

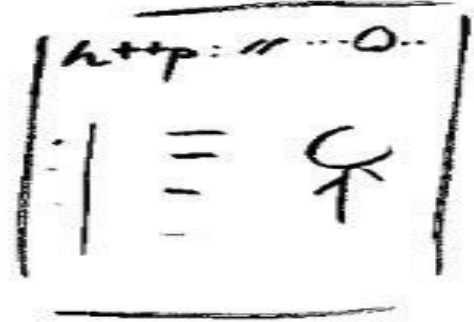
从开放获取到开放科学

CALIS 上海
2015-5-12
吴建中

是于拥有



是于知识

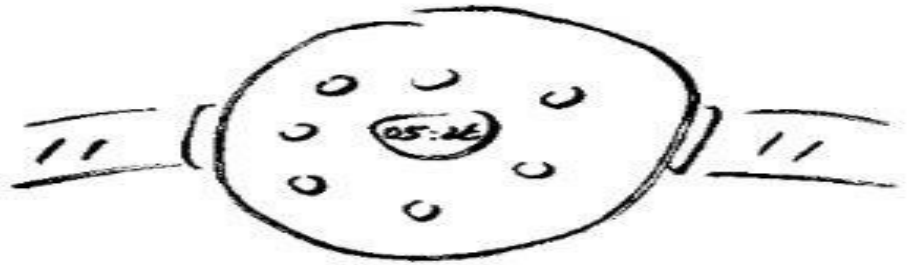


是于知识

不
不
不
不



是于知识



知识是流动的

是于知识

21 Apr. 2015

开放获取的缘起

- * 知识（数据）资源有意或无意地被垄断在少数人手里（包括政府、商业出版部门）
- * 大学图书馆买不起资源了，哈佛大学呼吁，实现开放获取，让大学回归学术中心的地位。
- * 让流动的知识（数据）释放出来，成为可共享资源。

开放获取

- * 2001年12月，开放获取国际研讨会发布“布达佩斯开放获取倡议”。2003年10月，德国马普学会发起柏林会议并发布了“开放获取柏林宣言”。
- * 实现途径——金色道路（**Gold Road**）；绿色道路（**Green Road**），即作者自存档（**Author self-archiving**），机构库（**Institutional Repository**）及各类机构典藏。

开放数据

- * 2013年6月，八国集团首脑在北爱尔兰峰会上签署《开放数据宪章》，提出了开放数据五原则：
- * 数据开放为本；注重质量与数量；让所有人使用；为改善治理而发布数据；发布数据以激励创新。
- * 美、英、法等国政府纷纷发布相应的行动计划，主动作出开放数据的各项承诺，逐步开放数据资源。

数字革命

- * 2014年8月，联合国秘书长潘基文要求独立专家顾问小组(IEAG)为掀起可持续发展领域数字革命提出具体建议。
- * 该小组提交了《一个可追责的世界：活用数字革命促进可持续发展》的报告。



《一个可追责的世界：活用数据革命促进可持续发展》

- * 数字革命涉及开放数据和开放科学两个方面。
- * 开放数据是指将原始的数据及其相关元数据以可以被二次利用的电子格式放在互联网上，让其他方自由使用。
- * 开放科学是将科学研究、数据和传播开放给包括专业和业余研究者在内的所有人的一场运动。

什么是开放数据？

- * 任何人都可以自由利用、发布数据。
- * 按白宫**2014年5月9日**发布的开放数据政策，开放数据并具有公共性、可获取性、描述充分、可再利用、完整性、及时性、可控的再发布。
- * 日本开放数据流通推进联盟于**2014年12月**发布《社会资本信息开放数据化和二次利用促进指南》。

开放科学

- * 每个人都可以成为科学家（市民科学家）。
- * 由微软公司出版的《第四范式：数据密集型的科学发现》一书，探讨在海量数据和泛在网络上发展起来的与实验科学、理论推演、计算机仿真这三种科研范式相辅相成的科学研究第四范式即数据密集型科学发现，以及这种新范式对科学研究、科学教育、学术信息交流及对科学家群体的长远影响。

结论无法验证的论文如何处理？

- * 2012年自然杂志的一项研究表明，临床前药物研究的结果无法验证的论文不少，这些结果无法再验证的数据对科学研究具有不利影响。在科研成果的再现性方面，国际生命科学领域已经把结论无法验证的论文看作是一个严重的问题。

《数据可获取声明》

* 美国科学公共图书馆从2014年3月开始要求数据都要公开，作者在论文开头签《数据可获取声明》，注明数据可在哪里、如何利用。但也认可个人隐私等数据可例外。

The screenshot shows the PLOS website interface. At the top, there are navigation links for PLOS.ORG, PLOS JOURNALS, ABOUT, CONTACT, and LOGIN. The main header features the PLOS logo and the text 'PLOS | BLOGS'. Below this, it says 'Friday, April 24, 2015 | PLOS Biology Community Blog'. There are navigation tabs for Home, STAFF BLOGS, BLOGS NETWORK, and COMMUNITY, along with a search bar and an RSS FEED icon. The main content area is titled 'PLOS BIOLOGUE' and 'COMMUNITY BLOG'. The featured article is 'Opening Up Data Access, Not Just Articles' by Emma Ganley and Jonathan Eisen, posted on February 3, 2014. The article text discusses the importance of data access and provides practical hints for improving public data archiving. An illustration of a person carrying a box labeled 'DATA REPOSITORY' is shown. The right sidebar contains a 'Biologue Search' box, 'Recent Posts' (including 'Stressed to Death: Overcoming Drug Resistance in Malaria Parasites'), 'About PLOS Biologue', and 'Recent Comments'. The footer of the browser window shows the URL 'http://0.gravatar.com/avatar/...' and the page title 'PLOS website.'

《数据可获取声明》

- * 通常作者发表的是研究结果，而把研究过程中的数据给忽略掉了，而且有些数据未必经过严密的验证，即使有同行评议也往往是依赖专家的判断而不是再现。
- * 仅在少数领域，评审专家试图再分析或核实数据或再现数学证明或其他程序中的所有步骤。而只有在文章从发表的期刊上被撤回的时候，人们才会追究评审专家的责任。

从paper到products

- * 在从研究成果的公开和流通向研究数据的再利用转移的今天，**知识的生产 and 再生产过程的基本单元，将从作为研究成果的学术杂志及论文，向作为研究过程的科学数据及预印本转移。**
- * 开放出来的数据不仅是为了得到验证，而且可通过再利用催生新的研究成果。因此，在开放数据和开放科学的推进过程中，**建立和完善确保科学数据流通的信息基础构架是科技创新的当务之急。**

三词箴言+网上发布

- * 法拉第有一个“三词箴言”：Work、Finish、Publish，现在要增加Release（网上发布），实验结果完成后发表还不够，还要把它发布到网上。
- * 一本书或一篇论文发表以后，可以延长其生命周期，作者和读者的对话、跟帖和评论会持续下去；更多内容 and 数据可被揭示出来，并得到再次利用。

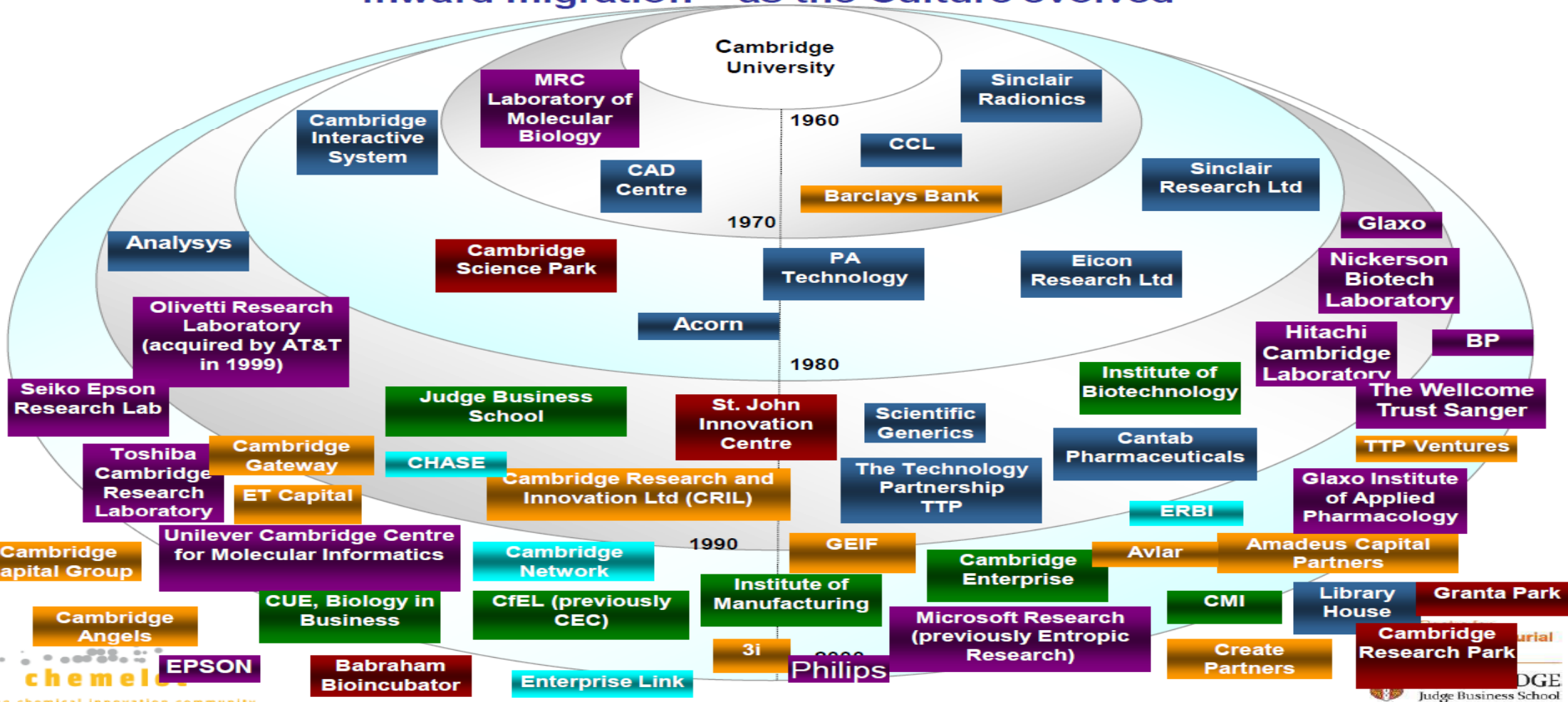
什么是开放科学？

- * 开放科学是学术资源开放获取理念与现代信息通信技术相结合的产物，通过利用现代科技手段促进学术成果的揭示、开放和再利用，形成有利于知识共享、大众创新和经济发展的科研环境

大学第三使命与科技创新

- * 从大学学术中心优势的弱化到学术中心地位（academic heartland）的回归。
- * 随着大学第三使命（教、研、合作）的增强，在推动科学数据开放、合作和共享的同时，吸引了大量企业向大学园区靠拢，形成大学技术园（technopole），如剑桥周边已形成了强大的与大学合作的企业群体。如今以大学为中心的区域创新集群正引领全球经济走出困境。

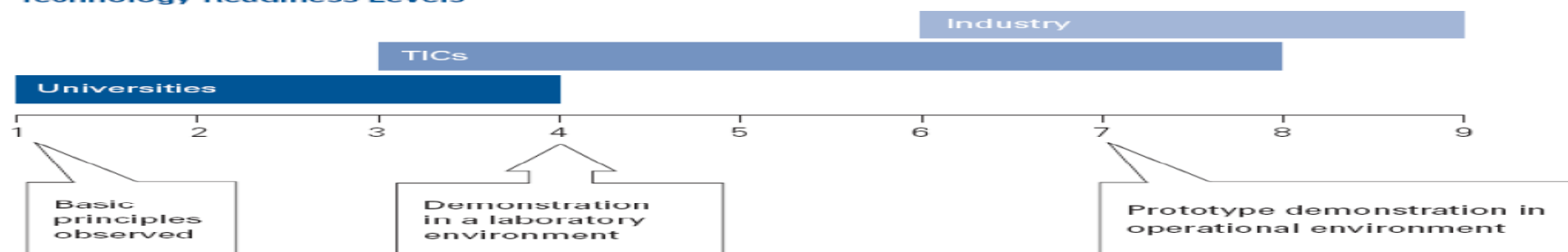
Key Components of Entrepreneurial Eco-system – Local start-ups and inward migration – as the Culture evolved



技术创新中心——TICs

The “Translational Gap” – between R&D and Commercial Products. Bridging the Gap with TICs

Figure 1
Technology Readiness Levels

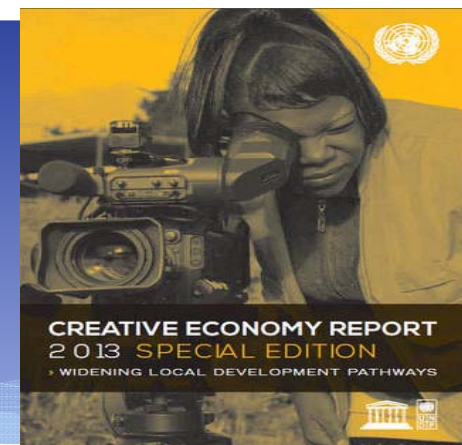


The Current and Future Role of Technology and Innovation Centres in the UK

5

Hauser Report Commissioned 2010
by UK Govt. Accepted and
£250millions allocated to start new
“TICs”

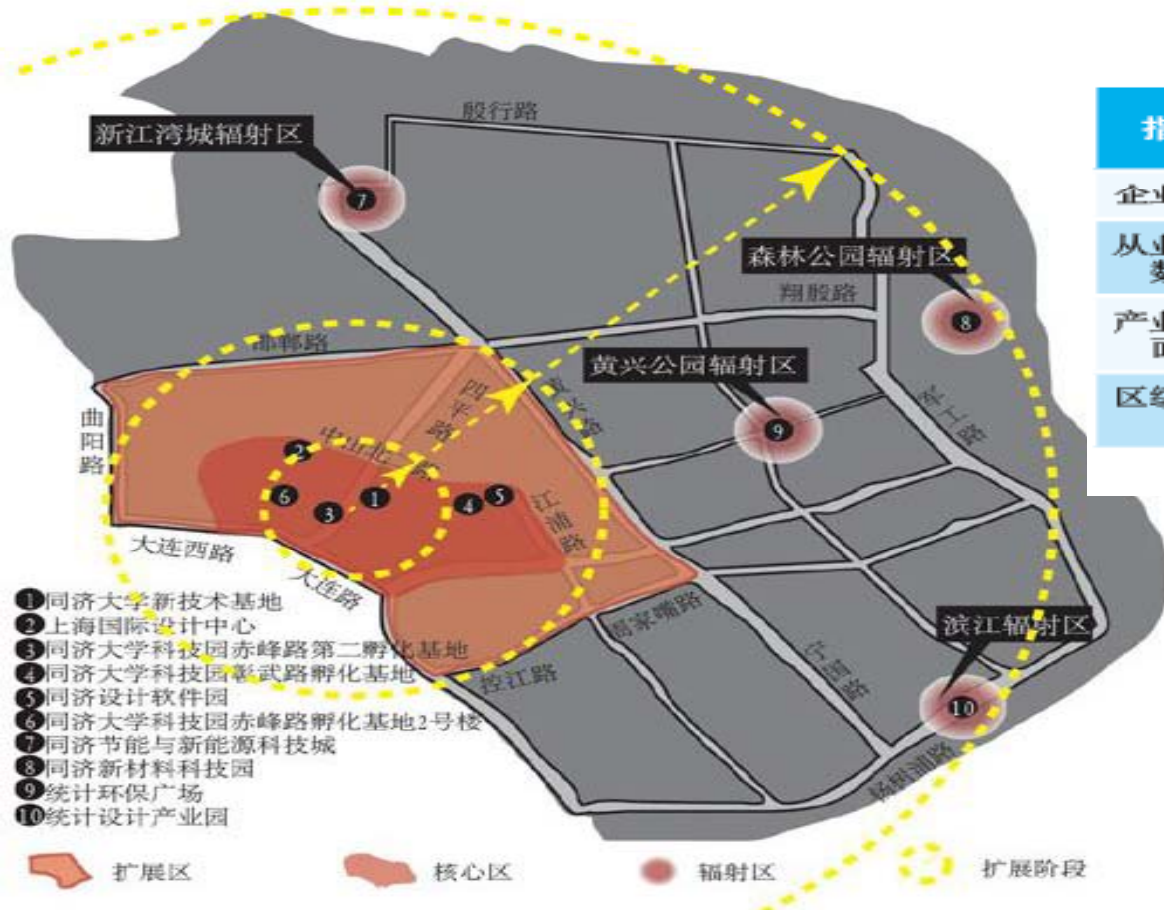
本地动力—从国际模仿到本地创新



- * 《全球创新指数》和由UNESCO和联合国开发计划署发布的《2013年联合国创意经济报告》特别版：扩大地方发展路径，都聚焦地方动力学问题。
- * 报告关注地方层面创意经济的互动、特殊性和政策，以及被纳入地方经济和社会发展计划主流的创意经济如何在社区、城市和地区进行有效的推进。

同济现象

环同济知识经济圈空间分布示意图



环同济知识经济圈占杨浦区的面积和财政收入

项目	环同济知识经济圈	杨浦区	占比
区域面积	2.6 平方公里	60.61 平方公里	约 1/25
区级财政收入 (2010 年)	5.04 亿元 (区级地方税)	50.07 亿元	约 1/10

数据来源:《杨浦区 2011 年政府工作报告》、《上海市场杨浦区环同济知识经济圈“十二五”规划》

环同济知识经济圈“十一五”期间发展情况及“十二五”发展目标

指标	“十一五”初期 (2006 年)	“十一五”末期 (2010 年)	增长	“十二五”目标
企业数量	227	800	252.42%	3000 家左右
从业人员数量	6135	31000	405.29%	—
产业载体面积	28.86 万平米	52.3 万平米	81.2%	100 万平方米
区级地方税	2.08 亿元	5.04 亿元	142%	—

数据来源:《上海市场杨浦区环同济知识经济圈“十二五”规划》

2014年总产出达到264亿元，产出强度之高在全球范围内罕见。

请各位批评指正！

