/technology and computer sciences, clinical medicine and pharmacy, and the social sciences. Arts and humanities are not ranked because of the technical difficulties in finding internationally comparable indicators with reliable data. Psychology and other cross-disciplinary fields are not included in the ARWU because of their interdisciplinary characteristics. Similar to the ARWU, institutions are ranked according to their academic or research performance in each broad subject fields. Besides the indicators used in ARWU, two new indicators were introduced: (a) the percentage of articles published in the top 20 percent journals of each broad subject field and, (b) the research expenditures (for engineering ranking). The list of top 100 universities in each broad field was published.

In October 2009, the Academic Ranking of World Universities by Subject Fields (ARWU-SUBJECT) was published, which ranks institutions in five subject fields, including Mathematics, Physics, Chemistry, Computer Sciences and Economics/Business. The list of top 100 universities in each subject was published.

Impact of ARWU

Although the initial purpose of ARWU was to find the global standing of Chinese top universities, it has attracted a great deal of attention from universities, governments and public media worldwide. A survey on higher education published by *The Economist* in 2005 commented ARWU as "the most widely used annual ranking of the world's research universities". Burton Bollag, a reporter at *Chronicle of Higher Education* wrote that ARWU "is considered the most influential international ranking".

One of the factors for the significant influence of ARWU is that its methodology is globally sound and transparent. The *EU Research Headlines* reported the ARWU work on 31st December 2003: "The universities were carefully evaluated using several indicators of research performance." Chancellor of Oxford University, Chris Patten, said "the methodology looks fairly solid ... it looks like a pretty good stab at a fair comparison." The ARWU and its content have been widely cited and employed as a starting point for identifying national strengths and weaknesses as well as facilitating reform and setting new initiatives. Martin Enserink referred to ARWU and argued in his paper published in Science that "France's poor showing in the Shanghai ranking ... helped

trigger a national debate about higher education that resulted in a new law... giving universities more freedom".

Ongoing Efforts to Diversify the ARWU

The ARWU tried to rank research universities in the world by their academic or research performance based on internationally comparable third-party data that everyone could check. No subjective measures were taken. It was done independently for our academic interests, with potential applications in the strategic planning of Chinese universities. Nevertheless, there are still many methodological and technical problems.

Methodological problems involve the balance of research with teaching and service in ranking indicators and weights—inclusion of non-English publications, the selection of awards, and the experience of award winners. Technical problems exist in the definition and name expression of institutions, data searching and cleanup of databases, and attribution of publications to institutions and broad subject fields. We have been working hard to study all the above-mentioned problems and to improve our ranking.

In addition to the field ranking and subject ranking, we are surveying the possibilities of providing more diversified ranking lists, particularly rankings based on different types of universities with different functions, disciplinary characteristics, history, size, and budget, as well as other topics. Furthermore, we have been doing theoretical research on ranking in general, seeking to contribute to the understanding of ranking. We have also been actively participating in international societies and communities related to ranking, such as the International Ranking Expert Group (IREG)—International Observatory on Academic Ranking and Excellence (http://www.ireg-observatory.org).

The New Methodology of THE's World University Rankings

Phil Baty

Editor of the THE World University Rankings, and Deputy Editor of THE

E-mail: phil.baty@tsleducation.com

Times Higher Education （THE）magazine has been publishing its World University Rankings since 2004, and the annual tables have become established as one of the key events in the international higher education calendar.

Conceived primarily as a basic tool for students to help inform their study choices, the rankings are now commonly used by academic faculty to help them select collaborative partners or new employers and by university managers to help set institutional strategies and to benchmark institutional performance. As national economies are re-shaped for a knowledge- and innovation-driven future, the rankings are also used by governments and policy makers to help shape or monitor national higher education policy.

In November 2009, THE accepted that the rankings it had been publishing, despite gaining great prominence internationally, were not suitable for the purposes for which they were being used. Because the rankings had become so influential, and because their reach had become so wide, the magazine's editors believed they had an obligation to make them more rigorous, sophisticated, balanced and transparent. As the international academic community was taking the rankings more seriously than when they were first conceived, the magazine's rankings needed to be more accountable to the academic community.

New Data Providers

In November 2009, THE has ended its six year arrangement with the company QS, which supplied the world rankings data between 2004 and 2009. From 2010 THE publishes the world university rankings with a new methodology and with all the data collected and analysed by the global research information specialists, Thomson Reuters.

The problems and concerns with the old 2004-2009 THE-QS methodology are well documented. Of most concern was the so-called "peer review" score of the old rankings. Some 40% of a university's overall ranking score was based on the results of a "peer review" exercise —in fact a simple opinion survey of academics, asking them which institutions they rated most highly. Many object in principle to the use of any subjective measures in rankings, arguing that they reflect past, not current, performance, they are based on stereotype or even ignorance, and that a good or bad reputation may be mindlessly replicated.

But we believe that reputation indicators can provide useful context, capturing information that simple quantitative data can't. Reputation is also crucial in the current, highly competitive global higher education marketplace. But to be useful such opinion surveys must be carried out very rigorously. The reputation survey carried out by our former ranking data supplier QS attracted only a small number of respondents. In 2009, about 3,500 people provided their responses—a fraction of the many millions of scholars throughout the world. A total of 9,386 responses were used to compile the 2009 world university rankings tables, because three years worth of survey data were aggregated. In THE's view, the sample was too small, and the weighting given to this indicator was too high.

Improved Reputation Survey

We have made major improvements to the rep-

utation survey for 2010. Thomson Reuters hired professional polling company Ipsos MediaCT to carry out a new reputation survey, with a clear brief to gather a higher number of respondents from a respondent pool that truly represents the international university community—in terms of both geography and subject mix.

Perhaps most significantly, the opinion poll was sent only to invited participants, selected to be properly representative of their discipline and their region, and selected on the basis of proven knowledge and experience in their field. Most respondents were sourced through Thomson Retuers' bibliometric databases to ensure that they have a record of research publication. We refused to adopt a "scatter gun", mass mailing approach to collect the responses. We rejected the idea of allowing academics to volunteer to fill in the reputation survey.

Thomson Reuters' Academic Reputation Survey was a resounding success. In just over three months, during a single year, it achieved 13,388 responses. That is around four-times more than anything ever collected in a single year under our old THE-QS ranking system. But it is not the size of the sample alone that is important—the survey also met our demands for a more representative sample. Some 30.2% of all responses came from the Asia Pacific region, compared to 38% from the Americas, including South America, 28.3% from Europe and 3.5% from Africa.

The responses also reflected the full breadth of subject areas—with most responses (22.8%) coming from the field of engineering and technology. Some 20.5% of responses came from the physical sciences, followed by 17.8% in life sciences and 17.8% in social sciences. 15.3% came from the field of Clinical, pre-clinical and health, while 5.8% came from the arts and humanities. Statistical analysis is applied to iron out any response biases. Some 6.5% of respondents said that their current role was in "senior institutional leadership". The most responses, 68.9%, described themselves as "academic staff", 14.9% as "research staff". Respondents had a high level of experience working in the higher education sector with the mean number of years self-reported at 17 years. When asked to report their main activities in their current role, 54% of respondents said they spent most of their time on research, 31% said teaching was their primary role and 12.6% replied "administration/services" and 1.8% said that they were not currently working.

For the first time in 2010, survey respondents were asked to judge reputation in both teaching and research—allowing the creation of the first ever indicator of teaching reputation for the 2010 World University Rankings. The survey questions were carefully prepared to elicit meaningful responses and to make clear what is being judged. Instead of a simple, generic, "who is best," the survey asked more detailed questions designed to elicit more informed and consistent answers, such as asking respondents which institutions produce the best graduate applicants, or where they might recommend their top undergraduates should apply for the best graduate programmes.

In THE's new ranking, the results of the reputation survey makes up 34.5% of the overall ranking score. This represents a reduction to the weighting given for reputation indicators compared to the old THE-QS methodology, despite dramatic improvements to the reputation survey.

Citations Data Normalized for Subject Variation

The other serious concern raised about the old THE-QS rankings was the way research excellence was examined. Under the old system, QS measured the number of times an academic's published work was cited by his or her peers. In an indicator worth 20% of the overall score, QS took the total number of citations for all papers published from an institution, and divided it by the number of full-time equivalent staff at that

institution. But this approach took no account of the different citation practices, and very different citation volumes, between disciplines. For example, the average citation rates by field (from Thomson Reuters' Essential Science Indicators database, 1 January 1998 to 31 October 2008) vary from just 3.06 in computer science and 3.83 in engineering to 20.58 in Immunology and 24.75 in molecular biology. QS's failure to normalize its citations data to reflect different citations volumes in different subjects, had the effect of disadvantaging institutions with strengths in areas where citation rates are lower, while boosting those institutions with strengths in areas where citations rates are much higher. This was unfair and misleading.

By working on the rankings for 2010 and beyond with Thomson Reuters, which owns citation databases covering 12,000 of the highest-impact journals and more than 110,000 conference proceedings, THE will draw on a very deep citations database. But it is not just the quality of the data which matters. By working directly with the owners of the citations data—rather than simply buying the data in from a third party source—THE can also draw on world-leading expertise in understanding and analyzing the citations data.

So from 2010 we adjust the data to take account of subject variations. Indeed, the full annual top 200 table is built from the subject level upwards. Data are collected and analysed for six broad subject areas: arts and humanities; clinical, pre-clinical and health; life sciences; physical sciences; engineering and technology; and social sciences. The 2010 rankings include six separate ranked tables in each of these subject areas. The overall top 200 table is built from the subject data. This represents another major improvement on the old THE-QS rankings system, where the five subject tables were based on just one indicator, the results of the QS reputation survey. So the THE-QS subject tables were a subjective measure of reputation only.

Methodology

Before the publication, THE has released the first draft of its proposed new rankings methodology for open consultation. THE uses 13 separate performance indicators to compile the league tables for 2010 and beyond—an increase from just six measures used under the old THE-QS methodology employed between 2004 and 2009. This wide range of individual indicators is grouped to create four broad overall indicators to produce the ranking score.

The general approach is to decrease heavy reliance on any individual indicators, which can be crude and easy to manipulate, and to instead use a large basket of indicators, grouped across broad categories related to the function and mission of higher education institutions. The four core aspects of a university's activities that we assess are: research; economic activity and innovation; international diversity; and a broad "institutional indicator" including data on teaching reputation, institutional income and student and staff numbers.

For the broad area of "research", five separate indicators are used. This category includes a measure of "citation impact", looking at the number of citations for each paper produced at an institution to indicate the influence of its research output. It also includes a lower-weighted measure of the volume of research from each institution, counting the number of papers produced per member of research staff. The research category also looks at an institution's research income, scaled against research staff numbers. This research category also includes the results, with regard to research, of the Academic Reputation Survey.

For 2010, the "economic/innovation" indicator uses data on research income from industry, scaled against research staff numbers. In future years, it could possibly include data on the volume of papers co-authored with industrial partners and a subjective examination of employers' perceptions of graduates.

"Institutional diversity", as under the old

THE-QS system, is examined by looking at the ratio of international to domestic students, and the ratio of international to domestic staff. It may also include, in future, a measure of research papers co-authored with international partners.

"Institutional indicators" include the results of the reputation survey with regard to teaching, and look at an institution's overall income scaled against staff numbers, as well as data on undergraduate numbers and the proportion of PhDs awarded against undergraduate degrees awarded.

Fit for the Twenty-First Century

We believe our new THE World University Rankings, powered by Thomson Reuters, represent a major improvement on past THE-QS ranking exercises between 2004 and 2009. We believe that the rankings is more sophisticated, balanced rigorous and transparent, reflecting strengths across the range of activities modern, research-intensive global higher education institutions are engaged in. We hope that the rankings will be a more suitable and serious tool to help the global sector understand itself in a rapidly changing world, rather than just an annual curiosity.

Ranking Web of World Universities

Isidro F. Aguillo and Noelia Gutiérrez Labajos

Cybermetrics Lab, IPP-CCHS-CSIC, Spain

E-mail: Isidro.aguillo@cchs.csic.es

Background

In the mid-nineties, the Cybermetrics Lab, a research group belonging to the Consejo Superior de Investigaciones Cientificas (CSIC, National Research Council), the largest public research organization in Spain, started to develop quantitative web indicators by using techniques derived from bibliometrics and scientometrics. The Lab focused its research on academic and research organizations, but its efforts were hampered by the still limited web presence of these institutions and the shortcomings of the main indicator then available, the Web Impact Factor. Some data collection was done using crawlers since 1999; however, after checking the huge human and computer resources needed, it was decided to use commercial search engines instead for the collection purposes.

In 2003 the introduction of the ARWU Ranking and specially its composite indicators developed by the Shanghai Jiao Tong University provided a model for the exploitation of the webometric data. The Ranking Web (www.webometrics.info), also known as Webometrics Ranking (WR), first appeared in 2004 with the explicit aim of promoting public dissemination of scientific knowledge, the so called Open Access initiatives. WR is basically a research product. Its methodology has been developed and improved continuously, even till now. Although the stability is not the target, the main factors of its methodology were decided and adopted in 2006.

Hypothesis and Model

The Web is becoming the most important scholarly communication tool and it makes more and more scientific information accessible. Electronic publication is cheap, fast and it could reach huge audiences all over the World. It is used not only for distributing academic and research papers in e-journals, but as a showcase describing the structure, activities and results of the academic and research institutions. The central hypothesis of the Ranking is that the university's web presence reflects its global performance, the quality of its departments and services, the impact of its outputs and its international prestige. The commitment to open access policies is an important asset and it is also considered in the Ranking design.

If the hypothesis is valid then the Ranking could be used for the description and evaluation of universities, adding an interesting tool to the design of science and higher education policies worldwide. Even if web presence is not yet a priority for some universities, the Ranking probably creates a virtuous circle pushing them to increase its electronic presence in the webspace.

The Ranking's model is build based on the following principles: (a) The Ranking Web is a ranking, an ordered league table of universities. It uses a composite indicator combining quantitative variables that intends to offer a simple, current, overall, not detailed description of the global performance of the university. (b) The Ranking Web is one ranking. Only one composite indicator is used as we do not find other feasible alternatives describing better the global scenario and more research is needed for analyzing the impact and contribution of the individual variables. (c) The Ranking Web is global. The objective is to rank all the world universi-

ties, a figure close to 20,000, including especially those from developing countries. (d) The Ranking Web takes into account all the university missions. The research performance is a major indicator for world-class universities but for most of the universities the other missions are also very relevant. Teaching, especially off-classroom support and distance learning can be indirectly measured from web activity whereas the so-called third mission, transfer of knowledge, could be uncovered through link analysis.

Qualitative analysis can be useful for describing smaller populations, but even in these situations benchmarking exercises (surveys) can be very biased. Quantitative approach has been quite successful for closed universes like research-based academia (bibliometrics). However, webspace is clearly a more difficult target. For example, link motivations can be diverse and varied; the number of citations and networks can be large; and they can be visited by unexpected stakeholders. However the data to deal with are far larger (in the order of millions). According to the big number theory, it is possible that meaningful patterns could appear even in this complex scenario. This is especially interesting for a ranking as it allows a better discrimination of the ranks.

Building the Ranking

The directory of universities' web domains consist of more than 20,000 entries and may take several years for compile, clean and update. Not all of the universities have only one webdomain. Some of them have several central domains. Even after a new domain is applied, the old domain may be still in use and enjoys a higher visibility. A number of 18,000 independent higher education institutions can be shown based on its domain names, with dozens of new institutions emerging every year and with a few hundreds of institutions changing domains every year.

Due to the large population and the number of variables involved, an indirect data capture method is chosen. The values are extracted from the main commercial search engines that usually have huge updated databases with easy operators with webometrics capabilities. In order to avoid the problems associated to different coverage of each engine, inconsistency of results and other mistakes, the data are collected from several sources, at least two times in a short period for error checking.

The composite indicator used in WR was inspired by the journal impact factor, which is obtained dividing citations (visibility) by papers (activity). In the web the visibility is obtained from counting external inlinks, whereas activity refers to the number of webpages. Due to the "power-law" distribution of these variables, a polynomial approach is chosen to maintain the ratio 1:1 between activity (50%) and visibility (50%). The importance of formal communication in the academia and the new citation bibliographic database freely available from the Web (Google Scholar) suggested making some changes in the way the activity was measured. The current WR is calculated as follows: (a) Activity: A weighted combination of Size (20%), the number of pages according Google, Yahoo, Bing and Exalead search engines (median of the log normalized values); documents (15%), the number of rich files (pdf, doc, ps & ppt formats) obtained in a similar way and Papers (15%), the number of entries in Google Scholar. (b) Visibility: The number of external inlinks according to Yahoo and Exalead search engines (50%). The current method gives extra value to those links coming from academic webdomains.

Till 2009 the Ranking published the ordinal values for each university and calculate the WR composite indicator from these ranks. The ranks for each indicator are still provided by each university, but the final rank is now calculated as the normalized values, which has not yet been made public. Several options are being studied to avoid problems with tie ranks that could greatly increase.

The weights are chosen a priori according to the 1:1 model. Meanwhile, both informal and formal communication are taken into account. In this case, at least two indicators are under scrutiny for improvement: (a) visibility should reflect impact in both non-academic and academic websites, but giving extra weight to the last ones. This can be technically difficult, as it involves calculation of very large matrixes. (b) Google Scholar still have many problems in its beta version: biased coverage, inability to recover only full texts or specific file formats, and faulty by-year or discipline filtering. If these problems are solved in the near future, weights to the variables should be re-adjusted.

This method intends to capture the three missions of the universities as a whole. However, weaknesses are inevitable: it is not easy to separate the contribution of each mission; bad practices of web naming have a huge negative impact on the result; and there are important disciplinary biases, despite it might favor traditionally bibliometric neglected disciplines, such as technologies, computer science, social sciences, humanities.

Results

Preliminary analysis of the WR's results shows a strong macro-level correlation with those provided by other Rankings. Major discrepancies between the Ranking Web's result and other rankings at individual level can be explained as due to different selection criteria (e.g. university hospitals inclusion), different methodologies adopted (e.g. surveys, weighting) and regarding webometrics, the degree of commitment with the web (e.g. unexpected under- or over-performing of some universities) and some controversial decisions regarding web domain names (e.g. many different domains, old discarded domains still used, shared domains).

However, the most important result is the evidence of an academic "digital divide". During the January 2010 edition, the distribution of universities in the Ranking Web by region was as follows: In the Top 100 group, North America was 76, Europe 17, Asia 3, Oceania 2, Latin America 2. In the Top 200 group, North America 114, Europe 60, Asia 15, Oceania 6, Latin America 4, Arab World 1. In the Top 500 group, North America 200, Europe 223, Asia 45, Oceania 14, Latin America 14, Arab world 3, Africa 1. In the Top 1000 group, North America 370, Europe 408, Asia 134, Oceania 35, Latin America 44, Arab world 4, Africa 5. In the Top 1000 group, North America 370, Europe 408, Asia 134, Oceania 35, Latin America 44, Arab world 4, Africa 5. In the Top 8000 group, North America 2577, Europe 2424, Asia 2009, Oceania 80, Latin America 744, Arab world 115, Africa 51. Checking the Top 100 and 200 groups, there are far more North American (USA and Canadian) universities in the Top 200 than their European counterparts. No other Ranking shows this concerning situation that is clearly related to the web policies of the elite universities. European and Asian universities should understand that they have to improve their support of web publication, including improving quality presence developing large Open Access repositories and increasing digital networking initiatives.

Research intensive institutions are known as the best ranked, the so-called world-class universities, but there are also large Latin-American and Asian universities in good positions with an important regional impact.

Final Comments

Web indicators are easy to build and they are able to offer a global vision of the performance of the academic institutions all over the world, including those from developing world and BRIC countries (Brazil, Russia, India and China). This make the Webometrics Ranking a very useful tool for students, scholars, researchers and policy makers, although more information is needed especially about the contribution and meaning of each of the criteria used.

Web rank is unrelated with the design of

the websites or the number of visits they received. Although it is size-dependent, the Ranking provides relevant information about the rank of a large number of universities from countries from all regions.

Future developments of the WR will be based on the results of our research agenda that includes: classification of the institutions (public/private, size, main discipline or field), weighting adjustment of the indicators, evaluation of the rich files formats (re-evaluating ps, xls), media files (video), new search engines (locally relevant like Baidu, Yandex, Yahoo Japan) and improved data visualization. Other proposals for the WR include merging universities and research centers' rankings, developing department or research groups rankings (by discipline), and integrating it with bibliometric based rankings.

The WR's family consists of the following websites: Universities (www.webometrics.info), Research centers (research.webometrics.info), Hospitals (hospitals.webometrics.info), Repositories (repositories.webometrics.info) and Business Schools (business-schools.webometrics.info).

The Impact of HEEACT's Rankings in Quality Enhancement of Taiwan Higher Education

Yung-chi Hou (Angela)

Director of International Exchange, Higher Education Evaluation & Accreditation Council of Taiwan
Associate Professor, Fu Jen Catholic University, Taiwan, China
E-mail: 035440@mail.fju.edu.tw

Quality Assurance and Academic Excellence —the Role of HEEACT as a Ranker

Most rankings used to be published by mass media, such as U.S. News & World Report. Nowadays, the type of ranking providers is still quite diversified. Some are produced by institutions themselves, such as Shanghai Jiao Tong University's Academic Ranking of World Universities. Some are imposed on tertiary education institutions externally by a governmental accreditor. In some countries, the ranking exercise is undertaken as part of the accreditation process, either by the accreditation agency itself, or by the authority in charge of tertiary education. Higher Education Evaluation & Accreditation Council of Taiwan (HEEACT) is one of the most successful cases.

Under Taiwan's "University Law" revised in 2005, all Taiwan universities and colleges are obligated for assessments regularly with regard to standards and procedures by accrediting agencies chartered by the Ministry of Education. Based on these bylaws, HEEACT was established in 2005 and began to accredit 76 4-year comprehensive universities and colleges in Taiwan. At the same time, HEEACT, as Taiwan's quality assurance agency, had been highly expected to offer related internationally comparable data and information for the Taiwanese government to be used in higher education policy making. Due to this, in 2007, HEEACT started to develop varying ranking systems, including "Statistical Analysis on Taiwan WOS Papers", "Statistical Analysis on Taiwan ESI Papers and h(m) Indicators", "Performance Ranking of Scientific Paper of World Universities", "Analysis on Patents by Universities and Colleges in Taiwan", "Performance Assessment on University and Industry Collaborations", and the "College Navigator". It is expected that each institution will be able to develop its own competitive edge and undergo self-positioning based on the features of HEEACT's different ranking tools.

HEEACT Global Ranking

Initially published in 2007, the major goal of the HEEACT's "Performance Ranking of Scientific Papers for World Universities" (SPWU) was to evaluate the current scientific paper performances of the top 500 world universities in order to find out the gap between Taiwanese universities and the rest of the world's universities. Also, the HEEACT global ranking attempts to provide universities in smaller, newly developing nations with insights into how to create influential research universities. HEEACT states clearly that, however, the ranking only presents the research performance of the universities instead of their overall quality in terms of governance, teaching and services.

The HEEACT global ranking employs data drawn from SCI and SSCI to evaluate universities' research performance. It considers publishing in international peer reviewed journals as the predominant mode of scientific research output, thus taking statistics on articles published in listed publications as an effective indicator of reflecting universities' research

performance. The 2007 performance measures were composed of nine indicators. The indicators together represented three different criteria of scientific papers performance to demonstrate the strengths of the institution: research productivity, research impact, and research excellence. In the 2008 edition, the indicator of "the number of subject fields where the university demonstrates excellence" was removed because of the publication of HEEACT's new type of field ranking. Now there are a total of eight indicators to measure the performance of research outputs of the universities, including number of articles in the last 11 years and in the current years, number of citations in the last 11 years and in the current years, average number of citations in the last 11 years, H-index of the last two years, number of highly cited papers, and number of articles in high-impact journals in the current year.

As HEEACT President Roger Chen has stated that the objective indicators used in this ranking system are designed to measure both long-term and short-term research performance of each university. He claimed that analyses of SCI and SSCI make global university ranking fairer, with an emphasis on both quality and quantity of publications. It also takes account of recent research performance in order to make a fair comparison between institutions with different lengths of history. It incorporates average number of criteria in its calculation of the score so as to prevent a predominance of large universities.

In 2008, HEEACT published an additional edition based on institutional size in order to minimize its impact on the final outcome. In addition, HEEACT also developed a global ranking by field and published the top 300 institutions in each field in 2008. The six fields were agriculture & environment sciences, clinical medicine, engineering & computing, technology, life sciences, natural sciences, and social sciences.

According to the HEEACT 2009 global ranking outcomes, institutions in the US and Europe still play a predominant position in the international higher education landscape. It is noteworthy that all of the world's top 10 universities are US universities, while only two universities in the Asian-Pacific region are ranked within the world's top 30 universities, and both of them are from Japan.

In fact, due to its global character and scientific approach, SPWU has become one of the most influential global rankings. More than 3,000 visitors browsed the website with HEEACT's ranking on the day it published. It not only has drawn much international and national attention but also has served as a useful benchmark for those who are eager to develop themselves into world-class research universities.

College Navigator in Taiwan

The other innovative ranking published by the HEEACT is "College Navigator". Though HEEACT publishes accreditation outcomes and the final reports on the website each year, not many users could fully understand the information. Hence, two major reasons underlay the pilot project "College Navigator in Taiwan" (CNT) launched by HEEACT in October, 2009: one is HEEACT acting as a marketing proxy for colleges and universities, and the other is to enhance the internationalization of Taiwan higher education.

Compared to the classic ranking methodology, users of CNT are given a certain extent of autonomy over selection of indicators and weightings. They are able to select the number of indicators, between three to ten within each broad criteria, and then weigh each one by their own judgment. In addition, users are able to rank the institutions they are interested in by region, type, size and program. More unranked information on universities that users are likely interested in, such as founding year, mission, total enrollment, number of programs, website, accreditation status, government funding, application, room and board, tuition, student clubs,

is presented as a supplement to the final ranking outcomes as well.

CNT with its consumer-based mission, in fact, should not be viewed a real ranking in the traditional sense. As key feature, it gives individual users the opportunity to establish their own rankings according to a number of self-chosen criteria. Moreover, the website just provides robust information (distinguishing only top—middle—bottom groups per indicator), such as CHE ranking, not spuriously precisely simple and overall ranking.

CNT has become a convenient college search tool for high school graduates in Taiwan and abroad. However, the same student users still expected to have more autonomy over the selection of indicator number and to have more transparent data about colleges and universities. HEEACT keeps updating the data of all ranked subjects and will include more than 80 universities of technology and sciences and technical colleges in the ranking system by 2010. In the long-term plan, HEEACT will attempt to invite universities in the Asian Pacific region to join the project. It is expected that more and more local and international students will be assisted with useful information in school selection at home and abroad.

HEEACT Rankings and Building World-Class Universities

Globalization in the 21st century presents universities and countries with a number of challenges and opportunities. Currently, the major concern for both of them is how to assure quality in higher education and to enhance global competitiveness through a variety of polices and actions. Hence, many governments attempt to use rankings as a basis of building world-class universities. There is no exception for Taiwan.

In response to the quest for a world-class university, the Taiwan government launched the Five-year 50 Billion Program for Developing First-class University and Top Research Centers in 2005. The program aims to develop at least one university as one of the world's top 100 universities in five years and at least fifteen key departments or cross-campus research centers as the top in Asia in ten years. The five-year 50 Billion Program can be linked with other strategies such as the wholesale restructuring of the higher education system for internationalizing Taiwan's higher education sector. More importantly, it marks Taiwan's intention to join the competition among other higher education systems in the region under the theme of building a 'world-class' university. To achieve this goal, 11 or 12 research universities granted are required to complete a five-stage evaluation ranging over the funding period in order to renew their projects in the following year. Also, the MOE consistently keeps its eyes on their performances of the funded institutions in the HEEACT global ranking.

According to the HEEACT 2009 global ranking, there are seven Taiwanese research universities on the top 500, including National Taiwan University (102) , National Cheng Kung University(307), National Tsing Hua University(347), National Chiao Tung University(456), Chang Gung Univesity (479), National Central University (483) and National Yang Ming University(493), as compared to five in 2008. As to the total number of universities on top 500 in the Asian Pacific region, Taiwan is ranked number 5 with a rate of 8%. Those five Taiwanese universities are all the recipients of the "five-year 50 Billion Excellence Program".

Although the number of Taiwanese universities moving into top 500 is gradually growing in the HEEACT ranking each year; however, in some senses, the HEEACT global ranking has provoked severe criticism over its methodology and purposes from Taiwan college presidents and some board members of HEEACT. Those Taiwanese universities which are not on the top 500 have been worried whether the HEEACT ranking would be adopted as the only criteria in the selection process in the future. Some, on the other hand, have criticized that teaching quality of these selected research universities do not seem to be assessed because some of the programs among 11 selected universities are granted as status of

"conditionally accredited".

Conclusion

HEEACT President Roger Chen recently explained the different functions of the HEEACT accreditation and ranking in this way, "different from those global ranking systems in which only research outputs is counted, HEEACT accreditation mainly focuses on teaching quality. I am hoping that universities will not misuse them, being misled by both tools". However, it cannot be denied that some research universities are under great pressure to provide students with a good learning environment and some other teaching universities set their institutional long-term goal of "moving into the Top 500". More and more Taiwanese institutions are using the performance indicators of the HEEACT accreditation and ranking as a tool of self–enhancement and have changed their institutional policies in some aspects.

Certainly, quality and excellence in higher education have become the major concerns in Taiwan. The pressure from international competition and accountability will continue to accelerate the importance of ranking and quality assessment in Taiwan's higher education.

The CHE University Ranking

Gero Federkeil

Project Manager, CHE, Germany
E-mail: gero.federkeil@che-ranking.de

Introduction

In the course of the last two decades, higher education rankings have emerged in many countries all over the world. Despite their now long tradition (the first ranking by U.S. News &World Report was published in 1983) rankings are still very controversial, in particular within higher education institutions. Alex Usher and Massimo Savino said, "wherever rankings have appeared, they have been met with a mixture of public enthusiasm and institutional unease". Rankings were established to create transparency about the higher education system in a competitive system for market actors —prospective students, their parents, employers. Rankings are simultaneously the medium and the outcome of competition. Rankings can be conceived as an imperative of the knowledge society. This means they are reproducing the competitive structures they are trying to measure. As rankings are constructing—with high public visibility—such hierarchies of higher education institutions in terms of better and worse and rankings might impact the market situation of single institutions, e.g. applications, it is no wonder that they are followed by those institutions both attentively and skeptically.

In Germany the first ranking was published in 1989 by a weekly journal. In the next few years, a number of rankings followed, mostly produced by journals. In 1998 the first ranking by the Centre for Higher Education Development (CHE) was published. CHE is a private, independent non-profit organisation that was founded by the German Rectors' Conference together with a private business foundation. The development of a ranking based on scientific methods of empirical social research was one of the founding tasks of CHE. The ranking followed a project of the Rectors' Conference to develop a set of indicators to describe institutional profiles.

Purpose and Development

CHE started its ranking after two years of intensive discussion with evaluation and methodological specialists as well as with students who gave insights on which information they expect from a ranking that is focusing on their needs for information. The main purpose of the ranking is to provide information on universities and programmes for prospective and mobile students and to help them making informed choices. Furthermore the ranking offers comparative information about their performance to universities themselves. It includes indicators on teaching and learning, resources and facilities, research activities as well as some information that is important for this target group but is not related to the performance of universities (such as local rents, size of the universities etc.). Research is included for two reasons: first, for some prospective students, information about research activities and performance is relevant for their decision making about their future university to study right from the beginning, and, second, a ranking without information on research would not be accepted very much by the universities and the higher education sector itself. As research is not in the centre of a ranking devoted to prospective students, CHE decided to publish a special a particular "Research Ranking" that gives more detailed insights into research performance for an academic target group. In this ranking the data on research are analysed an published in more detail. In addition to indicators on publications, citations, number of PhDs, research grants and patents, some bivariate and correlational analysis is included there.

The number of fields included in the ranking continuously increased. Today the ranking includes 32 fields. The selection of fields is mainly based on the number of students, that is, those chosen fields will have the highest number of students in Germany. Now the ranking covers the fields of about 80% of German students. There are limitations with regard to the inclusion of additional, in particular small fields. First, a further extension would be a challenge to resources both at CHE and the universities who have to deliver data for the ranking. Second, a ranking does not make sense for small fields that are offered only at a small number (less than about 15) of universities. In the context of the Bologna process there is also a growing demand for international comparative information on higher education institutions and programmes. CHE started to internationalise its university ranking in 2004 by including, step by step, universities from Austria, Switzerland and the Netherlands.

Basic Approach

The CHE ranking is characterized by three basic principles:

(a) Level of ranking: programme/field instead of whole institutions. In contrast to most rankings, CHE ranking is field-based. Evidence from the CHE ranking shows that universities can be very heterogeneous with regard to the performance of their individual departments. A university might perform well and hence be ranked high in physics and at the same time perform poorly and be ranked low in history. Prospective students who are the major target group of CHE ranking are interested in information about the field they want to study; averages on a whole university do not help them to make an informed choice. The finding, for example, that a particular university as a whole is ranked in the middle of the distribution is of no use for such prospective students interested in physics if it is ranked low.

(b) Multi-dimensional ranking instead of composite overall score. The number of indicators differs between rankings, but independently from that number most rankings calculate an aggregated overall score by giving particular weights to the indicators. By selecting a particular set of indicators and assigning specific weights to each indicator, rankings impose a specific definition of quality. According to the U.S. National Opinion Research Center, neither a theoretical nor an empirical basis is employed in developing such weighting procedures. With regard to the users /target group of the ranking, the heterogeneity of decision preferences should be taken into consideration. Some students are looking for a university with high research activities (as measures e.g. by research grants, publications etc.) while other students may look for a university with close contacts between students and teachers, good mentoring and short study duration. Calculating an overall score is thus to patronise the target group.

Furthermore institutional-level scoring levels out differences between particular aspects of a program or university's performance. This is most evident in rankings including indicators on both teaching and research. A university with good research performance does not necessarily provide good teaching and learning experiences to their students and vice versa. Multi-dimensional rankings can provide a differentiated insight into the strengths and weaknesses of a university. This is the only way to take into account the multi-faceted nature of quality. This view leads Usher & Savino (2007: 23) from their analysis of ranking systems to conclude that "one of the main reasons of institutional unease [with rankings] is the tendency of institutional ranking schemes to use weighted aggregates of indicators to arrive at a single, all-encompassing quality score".

(c) Groups instead of league tables. In the tradition of the U.S. News & World Report rankings most rankings order universities in league tables with individual rank positions. This approach suggests that each difference in the numeric value of an indicator marks a difference in quality/performance between the entities ranked. League table comparison inevitably involves the danger of misinterpreting small differences in the numeric value of an indicator in terms of differences in performance or in quality. In many cases, data are insuffi-

ciently precise to establish clear cut and unambiguous table positions in a reliable way. Or, to put it in statistical terms, such a procedure ignores the existence of standard errors in data.

Hence the CHE ranking orders universities only into three groups for each indicator: a top, a middle and a bottom group. There is no additional distinction made within groups; in all publications universities are ordered alphabetically within groups—so there is no league table.

Indicators

The set of indicators is crucial to rankings. Rankings can be distinguished according to the data sources to which they use and to the quality (relevance, validity) of indicators. Indicators should be relevant to the target group (data from student survey) of rankings. In an almost two year preparatory phase CHE tried to identify relevant indicators with the help of an advisory board (including evaluation experts and members of professional and university associations) and by group discussions with school leavers and students.

Out of this process a ranking model was derived containing nine components relevant in the decision process. Each component comprises several indicators—all in all between 20 to 30 (depending on the field). The components range from general information on towns (e.g. mean rents) and the university (size, year of foundation, type), student characteristics, central issues of courses and teaching, some aspects of employability, research and labour market to some overall judgments made by professors and students.

The nine components and its indicators are as following:

(a) Students: Ratio applicants/places.

(b) Study outcome: Percentage of graduates completing their degree within the norm time; Average time to degree; Graduation/drop-out rate.

(c) Internationalisation: Index of internationalisation of prgrammes, international (degree and exchange) students, internationality of academic staff, teaching in foreign languages; Support in studying abroad (data from student survey).

(d) Research: External research funds per full-time equivalent academic staff; No. of PhDs per professor; Number of publications per full-time equivalent academic staff; Citations per publication; Reputation with regard to research (data from professors survey).

(e) Teaching & Learning: Average time to degree /Percentage of graduates with norm period of completion; Student-staff-ratio: students per professor; Teaching quality (data from student survey); Organisation of programmes (data from student survey); Contact to students (data from student survey); Support by teachers (data from student survey).

(f) Facilities: Rooms (Lecture halls etc.) (data from student survey); Libraries (data from student survey); Laboratories (data from student survey); Employability, Labour Market: Career orientation and practical relvance of courses (data from student survey, data from graduate survey); Existence of career service/center; Preparation for labour market (data from student survey, data from graduate survey); Teaching of problem solving skills (data from graduate survey); Teaching of transfer competencies (data from graduate survey); Teaching of "learning to learn" (data from graduate survey); Teaching of team skills (data from graduate survey).

(g) Overall views: Overall satisfaction of students (data from student survey); Reputation of universities (data from professors survey).

(h) University and City: Size of the university; Local costs of accommodation for students; No. of places in student dormitories.

In addition to those indicators that are ranked into groups CHE ranking includes a number of qualitative, descriptive information about the profiles of institutions.

Publication of Results

Results of the CHE ranking are published in co-operation with a media partner in order to have a broad dissemination of results. Results are published both in their regular print edition, and in a separate study guide for students and online.

The most important feature of the web

version is an interactive, personalised ranking (called "my ranking") in which—according to our basic approach of a multi-dimension ranking that does not give general weights to indicators—the user can select up to five indicators according to his/her own priorities and preferences. In a second step users can decide for each of the (five) indicators if they want to have retrieved only those universities which are in the top group, or those which are at least in the middle group or all universities. With this instrument users can produce their own selection of universities based on their personal choice of indicators. As those personalised lists normally differ substantially depending on the selection of indicators (e.g. more focussing on teaching and learning versus research) this instrument can identify specific profiles and strengths and weaknesses of the institutions in all field.

Conclusions

The CHE ranking has developed a particular methodology that was appraised very positively by several comparative studies on rankings. This approach is distinct from ranking mainstream as it refers to fields/programmes instead of whole institutions. It is multi-dimensional and avoids the over-simplification of calculating a single composite indicator out of weighted indicators and avoids exaggerating differences in performance inherent in league tables by ordering universities into three groups. This approach, together with the interactive and individualised way of presenting the results in the web version, gives detailed insights into strengths and weaknesses, which means the profiles of departments and hence serves both the need of prospective students helping them to find the best university for them as well as the need of the faculties/departments and researchers themselves to compare with other institutions.

The Perspektywy Ranking: the Art of Evolution

Kazimierz Bilanow and Waldemar Siwinski

Kazimierz Bilanow: Coordinator of International Events at Perspektywy Education Foundation;
E-mail: k.bilanow@perspektywy.pl
Waldemar Siwinski: Founder and President of Perspektywy Education Foundation;
E-mail: w.siwinski@perspektywy.pl

Since the automobile was invented, every car has been equipped with a chassis, engine and wheels. Yet in today's cars none of these elements resemble the original. Though they serve the same basic functions, they have all undergone profound evolution. The best car of 1950 will not be the best car of 2010. Likewise the founder of the Perspektywy ranking, Waldemar Siwinski an engineer by education, understood that every instrument must be constantly updated to keep up with technological progress and expectations. The Perspektywy ranking owes its strong position in the Polish higher education market to years of adherence to good practice and continuous evolution.

The Origins

The origin of educational rankings in Poland goes back to 1992 when Perspektywy published the first ranking of secondary schools in Warsaw. It had instant success, and all copies of the Perspektywy magazine were sold out immediately. After ranking secondary schools in Warsaw, and later in Poland, for several years, Perspektywy decided to move onto ranking universities. In the early 1990' the situation in higher education created demand for such ranking. Following changes in its political system and the demographic boom, Poland witnessed an unprecedented outburst of activities in education in general and higher education in particular. The birth rate grew from 350,000 in 1982 to 720,000 in 1984. The demographic wave first appeared in secondary schools to hit higher education institutions like a tsunami in late 1990s. Poland experienced an unprecedentedly rapid, and unique on a global scale, expansion of the higher education sector. In a period of 15 years, the number of students in Poland had increased fivefold, from 380,000 to two million. To meet the demand 35 new public and 380 private higher education institutions were established.

Despite the fact that secondary school ranking had gained acceptance and popularity, it was clear from the outset that academic ranking was a far more challenging field that cannot be entered without preparation. Perspektywy needed a full two years to get ready.

First we had a series of meetings with the rectors of major universities, and the leader of the conference of rectors. At these meetings the construction and basic principles of the future ranking were discussed in details. The meetings were fruitful and made us aware that the universities expected Perspektywy to rigorously adhere to the agreed-upon principles. To ensure impartiality, independence and transparency a non-profit Perspektywy Education Foundation was established. This foundation has become responsible for selecting the criteria, the indicators, and all other matters related to ranking, while the Perspektywy Publishing House took over responsibility for publishing and marketing results.

Selecting Criteria and Methodology

Although the job of ranking may strive to be as neutral as possible there are subjective elements built into the process. The choice of the criteria and the weight assigned to them is just such an area. Perspektywy had to choose a group of people who would determine its criteria, and decided to rely on experts in the academic world. A special ranking board called Kapitula in Polish (from latin Capitulum) was set up. This board is composed of highly respected academics (former rectors) and other well re-

garded people in the public sphere. Their knowledge and experience helps in selecting proper methodology and leads to confidence in the results.

Our founder's experience in the engineering world tells us that every measuring tool, such as ranking, must be correctly scaled. In order to ensure that our first university rankings would not deviate too far from the current opinions of the time, we decided to take a conservative approach of giving prestige a relatively high assessment of 50%. Two surveys were conducted, one among academic staff and the other among employers. From the outset, however, it was decided that in the years to follow the weight given to prestige and peer review would gradually go down.

Since Perspektywy holds the belief that transparency is important in ranking, the magazine not only published the results of the first university ranking but also revealed details of their process, including a list of almost 1,000 professors who were part of the peer review survey.

The first Perspektywy University Ranking consisted of three criteria: prestige (50%), scientific strength (30%) and study environment (20%). These three criteria were measured by 15 indicators (see "Perspektywy—Ten Years of Rankings" Higher Education in Europe, Volume 27, no. 4, 2002). Since its first issue in 2000, Perspektywy has gone through a slow but continuous evolution during its first decade. Indicators, criteria and sources of data have been refined and modified to provide ever more accurate and true picture of the higher education in Poland.

In 2005, internationalization as a new criterion was introduced with the weight of 10%; at the same time the joint weight of criteria representing prestige was reduced to 30%, and weight of scientific strength increased to 40%. More changes followed. In the Perspektywy ranking 2010, a new criteria group is added—innovation—with the weight of 5%. Consequently the weight of prestige was reduced to 25% by 2010.

"Soft" and "Hard" Data

The adjustments in the importance of the initial criteria groups and introduction of new groups have been followed by an increase in the overall number of criteria from 15 in the first ranking to 32 used in preparing the lastest 2010 ranking. The growth in the number of criteria on the one hand reflected changes in the Polish higher education but on the other hand resulted from constant increase of available and reliable data that could be used for ranking purposes. The latter were the cases of internationalization, introduced for the first time in 2005, and bibliometric data introduced in 2009 ranking with the use of the SCOPUS database.

The general trend in the evolution of the Perspektywy ranking has showed decreasing weight of the "soft" data (gathered through questionnaires) and increased reliance on "hard" data available from ever more comprehensive electronic databases that contain detailed information on publications and citations. Even the "soft" data like employers' preference was gathered in a form of a professional survey conducted by an independent company Pentor Research International. Introducing new criteria and relying on new available date Perspektywy always strive to assure continuity that allows for year-to-year comparison of results.

The mathematical calculation of a ranking is quite simple; in most cases, calculations are based on the multi-criteria analysis (MCA). Far more difficult, however, is the process of gathering necessary data needed for the calculation. Designing its "ranking philosophy", Perspektywy followed the example of the US News & World Report Ranking, developed by Bob Morse and his team.

Five Rankings in One

Recognizing the diversity of Polish higher education institutions, both public and private, and differences in their missions and goals, Perspektywy, in addition to the overall university ranking, also prepared ranking within various groups of higher education institutions. Perspektywy University Ranking 2010 in fact is not just one league table, but it consists of five rankings: ranking of academic institutions which includes all universities and other higher

education institutions that have the right to bestow a PhD degree, with the exception of artistic institutions; ranking of private higher education institutions that have the right to give a master degree; ranking of private higher education institutions that can give a bachelor degree; ranking of public higher vocational schools; and ranking of fields of studies.

In preparing all the above sub-rankings, the same five groups of criteria are employed, that is, "prestige," "scientific strength," "study environment," "innovation" and "internationalization"; though the weight attributed to each criterion differs somewhat from one sub-ranking to another. In the 2010 Perspektywy ranking the University of Warsaw won the title of the best Polish university and Kozminski Academy was named the best private higher education institution in Poland.

The Perspektywy rankings do not only passively report on the state of Polish higher education institutions, but also contributes to the quality of higher education. Currently Polish higher education is far behind the other European countries in the aspect of internationalization. Since "internationalization" was first introduced in the Perspektywy ranking in 2005 as one of the criterion (with value of 10% first, and later 15%), there has been an increase in the number of English-taught programs, and a significant growth in the number of international students in Poland. Due to their great interest in their positions in the consecutive Perspektywy rankings, the higher education institutions closely follow the changes in criteria and tend to apply the criteria, if possible, to improve their positions. Since the criterion of innovation has been incorporated into the ranking just recently, it may take some time before its impact can be noticed.

Good Practice Pays Off

The Perspektywy university ranking has been published consecutively for the past eleven years. A number of other rankings have been published in Poland over this time but none has lasted long. What makes the Perspektywy ranking successful while others have failed? The Perspektywy ranking is believed to derive its strength from adherence to principles of good practice, many of which are rooted in the Berlin Principles on Ranking of Higher Education Institutions: transparency of ranking methodology and consultations with, stakeholders, including rectors and students; supervision over the ranking process is in the hands of an independent ranking board (Kapitula) consisting of renowned academics and people of high public standing; evolution of methodology to ensure that consecutive rankings reflect changes in the higher education field; consideration given to differences in missions realized by groups of higher education institutions, and publication of sub-rankings that take into account specific featuresof each group; publishing detailed data on ranking, including data related to the secific field of higher education institutions' activities, accompanied by analyses and commentaries on the criteria used in ranking and their strengths and weaknesses; securing partners that can help bring the results to the general public ("Rzeczpospolita," for example, is a major daily newspaper in Poland that is licensed to distribute a natinwide special insert with ranking results).

Through adherence to these principles, the Perspektywy ranking has enjoyed broad acceptance among the general public and within the academic community in Poland.

The Perspektywy team understands well that the Polish higher education is not an isolated island, but part of the European and global higher education space. In order to keep abreast with the ranking activity worldwide, the Perspektywy Education Foundation takes an active part in the international ranking expert group and serves as the secretariat of the IREG Observatory on Academic Ranking and Excellence.

"The University Ranking" of Asahi Shimbun Publications

Testuo Kobayashi

Editor of "University Ranking", Asahi Shimbun Publications Inc, Japan

E-mail: kobayashi-t7@asahi.com

Introduction

Asahi Shimbun Publications Inc first published "The University Ranking" in 1994, and has just published the 17th issue this year.

"The University Ranking" is designed to provide broader information of Japanese universities to the students preparing for the entrance examination.

In Japan, high school students usually make their university decisions by checking the minimum scores that they need to enter the universities. In fact, all universities in Japan are ranked by the scores of standardized tests that are provided by the commercial school enterprise in preparation for universities. The universities ranked in the top twenty are considered to be the most prestigious universities. Other aspects of the universities, such as their education and research activities, traditions, and histories, are also taken into consideration by prospective students, however, the ranking by test scores has been given priority for a long time.

In this context, a different system of the Ranking from different points of view is proposed to provide more varied information to high school students. For instance, the universities are ranked by the number of academic papers published by faculties of each university, by the number of papers cited in other papers, and by the amount of research funding given by the Japanese government.

However, at the beginning of the 1990s, the idea of the Ranking and its methodologies were not accepted by the public in Japan. So, universities were unwilling to provide information regarding on their performance at the very beginning.

Yet, the situation has changed. For the last decade, universities in Japan have faced a big challenge. Japanese universities are expected to take more responsibility in society. Increasing evaluation has been conducted on universities' contribution of education and research to the society. Universities are criticized that their education and research are not implemented in a competitive atmosphere; that they have weak linkage with industrial world; that their fundamental research is not enough; that they have produced only a quite limited number of Nobel Prize candidates; and that their students do not contribute to the community enough.

To solve these problems, the Ministry of Education and Science, universities, and the commercial companies are concerned about the future of Japanese universities, and agree to propose a new policy. The policy is to introduce the principle of competition into universities in order to revitalize universities' education and research.

Under these circumstances, the role of "The University Ranking" also dramatically changed. At the beginning, various information were mainly provided for high school students. But now, the more basic data for universities is collected to evaluate their own quality rather than for their customers.

The basic data includes: the number of the papers produced in universities; the number of the papers quoted in other papers; the amount of the research funding given by the Japanese government; contribution to society.

"The University Ranking" allows the comparison among universities by different criteria. For example, one can compare university "A" with university "B" in terms of the number of papers quoted in academic journals of physics.

Contents of "The University Ranking"

"The University Ranking" is based on the following three aspects of data: education, research, and contribution to society, all of which are universities visions in general.

Education: As for the faculty and facility, indicators used in the ranking include the number of students per teaching staff, the number of personal computers per student, the number of books of the libraries per student, the quality of the cafeterias and restaurants and the facilities and services for the physically challenged students, such as elevators, slopes, books in Braille, document enlargers, hearing aids, note takers, and so on.

As for the contents of education, a survey on students' satisfaction is adopted. There are many opinions such as "satisfied with the classes" and "lectures could be better." Also, questionnaires are sent to high schools to ask which universities they recommend for the students. Furthermore, questionnaires are sent to the human resource divisions of 300 companies. The companies are asked which university graduates have good professional knowledge, sales management skills, and planning abilities.

Research: Indicators on ranking research include the number of papers produced by scholars, the citation, the amount of research funding provided by the governments or other industries, and the number of joint-research projects of universities and companies. Special attention is paid to the joint research projects because the national government strongly recommends to increase the number of joint research projects.

The Contribution to society: What is the contribution of scholars to society? The scholars are expected to present their opinion in public based on their professional knowledge and insights to enlighten the people. To indicate this, a few indicators and factors are taken into consideration, for example, the frequency of appearance in general public media, the frequency of seminars open to the community, the number of acquired patents, and the number of memberships to governmental councils. The data is also collected as to how often scholars write articles for newspapers, opinion journals, magazines, and other publications.

It is very important to check out the frequency of appearance in general public media and the numbers of open seminars and lectures. These data shows the scholars' contributions to enlightening society. For example, through explanations of the historical and political backgrounds surrounding the terrorist attack in the United States on September 11th and the aerial bombing in Afghanistan frequently in the media, people might understand the details better than before.

The number of academic staff who are involved in governmental councils is also a good indication for the contribution to society. There are many types of governmental councils to discuss new policy-orientation in Japan. The ranking of membership to governmental councils is very useful to evaluate how much university scholars contribute to build a good society through their specialized knowledge. However, it has been noticed that some council members do not fulfil the expected task, but only justify the policy agenda set by the bureaucrats.

The other ranking indicators include tuition, the average age of academic staff ,finance of universities, the wage, and the alumni such as parliament, professional-sports players or Olympic athletes, and presidents of companies. In addition, the quality of the entrance examination, the universities' Internet web-site, and university-produced marketing literature such as brochures, are also taken into account.

Problems and Future

The biggest challenge in making the ranking is to establish the standard methodologies to evaluate Japanese universities. There are various sizes of universities in Japan. Some of them are very large and include all academic fields, but others are small colleges, and may only have a number of specialized academic fields. The number of students, faculties, and facilities across universities is so different; therefore, it is difficult to evaluate them on the same scale. An approach of overall rankings is not adopted here. It can be very difficult to answer what

"overall" is for Japanese universities.

Above all, assessing the quality of higher education leaves more room for improvement. Fair and objective evaluation methodologies have been sought, so as to evaluate the excellent lectures that cultivate and motivate students. This is a big challenge for "the University Ranking".

The best methodologies have been continuously explored and developed to assess higher education in Japan.

The Globe and Mail: Canadian University Report and the Campus Navigator

Alex Usher

President, Higher Education Strategy Associates, Canada

E-mail: ausher@higheredstrategy.com

In 2002, *the Globe and Mail*, Canada's national newspaper, first published its "University Report Card". Since that time, the publication and its associated web products have evolved considerably—this article traces their development over the past decade.

The Report Card, which as published annually in a separate colour supplement, was initially conceived as a response to Canada's first set of university rankings, published by Maclean's Magazine, a newsweekly. Maclean's rankings format closely parallels that of US news and World Report in that it relies primarily on data derived from a survey of institutions (as opposed to third-party data), supplemented by a reputational survey of administrators and employers, and that the weighting is based to a large extent on measures of financial inputs.

In contrast, *the Globe and Mail*'s product was somewhat different. Instead of trying to use data from institutions gathered via survey or third-party sources, *the Globe and Mail* chose to base its data exclusively based on student surveys. In these surveys, students are asked to evaluate universities on a variety of criteria, including quality of teaching, quality of student services, course availability, physical infrastructure, etc. The new project also tried to distinguish itself by billing itself as a "report card" rather than a ranking; and mean scores on each student-evaluated criteria were turned into a letter grade (e/g/ A, B+, etc). As in many European ranking systems, the Report Card attached no weightings across different categories within the Report Card. However, within the survey students were asked a summative question about how satisfied they were with their "overall experience"; as this usually gets reported first, the scores on this question may be seen as being functionally equivalent to a final weighted score in a more traditional ranking.

By basing its report around student survey results, *the Globe and Mail* was directly targeting undergraduate students (and their parents) as the market for the rankings. There was no attempt to measure reputation externally, nor was there any attempt to look at third party data on research income or bibliometrics. Partly, this reflected a reluctance to play with data with which they had little expertise; but in part it also reflected a genuine belief that institutional research-intensity at any rate was over-rated as a determinant of student experience. Indeed, by using satisfaction as a measure, it became quite clear that research-intensive institutions were at a disadvantage, as satisfaction on a variety of measures is inversely correlated with institutional size. For this reason, although all schools are graded on the same scale and all results on a single question are always shown on a single page, results are nevertheless grouped by institutional size to make comparisons fairer.

To begin with, the University Report Card was a joint project of *the Globe and Mail* and an established market research firm, the Strategic Counsel which was responsible for the survey. In 2005, Higher Education Strategy Associates, which operated as the Canadian office of the Educational Policy Institute from 2004 to 2009, was brought into the arrangement to create a web-based ranking along the lines of the German CHE ranking, which launched in the fall of 2006 and was known as the Globe Navigator. This web-based ranking had 25 indicators (later reduced to 17); users were able to select between one and five indicators on which to rank institutions, and could also filter the

results geographically. Canada being a very large country, inter-provincial mobility among students is still reasonably rare, with only about one in ten students moving to another province in order to study. For most students, truly national rankings are not necessary in terms of choosing an institution as most have a much more limited choice set bound by geography. As with the CHE rankings, institutions were not ranked on each indicator but rather divided roughly into terciles (i.e. top third, middle third and bottom third). A majority of the indicators were—as in the Report Card—based on survey data taken from the Strategic Counsel Survey; however, Higher Education Strategy Associates also included a number of pieces of empirical third-party data (including data on library holdings, research bibliometrics, cost of attendance, athletic team results, etc.) into the Navigator. The Navigator was thus not really just an online version of the Report Card, but provided additional analytics as well.

Prior to 2006, Strategic Counsel had relied on a company called UThink in order to generate student responses. UThink was the consulting arm of StudentAwards.com, a company which provides students with information on scholarship availability in Canada and which in return harvests email addresses and socio-demographic data. Uthink would email invitations to students in its database to participate in a web-based survey on Strategic Counsel's behalf and the latter would harvest the results. This usually resulted in about 10,000 - 15,000 responses per year, which meant that some individual institutional responses were quite small.

With the arrival of Higher Education Strategy Associates, which had closer ties to institutions, a different data strategy emerged. Again, modeling on the approach of Germany's CHE rankings, institutions themselves were brought into the picture. Instead of having UThink to send the survey invitation to students, institutions themselves were asked to do so, and nearly all chose to accept. There were a number of reasons why they chose to become participants in this project. The first was that they were offered free data (each participating institution received a copy of the entire data file for its own students). The second is that they had an interest in seeing the survey have improved response numbers and data quality. The third is that, at a time when there was growing disenchantment with the Macleans' approach to data collection (a mass institutional boycott began in 2006, the open letter which accompanied the announcement of the boycott made special reference to the Berlin Principles on rankings adopted by the International rankings Expert Group in June of that year.), there was also growing interest in alternative approaches to rankings and in particular to the CHE approach advocated by Higher Education Strategy Associates, which seemed to offer a more nuanced and acceptable approach to institutional comparison. With institutions now providing the access to students, annual participation jumped significantly, and total responses are now in the 35,000 – 40,000 per year range.

During the period between 2006 and 2009, there were only minor adjustments to the University Report Card and the Navigator, the most important of which was a re-branding exercise. The University Report Card became the Canadian University Report; the Navigator became the Campus Navigator and became part of a dedicated micro-site at *the Globe and Mail* called "Globe Campus" which launched in the fall of 2008.

In late 2009, Strategic Counsel left the partnership and Higher Education Strategy Associates took over the administration of the survey. A few small changes to the survey occurred at this point: notably, a reduction in the number of questions asked about satisfaction and an increase in the number of questions asked about individual characteristics (including an student's academic performance in university and secondary school and a set of questions which asked students to describe their own personality on a number of axes) and how they perceived their university (e.g. 'more focused in undergraduates vs. more focused on graduate students', 'a curriculum which is more theoretical than applied vs. one which is more applied than theoretical', etc.). These questions are not used for gathering data for purposes of

ranking or rating, but rather to help gain an additional understanding about the determinants of satisfaction. Satisfaction, essentially is about the delivery of services meeting the expectation of delivery. If students with higher grades have higher expectations, then it is quite possible that more selective schools (which in Canada tend to be the larger ones) will systematically have lower satisfaction rates. Evidence for this theory can be seen most clearly on the data for satisfaction with libraries; students at small institutions with tiny libraries nevertheless claim greater satisfaction with their libraries than do students at the highly selective University of Toronto, which has one of the 20 largest holdings in North America. The new data collection allows for more research which might allow for new ways of displaying satisfaction data which could hold constant (for example) student selection effects. A more important shift for 2010 is a move to portray more data on a field-of-study basis, both in the Canadian University Report and in the Campus Navigator. From this year, data on areas related directly to teaching (e.g. satisfaction with teachers, course availability, etc) will be displayed separately for seven broad areas of learning—Arts (i.e. humanities and social sciences), Sciences, Engineering and Architecture, Health Sciences, Education, Business/Commerce and Music and Fine Arts.

Ranking of Higher Education Institutions of Kazakhstan

Sholpan Kalanova

President, Independent Kazakhstan Quality Assurance Agency for Education (IQAA), Kazakhstan

E-mail: skalanova@mail.ru

Introduction

The main objective of ranking of higher education institutions of Kazakhstan is to present appropriate and objective information to various stakeholders, such as future students and their families, employers, governmental, public and international organizations, concerning the quality of higher education institutions, to foster competition among them, and to motivate them to improve.

Ranking of higher education institutions was first developed and published in Kazakhstan in 2006 by the National Accreditation Center of the Ministry of Education and Science. Then, in 2008, the Independent Kazakhstan Quality Assurance Agency for Education (IQAA) was founded. Its purpose is to work in accordance with the Berlin principles, which state that the quality assurance of higher education institutions must be conducted by independent, non-governmental organizations, which can function without influence of the Ministry of Education and Science, higher education institutions and other parties. IQAA developed and conducted the first independent ranking of higher education institutions of Kazakhstan in 2008, and is currently working on its 2010 ranking.

With the increasing popularity of higher education rankings, the Ministry of Education and Science firstly considered and studied about rankings in Kazakhstan in 2004-2005. Such global rankings as the ARWU and others were becoming prominent. The journal of Higher Education in Europe (vol. 27, no.4) published a series of literature on rankings and their methodologies, which had materials from a round-table discussion organized by UNESCO-CEPES in Warsaw in 2002. The literature in this journal issue had a significant influence on further developing the methodology for IQAA, as it emphasized that ranking methodology should be based not only on criteria of universities but also on the national educational system in place . Such criteria were considered to help Kazakh higher education institutions to reach international standards. Dr Jan Sadlak, who, then, worked as the director of European Center of Higher Education of UNESCO-CEPES (current president of IREG) has a considerable influence on our work and supported our methodology. The newly developed methodology in English was first presented during the IREG-3 Shanghai Conference in 2007.

Methodology

IQAA methodology collects data from various sources and adopts a multilateral approach to produce a balanced result. It combines both objective and subjective determinants and includes three parts: evaluation of academic resources of higher education institutions (60%); questionnaires of experts (10%); and questionnaires of employers in governmental organizations and business companies (30%).

The first part of the data is collected from universities. Higher education institutions have to present objective information on academic criteria in electronic and paper format with supporting documents which ensure the trustworthiness of the information presented. The data collection usually takes two-four months, and then the data will be analyzed. The data received from universities is also checked with information available from governmental organizations, higher education institutions' websites and other sources available.

Specific applications and instructions were

developed for evaluation of academic resources of higher education institutions. The format of the ranking has been changed to cut down the number of criteria and indicators in the last several years. IQAA used to be conducted and funded by the Ministry of Education and Science. At that time, more criteria and indicators were adopted and the data could be only collected and determined by visiting universities. Since the IQAA ranking is self-funded, it has to drop the number of indicators for this practical financial reason, while transparency, relevance and validity of its indicators are still ensured according to the Berlin principles. Seven criteria and 42 indicators were used in the previous version of the ranking, while six criteria and 36 indicators are used today. Those criteria and indicators that can be difficult to verify are excluded, so as to make sure the ranking results are objective and transparent. Specifically, the six criteria for evaluating the academic resources of higher education institution consist:

Criterion One evaluates qualitative composition of students and consists of five indicators. The indicators display qualitative composition of students. For example, the proportion of the number of students studying full-time to the overall number of students, including those studying part-time is measured. It is due to the fact that some students choose to study part-time, however, there is a concern with the quality of education offered part-time in comparison to full-time education. Also, the number of students studying on a governmental grant is counted. Kazakhstan has special governmental educational grants which are awarded to future students on the basis of United National Testing conducted by the Ministry of Education and Science prior to the admission of students to higher education institutions. Students, who receive a governmental grant, make a choice in which higher education to study. They may choose either a state or non-state higher education institution. Hence, private higher education institutions can have a chance to attract some governmental funding.

Criterion Two evaluates the selectiveness of students' admission and results of student educational achievements based on five indicators. A big role here is played by the Intermediate Governmental Control Test by the National Center of Testing of the Ministry of Education and Science, which students have to pass after finishing the second year of bachelor studies. Another indicator, which displays the level of education in higher education institutions, is students' achievements in national and international competitions.

Criterion Three evaluates the qualitative composition of faculty and consists of five indicators. This criterion displays various indicators, such as the relative number of faculty who hold scientific degrees, the relative number of faculty who received the governmental award "The Best Faculty of the Year". This competitive grant was set up by the government in 2005, to offer grants for faculty to research and work abroad. There are other indicators measuring the number of faculties who work full-time at a higher education institution and those who received other governmental awards.

Criterion Four evaluates research and innovational potential of a higher education institution and consists of eight indicators. Significant attention is attached to research funding. One of the indicators measures research funding for one faculty. The other important indicators include the number of scientific publications in Kazakhstan and other countries of the Commonwealth of Independent States (former USSR countries) per faculty and publications in foreign journals per faculty. Index of citations is not yet used in the ranking. Also, indicators of the number of patents, the number of dissertation councils, the presence of national laboratories and techno parks are included.

Criterion Five evaluates the level of international cooperation. A few existing indicators have been kept, such as the relative number of international exchanges of students and faculty, graduates who were granted the governmental grant "Bolashak" (the governmental grant, which allows students to receive education at top universities worldwide), the number of educational programs with foreign universities. A couple of new indicators were inserted, such as the percentage of foreign faculty invited to teach in universities and others.

Criterion Six evaluates the level of informational resources of a higher education institution. It consists of five indicators and five sub-indicators. The importance of this criterion is based on the websites of universities and their content. Indicators of this criterion determine the informational base of the websites, the frequency of updates, number of links to the website and visits of the website. This criterion used to be a part of Criterion 4 in the previous version of IQAA. However, due to the growing importance of Internet use and online informational resources of higher education institutions, it is necessary to concentrate on this criterion in particular. .

The second and third parts of the IQAA ranking are based on questionnaires of experts and employers, whereby all higher education institutions are evaluated by experts and employers.

The second part of the ranking is based on evaluation by experts in the field of higher education. Experts are chosen on the basis of their qualification and experience in accordance with the set requirements. There are five possible grades awarded to higher education institutions from "excellent" to "unsatisfactory" in the questionnaire. Instead of a satisfaction questionnaire to students and faculty, a questionnaire to experts in higher education sphere is adopted, due to the huge expense to visit to each higher education institution and conduct questionnaires of students and faculty.

The third part of the ranking is based on the evaluation of higher education institutions by employers by the qualification level of their employees. There are also five possible grades awarded to higher education institutions from "excellent" to "unsatisfactory" in the questionnaire. Questionnaires to employers are distributed among National Ministries, regional governments, national companies and representatives of small business. Telephone and e-mail questionnaires are also used.

Results of all parts of the ranking are summed up, in accordance with which universities' ranking positions are determined. The ranking system includes a general ranking, where all higher education institutions are included, and ranking by fields, where higher education institutions are ranked by groups. Besides the general ranking and ranking by fields, the other rankings by academic criteria, experts and employers evaluation are published. There are 132 higher education institutions in Kazakhstan today, 42 of which are state and 90 non-state higher education institutions. Higher education institutions which tend to participate in rankings are state higher education institutions and strong non-states higher education institutions. Weaker higher education institutions are more reluctant to participate in rankings because of the fear that they may display poor results. Consequently, stronger higher education institutions are able to attract more students and international grants, while weaker universities continue to struggle.

Our methodology has been improving and updating from time to time, although drastic changes are not welcomed. One of the influences for change is the cooperation fostered by IREG among ranking agencies by means of various conferences staged. For example, in the previous methodology, the percentage awarded for the quality of websites was raised by adding several more indicators, and also indicators for international cooperation were attached more weight to develop and promote internationalization in higher education institutions. Results of the Ranking are published in national newspapers such as "*Liter*" and "*Express K*" among others, and are also published on the website of the agency (http://www.iqaa.kz). Thus, the ranking results are available to whoever is interested in. Currently, rankings are popular among future students, governmental organizations and various public organizations.

Conclusion

Kazakhstan's current educational strategic plan of development includes the aim of several higher education institutions in Kazakhstan making the way into rankings of world-class universities. This adds to a high interest in rankings in Kazakhstan today. This is partly due to the development and publication of institutional rankings of national higher education institutions.

上海交通大学高等教育研究院简介

为了加快综合性、研究型、国际化的世界一流大学建设进程，上海交通大学于 2007 年 12 月 26 日在原高等教育研究所的基础上，成立了学校直属的"上海交通大学高等教育研究院"（简称高教院）。高教院下设世界一流大学研究中心、研究生教育研究中心、科学与技术政策研究中心以及《国际高等教育》编辑部。

高教院坚持以实证和定量研究为特色，以"世界大学学术排名（Academic Ranking of World Universities）"为标志，在高等教育研究领域已经具有良好的国际声誉；以教育部科技委《专家建议》为标志，在政府咨询方面已经产生广泛的国内影响。

以小规模、创新型、国际化的世界知名人才培养基地为目标，高教院致力于培养高等教育学、科学与技术管理等领域的专业人才。在校硕士生、博士生和在站博士后总人数的规划目标为 80 名，且大部分博士生有出国访学经历，留学生比例达到国际先进水平。

以小规模、高水平、国际化的世界知名社会科学研究基地为目标，高教院致力于构建具有国际竞争力的学术研究团队，有选择地追求卓越，在世界一流大学研究、研究生教育研究、定量科学与技术政策研究等领域打造国际学术品牌。

GSE 1985 Graduate School of Education Shanghai Jiao Tong University

The Graduate School of Education (GSE)，Shanghai Jiao-Tong University was established in 2007. GSE aspires to the highest standards of scholarly and professional practice, emphasizes quantitative research and international comparison, and focuses on applied research, policy studies and consultation.

GSE builds on the tradition of excellence of the former Institute of Higher Education, which was established in 1985. The "Academic Ranking of World Universities (ARWU)", which was first published in 2003 and updated annually, has gained international reputation. Consultation reports on strategies for building world-class universities and science policies, which are provided to Chinese governments on a regular basis, have attracted national attention.

GSE focuses on higher education and specifically on research universities in the context of national and international development. It has three research groups, namely the Center for World-Class Universities, the Center for the Study of Graduate Education, and the Center for Science and Technology Policy. The centers will have world-class research teams, leading international partners, and a highly internationalized student body.

GSE will uphold the best academic values-a commitment to academic freedom and original inquiry, service to the academic community in China and beyond, and provide the best possible education to graduate students in higher education and science policy.

Vol. 3, No. 4, October, 2010

Guoji Gaodeng Jiaoyu

Global Rankings

National Rankings