

作者: 李木子 来源: 中国科学报 发布时间: 2014-4-8 10:25:41

选择字号: 小 中 大

美国发布《从数据到发现：基因组到健康》白皮书

2014年3月，美国国家数据科学联盟（NCDS）发布了《从数据到发现：基因组到健康》白皮书，以便于解决基因组学面临的重大挑战。该白皮书是通过聚集数据科学领域与基因组学领域的领袖人物，共同探讨当前基因组学的挑战，并集思广益，就最为恰当且技术先进的建议达成的共识。

基因组学能够为数据科学相关研究提供某些美好的承诺。但新发布的白皮书指出，基因组学领域还存在数据来源、收集和管理，界定表型，裁定基因组变异，生物统计学和生物信息学，数据共享和生物伦理与法律六大关键挑战。

数据来源、收集和管理是在不知晓数据将如何被重新利用的情况下维护数据来源，为大数据集的收集和管理提出重大挑战。而由于缺乏标准化的数据元素与协调的数据集，同时缺乏从大型数据集中获取表型数据的相关技术，也使得界定表型问题进一步复杂化。

由于缺乏表型和变异体数据的标准，对变异的界定也十分模糊。另外，由于统计模型和软件不够充分，计算机处理能力不足，运行复杂模型时不可接受的时间延迟以及在促进数据集成与共享时，采用有限的联合分布式数据系统等，都造成了生物统计学和生物信息学上的重大挑战。

不仅如此，在基因组学中，尚未解决的生物伦理问题也比比皆是，其中包括某些偶然发现，披露基因测试结果以及敏感人群的隐私问题。对于物理性质、知识产权与信息财产的法律区分以及基因组数据的隐私性和机密性的法律区分，仍然需要公开讨论。

为积极地应对这些挑战，白皮书中也制定出关键性建议，力求不断推动基因组学的发展。

第一，在基因组学研究领域，促进跨学科合作并协调相关工作，形成“联盟”并协调各个研究组的科研成果，不断推动相关标准的制定，促进跨学科合作。

第二，推进分析方法和工具的相关标准和联合分布式数据系统的广泛采用，同时协调现有的数据集，综合分析，数据再利用以及进行科学发现。

第三，不断促进数据共享，同时通过激励机制和全新的技术解决方案，实现数据共享的不同技术方法的成本效益分析，维护其隐私性、安全性与来源问题。

第四，开发自动化、易于使用的、利益相关者驱动的、开源的临床决策支持系统。临床决策支持系统可以通过一个简单、合成的方式展示基因数据，反映出全部利益相关者的观点，并整合类似wiki功能，同时使包括授权临床医生和其他利益相关者能够解释与应用基因组的相关研究结果，从而充分地实现个性化医疗的巨大潜力。

第五，基于大数据的信息技术、数字存档与分析的培养教育和培训计划。针对广泛的专业、培训与职业发展，将基本的生物统计学概念纳入现有的培训项目中，能够显著增强科学家和临床医生有效地解释与应用基因组数据的能力。

第六，解决合理使用与滥用基因组数据之间的区别等生物伦理和法律政策问题。（李木子整理）

打印 发E-mail给:



以下评论只代表网友个人观点，不代表科学网观点。

还没有评论。