

开放科学背景下数字学术信息服务新场景

中国科学技术信息研究所 乔晓东

开放科学与数字学术信息服务如同双螺旋中的两根主键，互促互利，交替推进彼此的发展

- **开放科学**：使科学研究中开放透明可访问的知识（包括出版物、数据、实物样本和软件等），通过协作网络共享和发展，使出版和传播科学知识变得容易获取。



知识类型：科技出版物、科学数据、教育资源、软件代码、硬件等等。

关联过程：关联整个研究过程，并倡导越早越好

参与对象：传统科学界+科学界之外的社会行为者

支撑基础：数字技术+协作工具、基础设施

关键目标：参与研究过程的所有参与者之间的开放和无缝合作；社会参与者在研究过程中的有意义参与

1. European Commission, Directorate-General for Research and Innovation, Hessels, L., Koens, L., Diederens, P., Perspectives on the future of open science : effects of global variation in open science practices on the European research system, Publications Office of the European Union, 2021, <https://data.europa.eu/doi/10.2777/054281>

2. UNESCO Recommendation on Open Science. <https://unesdoc.unesco.org/ark:/48223/pf0000379949>

1、开放科学促成新的学术信息生产和传播模式

- 开放科学是由社会对科学知识共享的需求引发的，传统中学术信息的生产、交流和传播主要依赖于图书、期刊、学术会议等，其中，图书、期刊等传统出版形式又是重中之重。在社会对科学知识共享需求剧增的开放科学时代，传统出版为载体的学术信息生产与传播模式需要尽快适应需求，形成新的模式，解决传统模式下在生产、传播中的一些制约。
- 无偏见的快速出版与同行评议相结合的新型出版方式，缩短出版周期，促进最新研究成果的及时发布与传播。传统期刊的出版必须经历论文提交、同行评议、编审、发表等环节，整个周期至少需要2~3个月，常见的则需要半年以上。这种形式影响了知识的快速、及时传播。开放科学时代形成的开放出版模式，可以极大地改进此类问题。以欧洲开放研究计划Open Research Europe (ORE) 为代表的新型科技论文出版平台，通过在同行评议、正式入库检索前增加预印本发布（10日内发布预印本，供查看和引用）等环节，极大地缩短了研究成果面世的周期，提升了学术信息交流的时效性。

1、开放科学促成新的学术信息生产和传播模式（续）

- 开放同行评议，关联数据强制提交，改善研究中的可重复性、获取性、可验证性，提高研究的公开透明。当前，传统学术信息生产与传播中由于部分环节（如科学数据、同行评议等）的不公开、不透明，造成了越来越严重的学术不端与科研诚信问题。开放科学环境下，开放同行评议等新型评议模式将极大改进传统同行评议中的不公正隐患，如ORE中的同行评审中，增加了审稿意见和专家信息一同公布、允许作者回应和读者评论等环节；而传统论文出版所关联的科学数据被要求同时提交，则极大地改善了研究中的可重复性、获取性、可验证性等问题。
- 打破传统的学科边界，促进多学科研究成果的生产与传播，推进不同领域之间的联系、促进合作，以应对全球挑战和可持续发展目标。开放科学环境下，学科之间的边界越来越小，多学科共同合作的研究成果的生产与传播需求越来越大，以Scientific Reports为代表的多学科期刊正在努力推进生产与传播的学术信息的学科多样性。

2. 开放科学促进了新的学术信息资源的产生

- 开放环境，首先在内容生产端发生了重大变革，各种以传播、交流为目标，以各种新型平台为依托的新型“出版”正在颠覆传统信息服务的内容基座，一次知识内容的载体形态、类型、产生环节、产生方式均发生着重大的变化，由此引起的二次、三次内容生产也面临着变革。
- 首先，开放科学推进了出版模式的转变，这种转变不仅带来了预印本快速发展，更是促进论文中的“问题、假设、方法、数据、分析、解读、应用”分解为独立的研究模块进行出版，这些类型的学术信息构建起不同于传统文献资源的新型学术信息资源集合。

研究过程中的各类内容成为学术交流与传播的对象，以Open Research Europe为例，支撑**全领域14类出版**。Octopus一类平台尝试**研究要素**的出版。

研究论文	短评	注册报告	指南出版
视频出版	数据出版	软件出版	方法出版
案例出版	临床案例	同行评议
研究问题	研究假说	研究方法	研究结果
分析过程	结果解释	实际应用	

2. 开放科学促进了新的学术信息资源的产生（续）

- 其次，越来越多的工具类资源、事实数值类资源、新闻政策类资源等也成为了开放学术信息资源的重要组成部分。比如data.gov（开放政务）、CORDIS（开放资助信息）、re3Data（科学数据）、Open Science Grid（计算资源共享）等。新型的开放创新研究所需的诸多类型学术信息在开放科学环境下得到了极大发展，针对这些新型学术信息资源需要有清晰的认识与明确的定位。
- 为了更好地推进多种类型学术资源的生成与利用，需要从工作链的角度，深入研究不同类型的学术资源在开放科学各个阶段的角色和作用，基于此构建更加灵活的学术资源描述框架、学术资源融合策略、保存策略等，从而为开放科学环境下的异质异构资源提供重要的基础。
- 国家重点研发计划重点专项“基础科研条件与重大科学仪器设备研发”中所反映出的问题。

3、数字学术信息服务链条中各环节角色正在发生改变

- 信息技术与人工智能技术的急速发展，快速打破传统学术信息内容出版中发行、传播、服务的壁垒，原有的出版、发行、传播链条发生极大变化，出版即服务的趋势越来越明显，国际大型信息服务商逐渐形成覆盖整个科研生命周期的闭环，实现了从出版商到服务商再到问题解决方的角色转换，对传统知识内容传播模式发起了挑战。

流量概况

搜索引擎



新媒体



流量拓展方法



搜索引擎优化 (SEO)

- 站内优化：内容优化、链接优化、结构优化、内链优化、外链优化
- 站外优化：外链优化、数据提交、友链交换、蜘蛛池、站群

新媒体

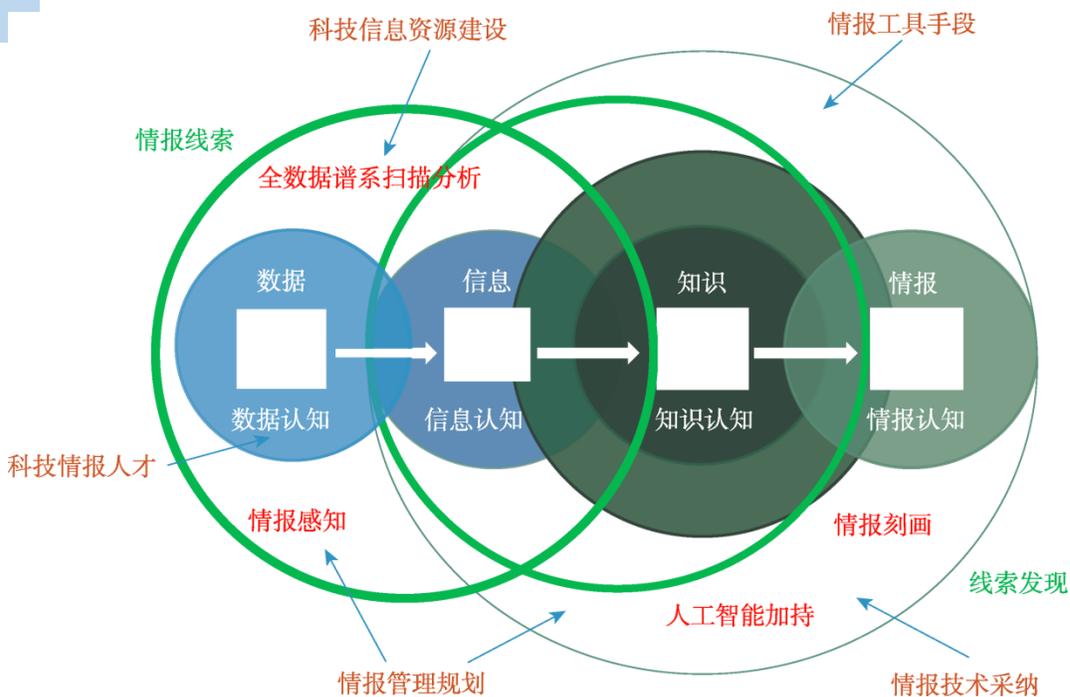
- 平台：历程、稳定性、流量评估、人群画像、流量逻辑、平台规则
- 内容：选题、定题、撰写文章、审核、发布文章、热门词汇
- 引流：话术、钩子、技巧

4、新的网络信息环境

- 复杂信息环境下科技情报理论体系构建：

复杂信息环境包括科技情报“外部”环境和“内部”环境两个方面。科技情报“外部”环境是指我国的外部环境日趋复杂，不稳定性、不确定性明显增加，此种国际环境给科技情报的信息源带来复杂性；科技情报“内部”环境是指科技情报生产所处的复杂信息环境，该环境下的科技数据与传统的科技数据截然不同，它是一种“不稳定状态”的数据，即数据来源、类型、存储结构、规模、质量等均发生了空前的变化，这也导致数据内容更为复杂化。具体来说，复杂信息环境主要涉及国际环境的复杂性、数据和信息的复杂性，科技情报技术基础的复杂性三个方面。

(赵志耘：论复杂信息环境下的科技情报卓智赋能)



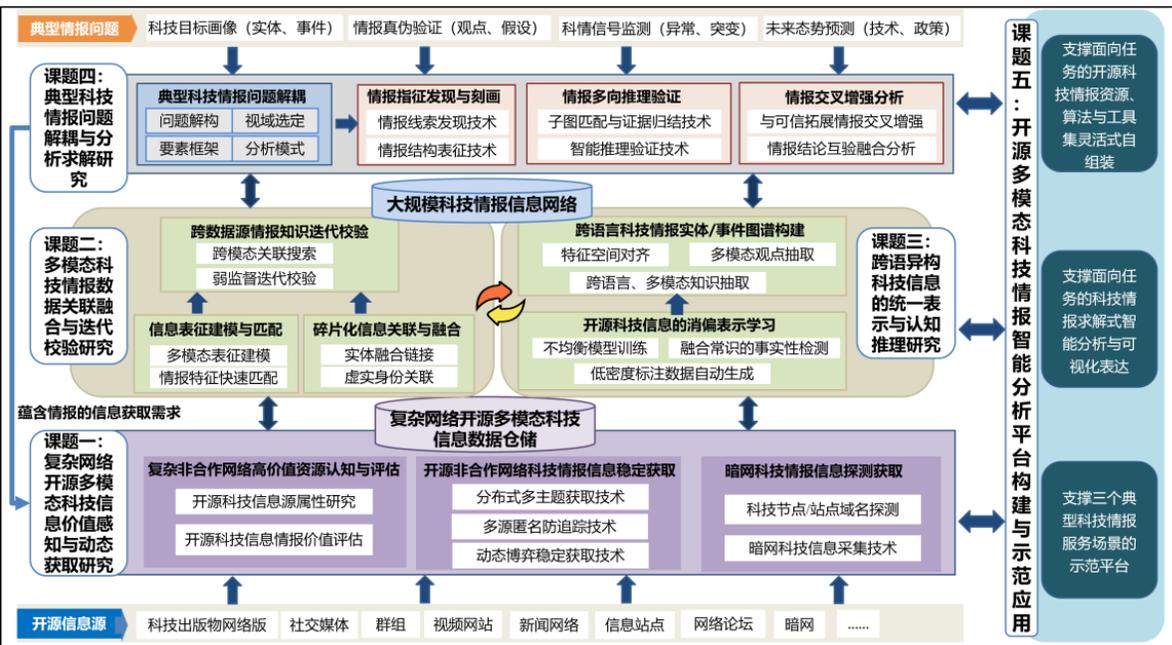
科技情报“卓智”赋能要素及实现路径

4、新的网络信息环境（续）

“网络开源多模态科技情报智能分析”立项教训：

拟解决的重大科学问题：

- 1、在复杂网信环境下，如何计算开源科技信息的情报价值，实现多类型开源信息稳定、安全、可靠、持续、隐匿式获取；
2. 如何实现碎片化、多语种、多模态开源科技信息的情报要素抽取和认知推理，融合形成高价值科技情报信息网络；
3. 在典型科技情报服务场景下，如何实现融合开源情报和传统情报的异质信息自动化加工和问题适配，形成智能科技情报分析新范式。



主要研究内容与课题设置

5、开放科学创造新的学术信息服务需求

- 推动开放科学环境下的新型学术出版，还必须同时考虑建设面向开放科学的新型信息服务基础设施和方法工具，包括但不限于专业性、综合性学术信息出版、评审和开放获取平台；支持学术研究和学术出版的科学数据、图像信息、软件工具等科研要素的登记、保存、揭示和开放服务平台；支持新型开放科学学术产出的科技评价方法工具；支持开放科学的科研诚信建设和学术不端治理信息服务；开放科学信息资源的集成发现和关联、分析、评价和挖掘等增值服务。

5、开放科学创造新的学术信息服务需求（续）

• 传统评价指标的新发展及新评价指标的涌现

典型产品数据库	指标名	可应用评价对象	指标定义	数据来源	适用场景	不适用场景
Cabell's Directories of Publishing Opportunities the Modern Language Association Directory of	Journal Acceptance Rate (期刊接收率)	期刊	出版的手稿与提交的所有手稿的百分比	期刊编辑部与出版商投稿系统	期刊的接受率取决于在特定期刊上发表的相对需求、现有的同行评审流程及发表时间等。因此，与可用性相比，它可能是感知声望和需求的代表。	不应被用作衡量特定手稿质量的标准
Journal Citation Reports	Journal Impact Factor (期刊影响因子)		反映该期刊最近发表的文章的年平均(平均)引用次数的指标	基于Web of Science中所有收录的论文的引用进行计算	仅适用于同学科期刊一级的比较	不可用于评价单篇论文的学术价值或单个作者的学术能力
Scopus	CiteScore		过去3年发表的文献在特定年份的总引用次数除以过去3年发表的文献总时指数的变体，仅对近五年发表的论文进行计算。	基于Scopus所有收录的论文的引用进行计算	仅适用于同学科期刊一级的比较	对发表社论、新闻、信函等前辅文(front matter)的期刊不利
Google Scholar	h-5 index		核心概念来自Google的PageRank算法，测量期刊的声望	基于收录库中所有收录的论文的引用进行计算	该指标对大刊有偏好。谷歌学术发布h5中位数就是为了增加公平性。	
SCImago	SCImago Journal Rank		发表于十年内各领域中被引用次数Top1%的文章。	Scopus数据库		
Essential Science Indicators	Highly Cited	论文/机构/科研人员/期刊	统计各学科10年间每年发表的论文达到某个百分点基准应被引次数	基于Web of Science中所有收录的论文的引用进行计算		这些指标都基于引用，因此它们继承了所有引用的局限性，在评估异质或跨学科的论文组时，可能会遗漏引用密度低的重要子主题。
	Percentiles (百分位)					
	归一化影响因子					
Web of Science	Enriched Cited References	论文	结合引用在全文中的引用意图、位置等形成引用位置、强度、语境和类型指标	基于Web of Science可获得全文的引文进行计算	除传统引文中显示的书目信息，并且提供引文陈述、引文语境、引文在被引文中的位置(如引言、方法、结果、讨论等信息)、引用意图(支持、争议或提及)和编辑信息(如更正、撤稿等信息)等。	
Scite	引用支持指数			从3000余万篇全文文章中提取和分析出来的9亿多条引文陈述，同时还包括超过10亿条原始引用		
Semantic Scholar	Influence Score (影响力分数)		Semantic Scholar获取全文授权的			
data citation index	科学数据引用	数据	科学数据集被应用次数	Web of Science收录的全球400多个仓储平台中的科学数据集及其索引库中的论文	辅助判断科学数据的价值	
Lens	专利引用	论文	论文被专利引用次数	Lens平台	辅助判断论文对技术的价值	
Overton	政策引用	论文	论文被政策应用次数	Overton平台	辅助判断论文对政策的影响	
figshare	软件数据引用	数据/软件等	一段时间内数据、软件的下载、浏览、引用情况	figshare平台记录的用户欣慰	指标用于追踪科学数据、软件的受关注程度	不是评价对象质量的直接衡量标准
	软件数据下载					
	软件/数据浏览					
Codeplex	软件下载	研究软件、脚本、代码片段	一段时间内的软件下载量	平台记录的用户下载行为	可被用作编程代码重用的反映指标。某些情况下它可以是用户数量的代替指标。	软件下载量不是软件使用、质量或影响的直接衡量标准。
Zenodo						
GitHub (API)						
Altmetric	Altmetric Attention Score (Altmetric注意力评分, AAS)	期刊文章、书籍和章节、预印本、数据集以及存放在可跟踪的存储库中的任何其他研究成果	注意力得分是一个自动计算的、加权的研究产出获得的所有注意力的计数	新闻文章、博客、推特、脸书、新浪微博、维基百科、政策文件(按来源)、问答、F1000、Publons、Pubpeer、YouTube、Reddit、Pinterest、领英、开放教学大纲、Google+。不包含Mendeley读者数、引用和CiteULike	个人可以基于该指标了解研究在网上受到的总体关注程度；出版商和机构也可以使用AAS来对其出版或制作的研究所受到的关注进行分组，以监测和衡量其影响力。	AAS不应被用作任何类型或质量的研究影响的直接衡量标准。
Altmetric	News Mentions新闻提及	期刊文章、书籍和章节、预印本、数据集	主流在线新闻站点提及研究成果的次数	来自在线新闻媒体和杂志的全文新闻报道和RSS提要。	新闻媒体的提及可以证明研究的影响力和关注度(电视台)	不应将新闻媒体提及作为质量或影响的直接衡量标准。大多数指标来源2011年或之后开始跟踪，因此更容易获得近期出版物的数据
PhumX Metrics						
Newsflo						
publons	Publons Score	同行评审的论文和会议记录	根据注册Publons用户提交的同行评议，以1到10的等级显示论文的感知质量和重要性。	重要性和质量的评级由Publons上的审稿人、编辑和作者社区评分	Publons Score可用于初步了解一篇论文在研究社区中是如何被接收和评估的。	Publons Score并不总是适合完整的同行评审过程，手稿根据审稿人的评论随着时间的推移得到改进，不应孤立地用作衡量研究影响的指标。

5、开放科学创造新的学术信息服务需求（续）

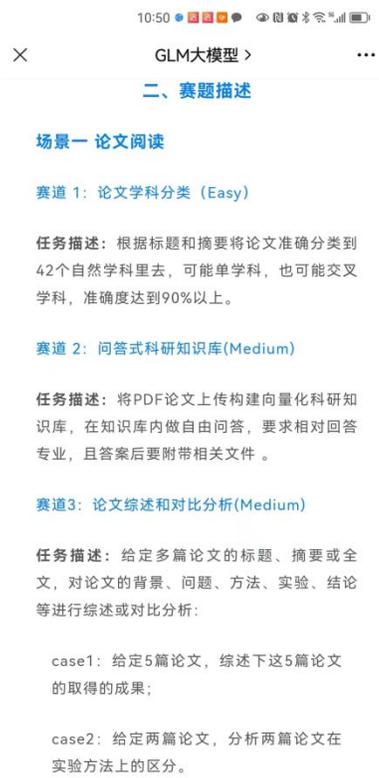
• 助力高校落实科研诚信建设主体责任

- ◆ 责任体系和诚信制度建设
- ◆ 科研诚信素养教育
- ◆ 关键业务点管理和把控
- ◆ 举报受理、调查处理
- ◆ 基础条件建设
- ◆ 信息开放共享
- ◆ 协调、配合与服务
- ◆ ○ ○ ○



6、为核心科技情报工作提供支撑

- 情报理论方法的创新
- 技术应用能力的提升（知识技术、数据技术、人工智能...）
- 学科和行业技术创新与行业发展基础设施和公共产品建设



7、发挥高校图书馆优势，构建多元开放数字科学信息服务新模式

- 鼓励建设多元的开放科学学术出版与信息服务，而非统一模式、统一程度。多比少好，部分开放也比不开放好。
- 开放科学信息服务可分领域、分阶段逐步实施，注重引领、示范作用。
- 高校图书馆可在开放学术资源建设，支持开放科学的信息基础设施建设，开放科学下信息素养培训，开放科学下新型信息服务能力构建等方面发挥更大作用。

8、高校图书馆馆员在开放科学背景下的新角色

UC SANTA BARBARA Recruit

Home · Open Recruitments · Open Access Collection Strategist (JPF02255)

Open Access Collection Strategist

Job #JPF02255
Library / Office of the Executive Vice Chancellor / UC Santa Barbara

We are no longer accepting applications for this recruitment. [Browse open recruitments](#)

POSITION OVERVIEW

Position title: Open Access Collection Strategist
Salary range: \$60,166-\$144,822
Percent time: 100%
Anticipated start: December 2022 or later
Position duration: Indefinite

- 加州大学圣巴巴拉分校“开放获取馆藏策划师”（17项岗位基本职责与责任）--传统的采访馆员职能的扩展：
 - 评估机会，并提出建议，使学术交流系统向研究的开放传播转变。
 - 参与实施创新的馆藏模式和方法，管理UCSB开放获取（OA）出版基金。
 - 通过eScholarship、ORCID和Dryad促进开放出版选择；
 - 监测出版商和供应商提供的信息和数据，以确定开放获取相关的趋势和模式
 - 基于证据的馆藏和内容获取、管理和评估方面进行合作
 -

EMORY UNIVERSITY

Open Access Librarian

Emory University: EUL: Emory University Libraries

Location	Open Date
Atlanta, GA	Oct 19, 2021

April 14, 2021

Publishing & Open Access Librarian

William & Mary Libraries · Williamsburg, VA

- Emory大学图书馆等发布“开放获取图书馆员”（主要推动开放获取和开放教育实践，及提供OA知识库的推广和支持服务，类似的还有“开放教育资源馆员”）--传统信息服务馆员或学科服务馆员的拓展：
 - 开放教育服务：包括帮助教员发现、采用和改编外部制作的开放教育资源；创建和维护开放教育计划的推广和教学材料；制定开放教育计划的目标、政策、最佳实践和评估指标；版权许可教育等；
 - 开放获取存储库服务：涵盖科学数据的科学数据管理服务；开放学术资料库的推广、宣传、培训与教育以及与各类数据资源的整合等。

FLORIDA STATE UNIVERSITY

News & Events
at College of Communication & Information

Home News Academics Events

Reproducibility Librarian @ UF Libraries

Posted by Lela Gibrake on September 18, 2019

The George A. Smathers Libraries at the University of Florida seeks a Reproducibility Librarian to develop an institutional strategy for education and support of transdisciplinary research: reproducibility and open science. This position will be located in the Health Science Center Library (HSCL) in Gainesville. The Reproducibility Librarian is a year-round tenure-track library faculty position. The person who fills this position leads in designing and implementing a multifaceted program to enhance campus-wide efforts to promote and improve research reproducibility from design to dissemination. The Reproducibility Librarian participates as an active member of the library- and campus-wide teams to develop programming for and support information retrieval/storage, data science, and research. The incumbent will provide interdisciplinary information consultation services in a variety of modes, design workshops to promote research reproducibility, perform course-integrated instruction, and participate in the Health Science Center Library's teaching program. The position is responsible for special projects as assigned, such as new service development and evaluation, and development of web-based resources. The librarian works collaboratively in group efforts and maintains professional relationships with faculty, students and colleagues.

- （围绕科学数据，图书馆新职位嵌入科研全流程，参与数据管理活动）
 - 数据图书馆员（基于ALA JobLIST 2006-2018年4月发布的美国学术图书馆相关岗位）：
 - 包括数据素养馆员、研究数据馆员、数据可视化馆员、数据馆员、数据仓储馆员、数据服务馆员、数据与政府信息馆员等各类图书馆岗位名称；
 - 美国佛罗里达大学图书馆2019年招聘“研究再现”图书馆员；
 - 霍普金斯大学图书馆形成了体系化的“数据服务”

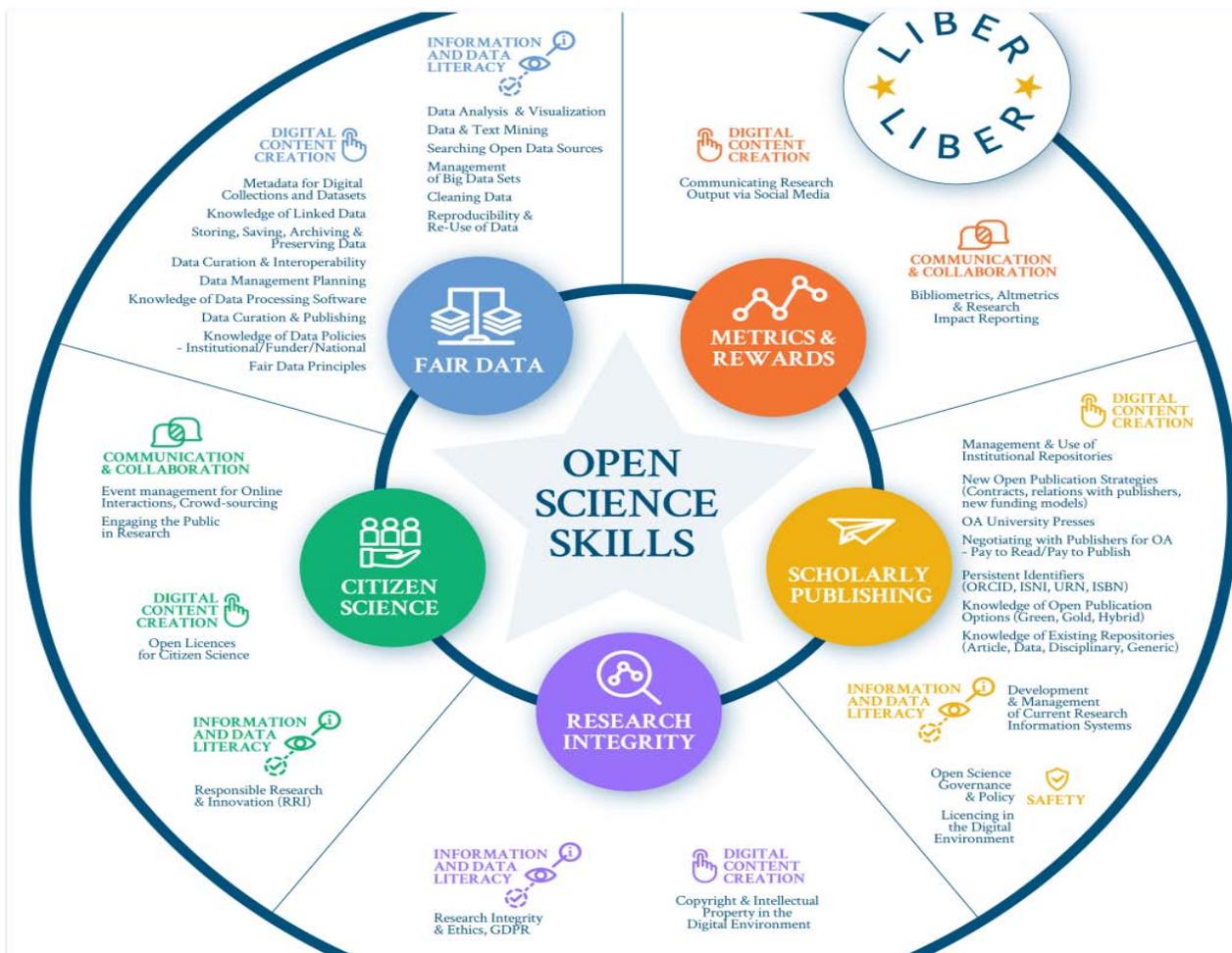
8、高校图书馆馆员在开放科学背景下的新角色（续）

- **传统图书馆岗位在开放科学环境下工作内容有了新的内涵**：如“开放获取馆藏策划师”、“开放获取图书馆员”、“开放教育资源馆员”等的出现，传统的采访工作、学科服务/信息服务需要结合教学、科研的需求，拓展多元化的开放资源的合理利用与集成，同时，需要参与到本校教学、科研资源的对外开放共享建设。
- **开放科学环境下催生了新型图书馆岗位**：科学数据相关的馆员岗位的出现，不同于传统图书馆馆员的角色，需要参与到科研的全流程中，从数据管理计划、数据素养、知识产权、数据开发利用、数据可视化服务等多个维度切入科研生命周期的不同阶段，推动开放共享，提高科研诚信与研究透明度。

8、高校图书馆馆员在开放科学背景下的新角色（续）

新的岗位需要新的技能、工具与服务的支撑，信息服务面临新机遇、新需求：

- FAIR数据：数据分析、数据可视化、数据挖掘、数据清洗、数据管理、关联数据、数据出版等等。
- 计量评价：新型可替代评价、研究影响报告等。
- 学术出版：了解新开放出版物政策、各类永久唯一标识、已有仓储等；
- 科研诚信：数字环境下的知识产权、科研诚信与伦理等。
- 公众科学：公众科学的开放许可等；



9、新的数字学术信息服务业务逻辑

- 资源的数据化（数据治理、知识组织）、数据的工具化（知识计算、方法与算法）、工具的平台化（技术能力与业务模式）、服务的场景化与工程化（开放共享与可持续发展）
- 知识的智慧化、智慧的场景化、智能的泛在化（张晓林）

知识对象已内在地结构化、语义化、可定制、可关联,已涵盖从内容对象到知识关系网络到知识创造与利用流程,成为可交互、可计算的智能体。

通过场景驱动的创新机制,可在用户解决问题的过程中,通过数据、模型、计算和交互来支持感知智能、认知智能和决策智能。

通过泛在智能机制,推动智慧知识对象生态体系的构建与管理、知识与数据双驱动的问题分析与决策推理、多源资源环境下AI赋能的个性化知识服务的动态构建,以有效推动智慧知识服务的普惠可及。

总之，推动开放科学环境下的数字学术信息服务，图书馆和科技信息服务机构应该发挥更好的引领和支撑作用，同时也应鼓励各类学术研究、学术出版和信息服务主体进行多种模式的大胆尝试，推动开放科学中各参与角色承担、落实其主体责任。除了享受开放科学带来的好处外，我们也应是一个贡献者。

开放资源和传统内容服务学术信息资源互为补充。开放科学资源为专业化学术信息服务提供了更多的知识库、实体数据、事实性资源、工具方法等。同样，开放资源也需要深度数据治理、知识组织、关联揭示等图书馆和信息服务机构的专业化支持，需要如资源审核、注册、保存、检测、评价等围绕开放科学资源的外围服务。积极促进各类配套的开放科学基础设施和可持续服务能力建设，大处着眼、小处着手，将开放科学学术信息服务纳入开放科学生态建设的大环境下思考。



不对之处还请批评指正，谢谢。

电话：13301170911

邮件：qiaox@wanfangdata.com.cn