arXiv:2512.22159v1 [cs.DL] 16 Dec 2025

Oignon：引文图工具

H. Ballington

大气与环境研究所，伍珀塔尔大学，德国

2025年12月30日

摘要

引文图可视化是学术研究中用于了解研究背景的实用工具。遗憾的是，现有解决方案可能存在若干缺点，例如扩展性较差、网络遍历浅、受到免费增值功能限制，以及构建速度缓慢。Oignon 是一款免费的开源工具，可用于系统地探索学术研究。它采用带有时效性加权的双路径排名系统，生成能够呈现与特定出版物相关的奠基性著作和近期突破的图。

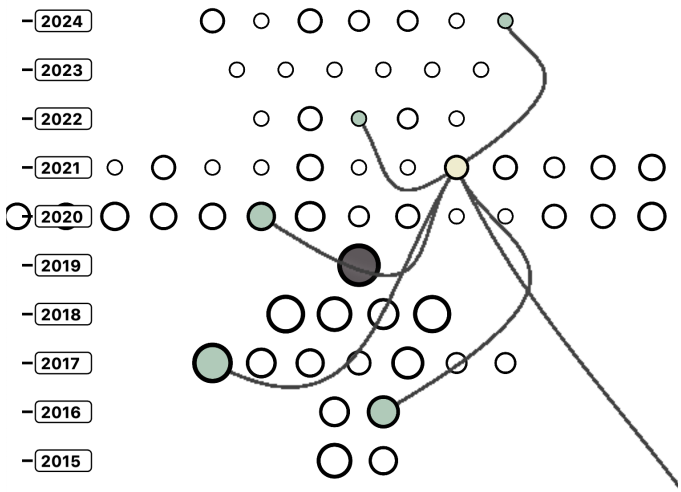


图 1：使用 Oignon 创建的图的子区域示例。

关键词——引文分析、文献计量学、图可视化

1 背景

目前已有 *many* 用于创建引文图可视化的工具，但几乎所有工具都至少存在以下一项缺点（按造成的不便程度排序）：

1. **入门门槛：**尽管 OpenAlex [1] 和 Semantic Scholar [2], 等易于获取的数据免费开放，但许多引文图绘制工具仍设置了人为障碍：强制创建账户、通过免费增值模式限制功能，或要求机构身份验证。

2. **力导向图：**大多数现代工具都将图显示为力导向图，其中代表论文的节点通过迭代式物理模拟进行定位。节点之间相互排斥，而连接则彼此吸引，这自然会形成一种优先呈现拓扑结构而非语义含义的布局：论文的位置除了传达其直接邻域信息外，几乎没有更多含义。这类布局还存在 $O(n^2)$视觉复杂度问题——随着图规模增大，边交叉和节点遮挡会呈二次增长，导致难以系统阅读图的“毛球效应”可扩展性问题。

3. **前向覆盖不足：** 许多工具专注于后向遍历，生成冻结在发表日期的存档快照。前向覆盖往往系统性地偏向较早的著作，因为它们有更长时间积累引文。在这一领域，时间归一化是一个需要纳入考量的重要因素。此外，后向遍历本身更易处理，因为它受益于汇聚结构——参考文献往往汇聚至共同的奠基性著作，从而放大共被引信号。相反，前向遍历具有发散结构，施引论文会分支到各种研究方向中，且共被引信号较弱；这通常会导致图构建速度缓慢，而大多数工具要么忽略这一问题，要么通过截断前向范围来解决。

4. **深度与完整性：** 学术写作的严谨程度与对其所处更广泛学术思想脉络的理解密切相关。识别并致谢奠基性著作和开创性论文，是学术深度的重要体现。现有工具大多无法掌握这类传统上依赖隐性专业知识的内容，通常是因为引文网络遍历不够深入，未能揭示参考文献之间共同的深层根源。

2 Oignon

Oignon（*"onion" in French*）是一款用于创建学术研究引文图的工具。源代码依据 MIT许可证 [3]发布。图中的节点呈现在二维网格上，纵轴表示发表年份。图既可以根据源论文创建，也可以根据作者创建。根据源论文创建时，Oignon 会使用 OpenAlex，通过一个排序系统（灵感来源于 [4]）搜索相关的根节点（更早）和分支（更晚）出版物。

- **根节点排序：** 根节点种子论文定义为源论文的直接参考文献。根节点论文的排名由以下指标简单求和得出：

  – *Cocited count:* 根节点候选论文与一篇种子论文同时被引用的次数（相关性）– *Cociting count:* 该论文与种子论文共享的参考文献数量（相似性）– *Cited count:* 引用根节点候选论文的种子论文数量（重要性）

- **分支节点排序：** 分支节点种子论文定义为直接引用源论文的出版物。分支论文的排序是以下指标的简单总和：

  – *Citing count:* 分支节点候选论文引用的种子论文数量（相关性）– *Cociting count:* 该论文与源论文共享的参考文献数量（相似性）– *Cocited count:* 该论文与源论文同时被引用的频率，并按时效性加权（相关性 + 重要性）

时效性加权采用对数加权： $w(t) = 1 + \ln\left(1 + \frac{h}{t}\right)$，其中 t 为自出版以来的年数（限制为至少 1）， $h = 4$ 为半衰期。

方向	指标	定义	解读
根节点 （向后）	被引次数	引用本文的种子文献	重要性
	共被引次数	参考文献中与种子文献共同出现	相关性
	共同引用次数	与种子文献共享的参考文献	相似性
	总排名	上述各项之和	总体得分
分支 （向前）	引用次数	被本文引用的种子文献	相关性
	共同引用次数	与源文献共享的参考文献	相似性
	共被引次数	与源论文的共被引**（按时效性加权）**	相关性 + 重要性
	总排名	上述各项之和	总体得分

表 1：根节点和分支论文选择的排序指标。

排序指标摘要见表1。最终选入图中的节点包括源论文、源论文的所有直接参考文献、分支种子（数量设有上限），以及排名靠前的根节点和分支。图1展示了使用 Oignon 创建的图的一个子部分。出版物以圆圈表示，并按发表年份排序。圆圈大小代表全局引用次数。源节点以深灰色实心节点表示，浅黄色表示选中以供检查的节点，绿色圆圈则表示参考文献或施引文献，具体取决于 y 位置。

参考文献

[1] Jason Priem、Heather Piwowar 和 Richard Orr。《Openalex：学术作品、作者、出版场所、机构和概念的完全开放索引》，2022。[2] Suzanne Fricke。《Semantic Scholar》。载于 *Journal of the Medical Library Association*，第 106 卷，第 145–147 页，2018。[3] H. Ballington。《Oignon：引文图工具》。 https://github.com/hballington12/oignon，2025。[4] Tim Wohlin。《本地引文网络》。 https://LocalCitationNetwork.github.io，2024。开源引文网络可视化工具。