RecruitScope：用于多维招聘数据分析的可视分析系统

Xiyuan Zhu
华中科技大学 武汉，湖北，中国
AdInistrator.FallinDepart@gmail.com

Wenhan Lyu
华中科技大学 武汉，湖北，中国
3980181523@qq.com

Chaochao Fu
华中科技大学，中国湖北武汉
2460468415@qq.com

Yilin Wang
华中科技大学，中国湖北武汉
2208469290@qq.com

Jie Zheng
华中科技大学，中国湖北武汉
2639840745@qq.com

Qiyue Tan
华中科技大学，中国湖北武汉
2033935380@qq.com

Qianhe Chen
华中科技大学，中国湖北武汉
qianhechen01@gmail.com

Yixin Yu
华中科技大学，中国湖北武汉
sakaaanayu@gmail.com

Ran Wang*
华中科技大学，中国湖北武汉；
教育部大数据与国家传播战略哲学社会科学实验室，中国湖北武汉；
rex_wang@hust.edu.cn

arXiv:2601.04630v1 [cs.HC] 8 Jan 2026

摘要

在线招聘平台已成为现代招聘的主要渠道，但大多数平台仅提供职位名称、关键词和薪资范围等基础筛选功能。这限制了对不同尺度下多属性关系和就业市场模式的全面分析。我们提出了***RecruitScope***，这是一款可视分析系统，旨在支持求职者和雇主，尤其是人力资源专家，对招聘数据进行多维度、跨层级探索。通过协同可视化，***RecruitScope*** 使用户能够从多个视角分析职位和薪资模式，在宏观层面解读行业动态，并在微观层面识别新兴职位。我们通过案例研究展示了***RecruitScope***的有效性，这些研究揭示了地区薪资分布模式、刻画了行业增长轨迹，并发现了就业市场中需求旺盛的新兴职位。

CCS 概念

• **以人为中心的计算 → 人机交互（HCI）**；可视化系统与工具。

关键词

招聘数据分析、劳动力市场分析、探索性数据分析

1 引言

在线招聘平台已成为现代招聘的主要渠道，雇主和求职者日益依赖这些平台发布和搜索职位 [1]。这些平台的广泛普及产生了大量多维度的结构化招聘数据，例如工作岗位、薪资、地点和技能要求。对这些多维信息进行有效可视化，能够全面洞察就业市场。

然而，现有的招聘数据可视化方法存在明显局限。从平台功能的角度来看，大多数在线就业市场未能充分整合求职者的多方面要求，例如薪资、地点和福利待遇等。因此，它们无法全面呈现职位相关信息[2, 3]。造成上述局限的一个根本原因是，当前系统通常最多只允许用户同时考察两个属性，并将多个因素之间的相互依赖关系简化为二元关系 [4]。因此，传统方法难以揭示招聘记录多个维度间复杂的交互作用和潜在模式。

这些局限阻碍了人们全面理解就业市场。一方面，用户难以同时考察多个维度之间的交互关系，这限制了他们深入了解工作差异和薪资分布模式。另一方面，传统分析方法使得

*通讯作者。

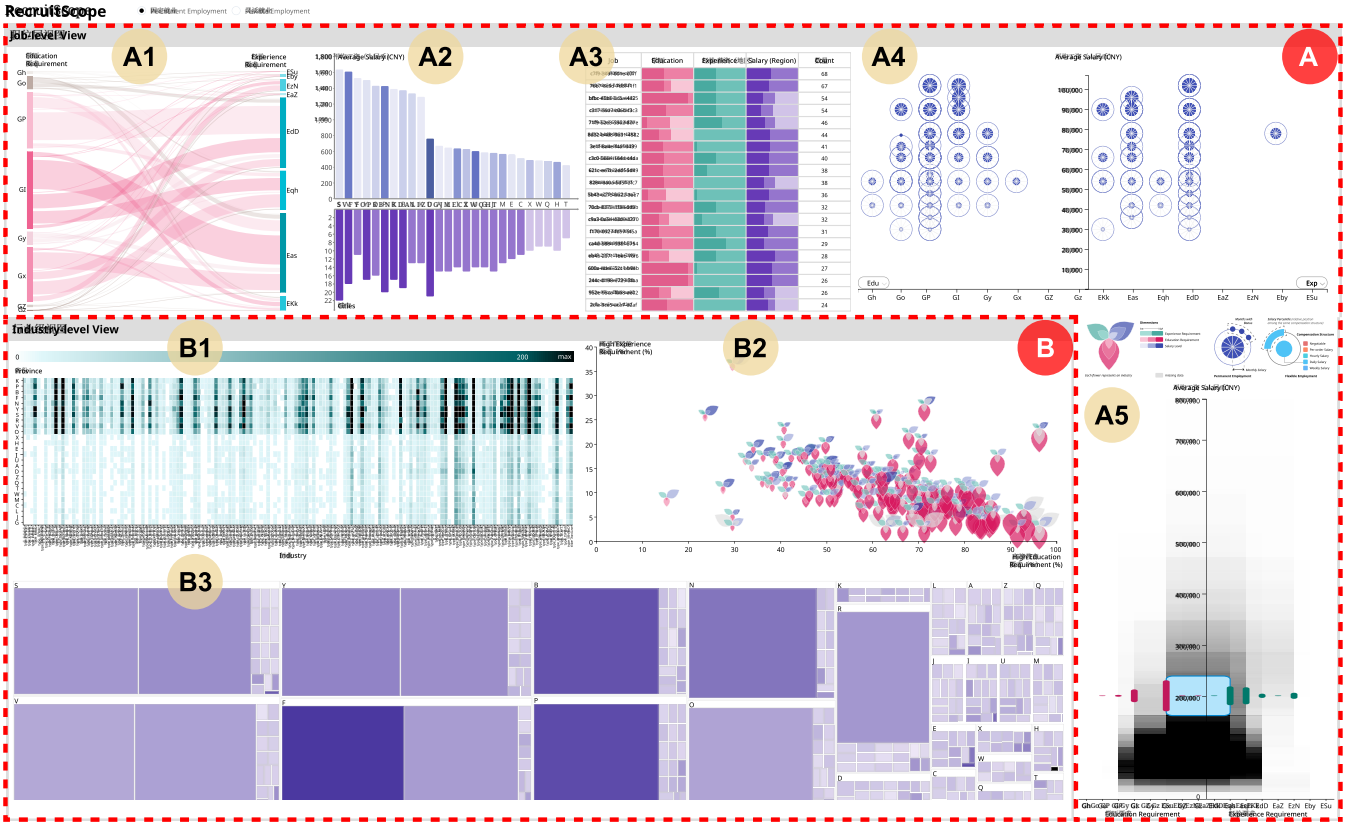


图 1：*RecruitScope*：一个用于招聘数据的多层次可视分析系统。A）职位层视图整合学历、经验、薪资和区域分布，支持对职位进行详细探索。B）行业级视图支持跨行业和跨区域比较，以揭示行业动态和薪资模式。图中所示的职位名称、学历和经验要求以及地区均已匿名化，以保护数据隐私。

推断行业动态和就业市场不断演变的趋势变得困难。更重要的是，在整体行业格局与具体职位详情之间切换视角时，现有分析工具缺乏足够的灵活性，因而妨碍用户形成对招聘市场的整体认识。

为解决上述局限性，我们提出：

- ***RecruitScope*** 是一款面向招聘数据的可视分析系统，支持从多个维度进行探索。它帮助用户分层次地理解劳动力市场动态，从宏观行业趋势到微观职位差异，从而克服分析多个因素之间相互作用时的局限。

- **一种新颖的花形散点图设计**，灵感源自日益兴起的基于图元的可视化趋势，其中每条数据记录都被编码为紧凑的视觉标记，以呈现多个属性。在招聘数据分析中，展示不同行业或地区之间的整体关系，以及各类别内要求的多维构成，至关重要。因此，我们将散点图的空间分布与雷达图的多维编码能力相结合。每个花瓣都编码行业特定要求例如学历、经验和薪资，而散点图上的空间位置则反映各行业在招聘市场中的相对位置。这种设计不仅呈现了不同行业之间的相对关系，也揭示了各行业内部的要求模式，弥补了传统雷达图仅支持比较少数对象的局限。

- **典型案例研究**通过揭示薪资分布模式、识别行业演变趋势并发现新兴职位角色，验证了我们系统的有效性。

2 相关工作

2.1 劳动力市场分析

劳动力市场分析（LMA）利用数据、指标和统计模型，为劳动力供需提供洞察、预测和决策支持。它有助于指导国家就业政策 [5]并协助企业制定有针对性的人才招聘决策 [6]。

劳动力市场分析的理论基础源自传统经济模型，例如戴蒙德-莫滕森-皮萨里德斯模型

基于劳动者异质性 [7]。随着互联网的发展，在线招聘数据已成为劳动力市场分析的重要来源 [8], ，能够提供最新的就业市场趋势指标。然而，传统方法通常依赖具有固定模式的人工标注数据集，可能无法捕捉劳动力市场的动态变化 [9]。后来，动态图、超网络和合成增强等新方法被应用于劳动力市场分析 [10—12]。

尽管如此，现有研究仍存在明显局限。一方面，许多研究仍聚焦于宏观统计数据，例如职位数量或特定技能的出现频率分析，未能揭示多个维度之间复杂的相互依赖关系 [13]。另一方面，劳动力市场分析既需要识别行业层面的宏观趋势，也需要捕捉个体职位层面的微观细节 [6]。然而，当前方法往往只关注单一尺度，难以支持跨层级探索 [13, 14]。

为解决这些局限，我们提出了***RecruitScope***，这是一个面向招聘大数据的可视分析系统，支持用户开展多维度、跨层级探索。

2.2 职位可视化

职位可视化 是基于在线招聘数据构建的交互式可视分析系统 [15, 16]。它将大量招聘信息转化为多维结构化信息，并以可视化方式呈现，帮助求职者和雇主 [17]。

可视化设计领域的现有工作已从静态展示发展到交互式分析。早期原型通常是静态的，或仅提供有限的交互功能，主要侧重于关键词提取和频率分析，以识别基本关系 [18, 19]。近年来，涌现出更多由大数据驱动的多视图可视分析原型，使用户能够跨多个维度关联探索招聘数据。一些研究还引入了推荐和匹配解释，以增强决策支持和可解释性 [20]。与此同时，JobViz 等系统 [3] 体现了职位探索从关键词搜索向技能驱动分析的转变，而 Skills2Graph 等方法 [21] 则展示了利用大规模语料库支持多维关系探索的可能性。

总之，招聘数据包含丰富的信息，例如职位异质性、薪资模式和行业趋势，但当前的分析方法仍难以揭示这些因素之间错综复杂的相互关系。因此，***RecruitScope*** 旨在揭示复杂招聘数据中的潜在关系，从而更好地支持用户在就业市场中的决策。

3 设计目标

根据初步研究结果，RecruitScope 面向招聘生态系统中的两类核心用户群体：求职者和雇主，尤其是人力资源专家。该系统旨在通过多层次可视化分析，帮助求职者高效探索职业机会，同时为雇主和人力资源专家提供有关人才市场结构与趋势的洞察。因此，我们提出以下设计目标：

从求职者的角度出发，我们提出**G1**，即**支持对职位异质性和薪资模式进行多维分析**。

- **G1.1：量化多个维度之间的差异**。系统应整合招聘记录中的多种特征，帮助用户评估不同职位之间的差异。这有助于从宏观层面理解招聘市场的需求结构，并为更明智的决策和规划奠定基础。

- **G1.2：揭示潜在的薪资分布模式**。招聘数据中的薪资往往呈现显著偏度、长尾效应或结构性差异，因此很难识别高薪职位集群或低薪资风险区域。因此，系统应将隐藏的薪资分布模式可视化，帮助用户制定有效的政策和求职策略。

从雇主或人力资源专家的角度出发，我们提出**G2**，旨在**支持从宏观到微观的层次分析，从招聘数据中发掘潜在洞见**。

- **G2.1：对行业动态进行宏观层面的解释**。行业发展通常遵循增长、成熟和衰退的周期性模式 [22]。然而，现有的招聘数据分析往往只关注职位发布数量的短期变化，缺乏对长期轨迹的描绘。这使用户难以判断哪些行业正在兴起，哪些行业正在衰退。因此，系统应从宏观视角揭示扩张行业和收缩行业的演化路径。

- **G2.2：微观层面发现人才需求紧迫的新兴职位**。 新兴职位通常规模较小且分布广泛，因此难以通过传统的汇总统计方法发现。尽管如此，它们对于预测劳动力市场变化具有很强的价值。因此，系统应支持细粒度的职位级分析，以发掘快速增长的职位。

4 数据处理

我们使用了由 2024 ChinaVis 挑战赛提供的招聘数据[1]，其中包含 430,664 条招聘记录，涵盖 169,540 个职位、267,296 家公司和 158 个行业，分布于 26 个省级和 371 个市级行政区。该数据集包含 8 种经验要求和 8 种教育程度要求。所有数据均已匿名化；职位、公司和行业均由唯一标识符替代，行政区则采用由一个字母表示省级代码、由三位数字表示市级代码的方案。个人信息及其他敏感字段均已删除。本文中，我们使用简写符号来指代这些匿名化实体。表1 展示了完整的符号系统及其示例。

[1] 2024 ChinaVis 挑战赛

表 1：最终数据结构

字段	定义	数量	表示法示例 / 枚举
_id	职位的唯一代码	1,713	Pos-([a-z0-9]4-)3[a-z0-9]4
省份	省级行政区划-sion	26	Prov-[A-Z]
city	市级行政区划 sion	370	城市-[A-Z][0-9]3
薪资	–	–	–
薪资_类型	薪酬结构	7	–
薪资_基数关系	永久雇佣的薪资周期	–	–
上限_	薪资上限	–	–
下限_	薪资下限	–	–
公司	公司的唯一代码	136,149	–
行业	行业类别的唯一代码	158	Ind-[A-Za-z0-9]6
学历	学历要求	8	博士 *(Gh)*，硕士 *(Go)*，学士 *(GP)*，大专 *(GI)*，高中 *(Gy)*，技校 *(Gx)*，初中 *(GZ)*，无要求 *(Gz)*
经验	工作经验要求	8	*10+* 年（*ESu*），*8–9* 年（*Eby*），*6–7* 年（*EzN*），*5* 年（*EaZ*），*3–4* 年（*EdD*），*2* 年（*Eqh*），*1* 年（*Eas*），无要求（*EKk*）

数据筛选。我们首先按关联招聘记录的数量对职位进行排名。结果显示，有 148,201 个职位仅关联一条招聘记录，呈现出明显的长尾分布。由于单条或极少量记录可能反映的是临时或偶发的招聘需求，而非稳定的市场需求，我们选取招聘量排名前 1% 的职位作为分析基础。

要求编码。我们将学历要求映射为三个等级：高等教育（学士学位*(GP)* 及以上）、中等教育（高中毕业证书*(Gy)* 和中专文凭*(Gx)*），以及低等教育（初中毕业证书*(GZ)* 和无教育要求*(Gz)*）[23]。同样，我们将经验要求合并为两个等级：高经验（≥5 年）和低经验（<5 年）。

薪资标准化。原始数据中的薪资采用了各种格式（年度、月度、周度等）。我们首先将就业类型分为固定就业和灵活就业，然后提取薪资上下限 [24]。所有薪资均转换为年度口径 [25]。

异常值检测。 鉴于正式就业与灵活就业在结构上的差异，我们分别对这两组数据进行异常值剔除。首先按省级行政区划和职位对招聘记录进行分组，然后在每组内应用四分位距（IQR）方法。

$$L = Q_1 - 1.5 \times \mathrm{IQR} \tag{1}$$

$$U = Q_3 + 1.5 \times \mathrm{IQR} \tag{2}$$

其中， $Q1$ 和 $Q3$ 分别表示第 25 和第 75 百分位数。区间 $[L, U]$ 之外的记录被视为异常值并予以剔除。我们采用 IQR 方法，是因为它不要求数据服从正态分布，且凭借其稳健性，已广泛应用于招聘和薪资分析 [26]。

最终数据集的记录结构如表 1所示。

5 系统概览

基于设计目标，我们开发了 ***RecruitScope***，这是一个由八个协同视图组成的可视分析系统，分为两个分析层级：**职位层视图（A）**和**行业层视图**

5.1 职位层视图

该**职位层视图**帮助用户从多个维度量化不同职位之间的差异**(G1.1)**，并快速把握整体招聘模式**(G1.2)**。它包含五个部分：

教育—经验桑基图 (A1)。学历要求和工作经验要求是招聘信息中最常见的两类要求，通常会结合呈现，以设定候选人的入职门槛。**教育—经验桑基图**展示了不同职位的学历要求与经验要求之间的对应关系。左侧矩形表示学历要求，右侧矩形表示经验要求。两者之间流带的宽度表示符合相应学历和经验条件的招聘记录数量。通过选择相应的矩形标签或流带，可以筛选特定的学历要求与经验要求组合。与热力图或分组柱状图等替代方案相比，桑基图能更有效地揭示这两个类别型维度之间的匹配模式[27]，并提供更好的视觉连贯性，便于理解各职位入职门槛的整体分布。

地区—城市双向条形图（A2）。**地区—城市双向条形图**展示了各省级行政区划的招聘规模和区域结构。横轴表示省级区划。上方条形图编码了各省的招聘记录数量，

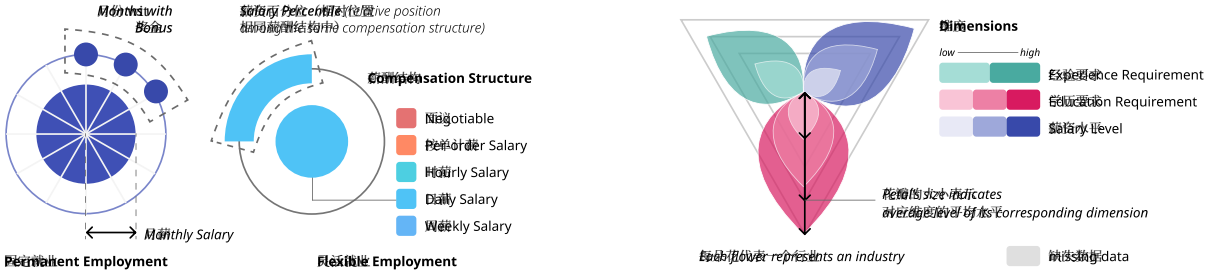


图 2：字形　　设计在（左）薪资模式散点图和（右）行业要求分布视图中　　**w**

条形的不透明度表示相应省份的平均薪资水平。下方条形图表示各省地级行政区的数量。

职位比较列表 (A3)。学历和工作经验要求主要决定职位的入职门槛，而地区招聘分布模式则揭示了市场趋势和薪资差异。与分别考察学历—经验关系 **(A1)**或区域/跨行业薪资分布 **(A2)**不同，**职位比较列表** 将学历、工作经验和区域薪资分布整合到统一视图中。这种扩展的分组柱状图设计在紧凑的布局中纳入多个属性维度，从而能够高效比较组合特征。该视图与**学历—经验桑基图** 和**地区—城市双向柱状图**联动，使用户能够聚焦于筛选后的数据条目。每个矩形的面积表示所选职位中各类别的占比；在“学历”和“工作经验”列中，透明度表示要求等级；在“薪资（地区）”列中，颜色区分高、中、低薪资等级。

薪资模式散点图 (A4)。**薪资模式散点图** 旨在满足支持对两种薪酬结构进行可视化比较的需求 **(G1.2)**：固定就业和灵活就业。在招聘数据中，前者通常以月薪为特征，而后者则包括周薪、日薪和时薪等多种形式。为满足这一需求，该散点图保留了传统散点图揭示模式和异常值的优势，同时引入环形编码，在有限的视觉空间中整合多个维度的信息。用户可以通过下拉菜单，为横轴的正负两侧分别选择一个维度。每个点对应一条招聘数据记录。符号设计如图 2所示。对于固定就业，内圆的大小表示月薪，外环上的小圆数量表示提供的奖金月数。对于灵活就业，颜色用于区分周薪、日薪和时薪，而外环上的弧线则显示该记录的薪资在所有同类型招聘记录中的相对位置。

职位要求分布视图(A5)。**职位要求分布视图**扩展了箱线图，以更好地呈现招聘数据中离散属性与连续属性之间的关系。由于学历要求和经验要求属于类别型数据，而薪资属于连续数据，因此标准箱线图中的须线并不适用。

为解决这一问题，该设计以带状区域取代线条，更直观地展示各职位的学历、经验和薪资分布。每个色块代表一个职位：其横向跨度对应该职位的学历和经验范围，纵向跨度则表示相应的薪资范围。重叠的色块会以视觉方式聚合，重叠越密集，不透明度越高。横轴上的每条带状区域对应一个学历或经验等级，其长度表示要求达到该等级的职位比例。该设计将入职要求与市场回报整合在同一视图中，增强了多维招聘数据的可比性，并便于快速识别薪资分布中的集中区域、差异和异常趋势。

5.2 行业级视图

行业级视图 强调跨行业和地区的宏观层面对比。它支持从多个角度分析行业动态，包括薪资模式、职位需求和要求分布（**G2.1**），并有助于发现人才需求紧迫的新兴职位（**G2.2**）。

行业—地区分布视图（B1）。**行业—地区分布视图**根据薪资水平对省级行政区划和行业进行排名，揭示不同行业的地域偏好以及不同地区的行业偏好。x 轴按照平均工资降序排列行业，y 轴以同样方式排列 26 个省级单位。每个网格单元格的颜色不透明度表示招聘数据记录的数量。

行业要求分布视图 (B2)。**行业要求分布视图** 旨在从学历要求、经验要求和薪资三个维度分析行业。每个花形符号代表一个行业：红色花瓣编码学历，绿色花瓣编码经验，蓝色花瓣编码薪资，灰色花瓣表示缺失值。花瓣大小反映对应属性的平均水平。x 轴和 y 轴分别表示各行业招聘记录中要求高等教育和高经验的比例。我们扩展了传统雷达图的视觉设计。雷达图是可视化多维数据的经典方法，但在大规模行业比较中面临显著挑战。当需要同时展示数百个行业时，重叠多边形会严重降低

可读性。雷达图也侧重于比较少量对象的整体轮廓，因此难以快速识别特定维度上的差异。对行业分组并以网格形式排列多个雷达图，可以在一定程度上缓解重叠问题，但每张图会变得过小，难以辨认细节；此外，网格布局将行业分割成彼此孤立的单元，掩盖了它们在更广阔市场中的相对位置。已有研究提出了其他设计，例如 **Origami Plots** [28] 和 **Flow Radar Glyphs** [29]。Origami Plots 通过固定坐标轴顺序，减轻了对轴顺序的敏感性，但在大规模场景中仍存在重叠问题。Flow Radar Glyphs 突出了流场数据中的时间和方向特征，但密集放置的字形常常导致严重的视觉杂乱。因此，这些方法并不适用于大规模异构招聘数据。为克服这些障碍，我们提出了**花形散点图设计**。该设计将每个行业编码为紧凑的花形符号，并利用散点图中的空间分布来避免重叠，同时在大规模展示时保持视觉清晰。它使用户能够快速发现单个维度中的异常，同时维持行业有序的二维分布。因此，该视图既展现了各行业在招聘市场中的相对位置，也揭示了每个行业要求的内部构成。

区域概况视图（B3）。区域概况视图通过矩形树图布局展示省级或市级划分中平均工资、高需求职位和行业偏好的分布。**省级（B3.1）**区域视图中，每个灰色矩形代表一个省级单位，其中包含代表其下属城市的小矩形。每个矩形的面积反映招聘信息的数量，其颜色不透明度表示相对于其他地区的平均工资。点击主要地区后可进入市级视图，该视图采用相同的矩形树图方案。在每个矩形树图单元格中嵌入一个三层环形图：外环和中环分别按占比展示排名前五的职位和排名前五的行业，而内环的不透明度则

编码平均薪资水平。该层级结构与嵌套的行政区划自然契合，使矩形树图成为可视化多层级区域分布的理想选择。矩形树图能高效利用空间并呈现层级关系，而嵌入其中的环形图则进一步在区域框架内叠加行业和职位信息，增添多维背景信息，同时不影响可读性。[30]通过结合展示区域、职业和行业信息以及薪资水平，该视图支持从宏观市场结构到微观就业偏好的探索，呈现全面的区域劳动力概况。

6 典型案例研究

我们开展了两项典型案例研究，以展示所提出系统的有效性。

6.1 案例 1：发掘劳动力市场中的隐藏机遇

这位求职者*P1*拥有学士学位（*GP*），并具有*3–4*年工作经验（*EdD*），希望找到一个能提供高薪资和良好职业前景的职位。然而，在竞争激烈的招聘市场中，他觉得自己的资历可能使自己处于劣势。

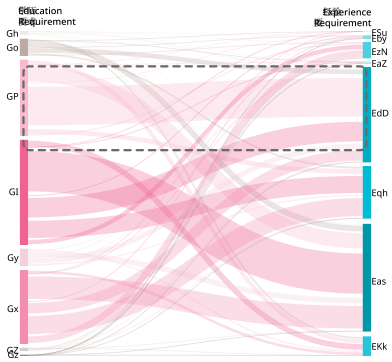


图3：*GP–EdD*组合的流量明显比大多数其他组合更粗，表明这一学历组合具有显著的市场需求。

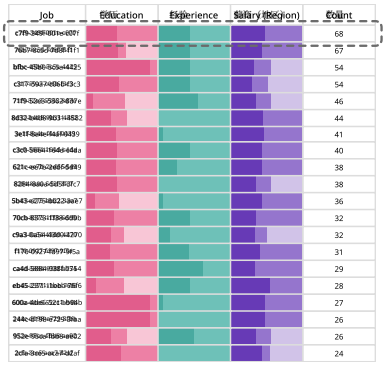


图4：*GP–EdD*对应的职位按招聘数量排序。突出显示的职位（Pos-*c7f9-349f-001e-c07f*）发布频率最高，共有68条招聘信息。

P1 转而查看**教育—经验桑基图**，并选择*GP* 和*EdD*这一组合。相对较粗的流量表明市场需求显著，这稍稍缓解了他的担忧，并促使他进一步探索职位和行业的分布情况。

行业差异与机遇。在**教育—经验桑基图**中选择*GP–EdD* 后，系统会在**职位比较列表**中显示所有相关职位。*P1*选择排名第一的职位 Pos-*c7f9-349f-001e-c07f*进行深入分析。在**薪资模式散点图**中，该职位主要集中于 Ind-*vrMpBQ* 和 Ind-*YbtQwp* 行业，在 Prov-*K* 和 Prov-*F*的薪资水平相对较高，年薪范围为50,000至160,000人民币。大多数招聘记录提供固定月薪，另有一些提供一个月奖金。

为进一步评估职业前景，*P1* 切换至**薪资模式散点图**，并观察到 Ind-*qrSFSq* 行业

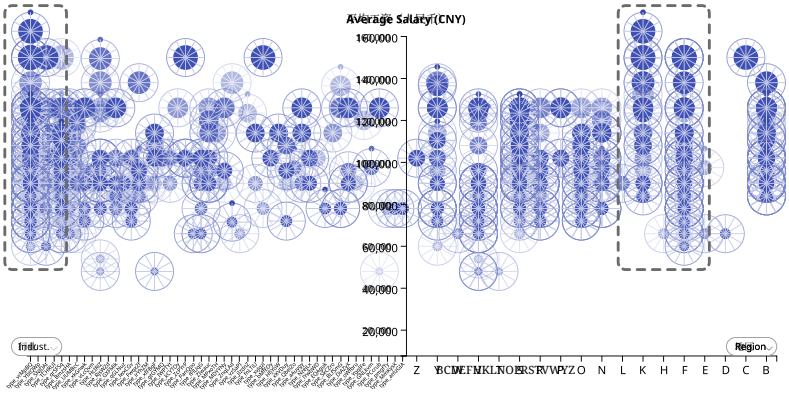


图5：所选职位（Pos-*c7f9-349f-001e-c07f*）集中在 Ind-*vrMpBQ* 和 Ind-*YbtQwp*，且在 Prov-*K* 和 Prov-*F*的薪资较高。

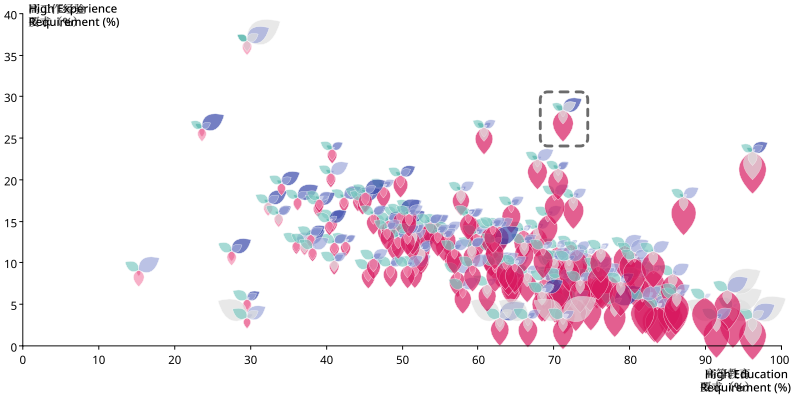


图6：Ind-*qrSFSq* 位于右上区域，花瓣较大且呈粉色，表明其资历要求较高，对长期职位的需求也很旺盛。

密集分布于经济发达地区，表明这是一个正在扩张的行业。在**行业—地区分布视图**中，Ind-*qrSFSq* 的散点位于中右侧区域，且花瓣较大、呈粉红色，表明该行业正进入增长阶段。

地区选择的重要性。*P1* 最初计划以一线城市为目标。然而，**地区—城市双向条形图** 显示，Prov-*K* 和 Prov-*F* 提供的平均薪资显著高于其他地区。结合**行业—地区分布视图**所呈现的信息，*P1* 确认，所选职位在 Prov-*K* 和 Prov-*F*的薪资位居最高之列。这改变了他的职业选择方式，并凸显出地理位置的选择会直接影响薪资。

反思与决策。*P1* 意识到，传统求职策略往往过分强调学历和工作经验，却忽视了其他关键维度，例如行业发展阶段、地区薪资差异和薪酬结构。综合考虑这些因素后，*P1* 决定关注 Prov-*K* 和 Prov-*F*的 Ind-*qrSFSq* 机会，这些地区的薪资相对较高，行业前景也

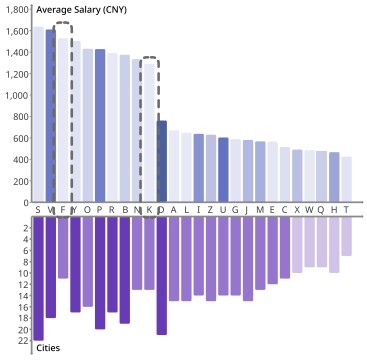


图7：Prov-*K* 和 Prov-*F* 的平均薪资较高。

实力强劲。此外，由于大多数职位提供固定月薪且奖金有限，*P1* 计划在薪资谈判中强调基本工资。

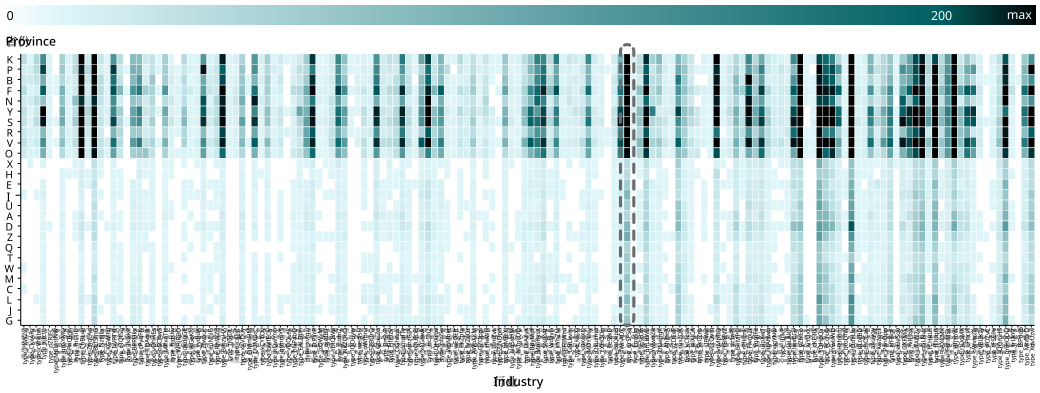


图 8：Ind-*qrSFSq* 聚类中，较深色的单元格代表就业机会充足的经济发达地区。

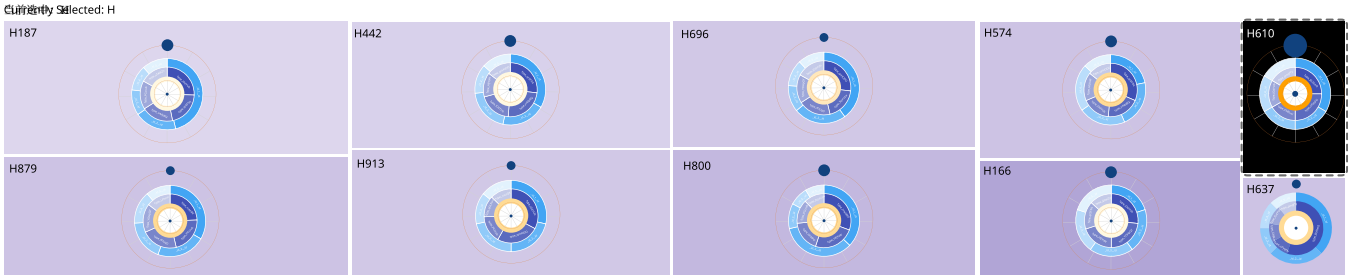


图 9：City-*H610* 脱颖而出，其薪资强度远高于周边城市。

Job	Education	Experience	Salary (Region)	Count
c7f9-349f-001e-c07f				68
76b7-8c5d-7d8f-f1f1				67
bfbc-45b0-5c5a-4425				54
c317-59a7-c06b-f3c3				54
71f9-52e3-5362-8d7e				46
8d32-b4d8-9b31-4582				44
3e1f-8a4e-f4af-0439				41
c3c0-5664-164d-c4da				40
621c-ee7b-2edf-5d49				38
8284-8a0a-5d5f-3fc7				38
5b43-e275-bb22-3ae7				36
70cb-8373-1f38-6d9b				32
c9a3-0a54-43d0-4270				32
f170-0927-fd97-9f5a				31
c4dd-5884-938f-b754				29
eb45-2371-1beb-76f6				28
600a-4de6-52c1-b94b				27
244c-d198-e729-3fba				26
952e-95ca-fbb8-ae02				26
2cfa-3ce5-aca7-d2af				24

图 10：Pos-*eb45-2371-1beb-76f6* 的薪资排名最高，但要求的学历仅为中等，经验要求较低。

6.2 案例 2：检测新兴职位中的人才缺口信号

位于 Prov-*H* 的科技公司*P2* 希望找出竞争相对较低的细分劳动力市场，以制定差异化的招聘策略并获得优势，竞争激烈的人才格局。*P2* 认为，薪资异常高但入职门槛相对较低的职位通常意味着劳动力短缺。

城市薪资异常城市-*H610*。 通过**区域概况视图**考察 Prov-*H* 的整体招聘格局时，*P2* 发现，尽管该省的招聘总量并非最高，但其中一个地级行政区的薪资水平显著高于周边地区。这一异常引起了公司的注意，*P2* 据此推断，薪资水平如此之高，表明特定职位存在劳动力供给短缺。

薪资压缩。 *P2* 随后在**教育—经验桑基图**中选择技校毕业证书（*Gx*）与*1* 年工作经验（*Eas*）这一组合。结合**职位比较列表**和**职位要求分布视图**中的信息，*P2* 发现，Pos-*eb45-2371-1beb-76f6* 的薪资排名最高，但其要求属于中等学历和低经验类别。进一步通过**薪资模式散点图**分析发现，该职位的年薪为 80,000 至 150,000 人民币，显著高于学历和经验要求相近的其他职位。

新兴职位。 *P2* 随后在**行业要求分布视图** 中分析该职位所属的行业，发现该行业的散点出现在

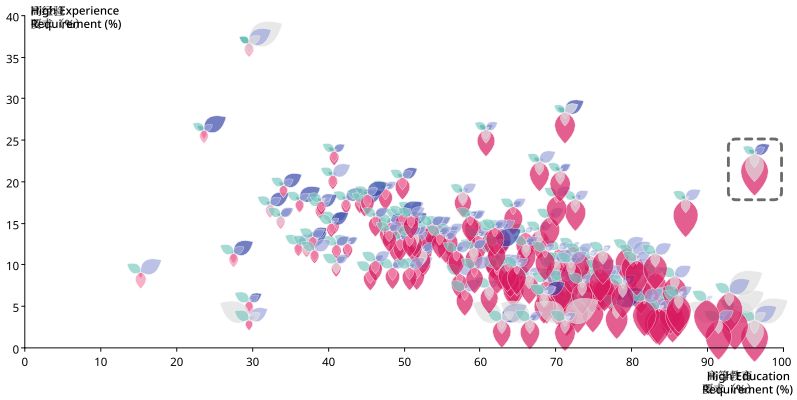


图11：该行业的标记出现在右上象限，并带有较大的粉色花瓣，表明这是一个快速增长且技术密集型的行业。

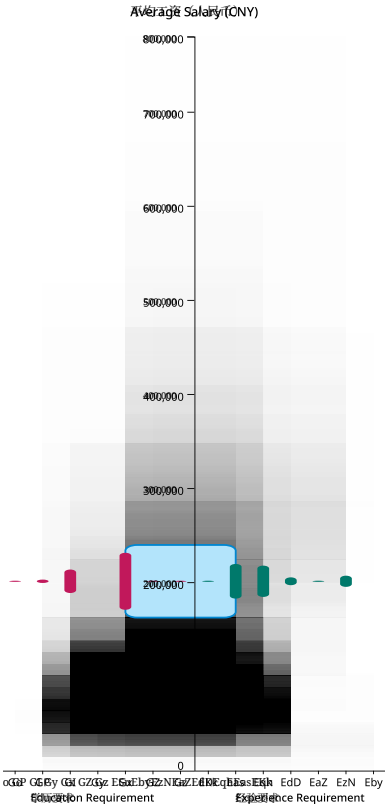


图12：Pos-*eb45-2371-1beb-76f6* 位于中等学历、低经验象限，代表着一个准入门槛适中且薪资较高的机会。

右上区域，且带有一片较大的粉色花瓣，表明这是一个快速增长、技术密集型行业。这说明该职位不仅目前稀缺，而且在行业的长期发展中还可能成为一个新兴职位。

反思与决策。 基于分析结果，该企业制定了以下招聘策略。首先，*P2*决定降低Pos-*eb45-2371-1beb- 76f6*的学历要求。其次，*P2*计划与职业技术学校建立合作关系，提前招揽相关专业的优秀毕业生。最后，*P2* 打算提供具有竞争力的起薪，并配套建立技能培训体系，以吸引和留住人才。

7 结论与未来工作

我们提出了***RecruitScope***，这是一个招聘数据可视分析系统，支持从行业层面到具体职位的分析，有助于揭示招聘市场中的多维模式与趋势。一方面，***RecruitScope*** 将数据分析与可视化呈现整合于同一平台。另一方面，我们通过一种新颖的花形散点图设计，解决了大规模多维数据可读性难以维持的挑战。该设计揭示了多个维度之间的复杂关系，包括职位异质性、薪酬结构和地理特征，从而为理解劳动力市场动态提供了新的视角。

然而，***RecruitScope*** 仍存在一些局限性。首先，尽管该系统可应用于实际场景，但其处理不同结构的招聘数据集以及跨数据源泛化的能力仍有待验证。其次，尽管***RecruitScope*** 支持对多个属性进行协同探索，但缺乏简洁的总结。例如，学历、经验和薪资之间的非线性关系（例如，学历更高并不总是意味着薪资更高）往往需要通过跨视图推理才能发现。

在未来的工作中，我们将把系统扩展到更多样、更贴近现实的数据集，整合来自多个平台的异构数据，并纳入职位描述、必备技能和福利待遇等更丰富的字段。我们还将研究改进后的视觉编码和交互技术，以降低学习曲线，提升系统的整体易用性。

致谢

本研究部分得到中国国家社会科学基金项目（项目编号：25BXW041）资助；部分得到华中科技大学中央高校基本科研业务费专项资金项目（项目编号：82400049）资助。

参考文献

[1] Tawanna R Dillahunt、Aarti Israni、Alex Jiahong Lu、Mingzhi Cai 和 Joey Chiao-Yin Hsiao。2021。考察在线平台在就业中的使用情况：一项针对美国求职者的调查。载于《2021年CHI计算机系统中的人为因素会议论文集》（日本横滨）（CHI '21）。美国计算机协会，纽约州纽约市，美国，文章562，共23页。doi:10.1145/3411764.3445350[2]Yuanyuan Chen 和 Ruijie Pan。2022。基于文本挖掘的招聘职位数据分析与可视化研究。《多媒体进展》2022，1（2022），9047202。arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1155/2022/9047202 doi:10.1155/2022/9047202[3] Ran Wang、Qianhe Chen、Yong Wang、Lewei Xiong 和 Boyang Shen。2024。JobViz：招聘广告的技能驱动型可视化探索。《可视化信息学》 8, 3 2024 18 28 10 1016 2024 07 001 ( ), –。doi: . /j.visinf. . .

[4] Toshiyuki T. Yokoyama、Masashi Okada 和 Tadahiro Taniguchi。2021。Panacea：利用时变图分析年度招聘趋势的可视化探索系统。《PLOS ONE》 16 03 2021 1 22 10 1371 0247587 ( ), –。doi: . /journal.pone.

[5] Anna Giabelli、Lorenzo Malandri、Fabio Mercorio 和 Mario Mezzanzanica。2022。GraphLMI：一种通过图数据库探索劳动力市场信息的数据驱动系统。《多媒体工具与应用》81，3（2022），3061–3090。doi:10.1007/s11042-020-09115-x[6] Chuan Qin、Le Zhang、Yihang Cheng、Rui Zha、Dazhong Shen、Qi Zhang、Xi Chen、Ying Sun、Chen Zhu、Hengshu Zhu 和 Hui Xiong。2025。人才分析人工智能技术综合综述。《IEEE会议录》113，2（2025），125–171。doi: . /JPROC. 3372744.

[7] Arghya Bhattacharya、Paul Jackson 和 Brian C. Jenkins。2018。重新审视中级宏观经济学中的失业问题：教授戴蒙德-莫滕森-皮萨里德斯模型的新方法。《经济学教育杂志》49，1（2018），22–37。doi:10.1080/00220485.2017.1397574[8] Lucia Mytna Kurekova、Miroslav Beblavý 和 Anna Thum-Thysen。2015。利用在线职位空缺和网络调查分析劳动力市场：一项方法论探究。《IZA劳动经济学杂志》4（2015年12月）。doi:10.1186/s40172-015-0034-4[9] Sreerama Murthy Velaga、Sai Srikara Prabhas Satti、Sai Karthik Balivada、Harika Atta 和 Yaswanth Battula。2024。自动化人力资源招聘系统：一种多阶段评估方法。载于《计算创新与新兴趋势国际会议论文集》（ICCIET-2024）。Atlantis Press，1143–1152。doi:10.2991/978-94-6463-471-6 109_

[10] Xi Chen、Chuan Qin、Zhigaoyuan Wang、Yihang Cheng、Chao Wang、Hengshu Zhu 和 Hui Xiong。2024。Pre-DyGAE：用于职业技能需求预测的预训练增强动态图自编码器。载于《第三十三届国际人工智能联合会议论文集》（韩国济州）（IJCAI ' 24 222, 9 10 24963 2024 222 ）。文章页码。doi: . /ijcai. /

[11] Zhuoning Guo、Hao Liu、Le Zhang、Qi Zhang、Hengshu Zhu 和 Hui Xiong。2022。利用动态异构图增强元学习进行人才需求供给联合预测（KDD '22）。美国计算机协会，纽约州纽约市，美国， 2957 2967 10 1145 3534678 3539139 –。doi: . / . Viktor Wolf、

[12] Felix Neubürger 和 Ralf Lanwehr。生成用于技能需求预测的合成数据，以改进预测建模。载于《2023年IEEE世界应用智能与计算会议》（AIC）。313–318。doi:10.1109/AIC57670.2023.10263811[13] Brian Fabo 和 Lucia Mytna Kurekova。2022。在线劳动力市场数据使用相关的方法论问题。技术报告。国际劳工组织研究部，日内瓦。doi:10.54394/ZZBC8484[14] Xi Chen、Chuan Qin、Chuyu Fang、Chao Wang、Chen Zhu、Fuzhen Zhuang、Hengshu Zhu 和 Hui Xiong。2024。Job-SDF：用于岗位技能需求预测与基准测试的多粒度数据集。载于《神经信息处理系统进展》，A. Globerson、L. Mackey、D. Belgrave、A. Fan、U. Paquet、J. Tomczak 和 C. Zhang（编），第37卷。Curran Associates, Inc.，129329–129356。https://proceedings.neurips.cc/paper_files/paper/2024/file/e997325c6f4045aa646c81e674076297-Paper-Datasets_和_基准测试_Track.pdf[15] Francisco Gutiérrez、Sven Charleer、Robin De Croon、Nyi Nyi Htun、Gerd Goetschalckx 和 Katrien Verbert。2019。解释和探索职位推荐：一种用户驱动方法，用于与基于知识的职位推荐系统交互（RecSys '19）。美国计算机协会，纽约州纽约市，美国，60–68。doi:10.1145/3298689.3347001

[16] Sven Charleer、Francisco Gutiérrez 和 Katrien Verbert。2019。通过可操作仪表板为就业中介和求职者提供支持。载于《第二十四届国际智能用户界面会议论文集》（加利福尼亚州 Marina del Ray）（IUI '19）。美国计算机协会，纽约州纽约市，美国，121–131。doi:10.1145/3301275.3302312[17] Dena F. Mujtaba 和 Nihar R. Mahapatra。2020。从招聘信息中挖掘和分析职业特征。载于《2020年国际计算科学与计算智能会议》（CSCI）。683–685。doi:10.1109/CSCI51800.2020.00124[18] Norhaslinda Kamaruddin、Abdul Wahab Abdul Rahman 和 Ramizah Amirah。2019。利用自动关键词选择和可视化方法实现求职者-行业匹配系统。《印度尼西亚电气工程与计算机科学期刊》 13 03 2019 1124 1129 10 11591 13 3 1124 1129 ( ), –。

[19] doi: . /ijeecs.v .i .pp - Guoliang Si、Hengyi Lv、Hangfei Yuan、Dan Xie 和 Ce Peng。2021。一种基于求职者与职位多维结构数据匹配的高效可解释可视化方法。《自然与社会中的离散动力学》2021，1（2021），2215280。arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1155/2021/2215280 doi:10.1155/2021/2215280[20] Ling Xin、Bin Zhou 和 Pan Liu。2024。基于蚁群的职位信息可视化分析与个性化推荐。《EAI〈可扩展信息系统认可汇刊〉》11，3（2024年2月）。doi:10.4108/eetsis. 5061[21] Anna Giabelli、Lorenzo Malandri、Fabio Mercorio、Mario Mezzanzanica 和 Andrea Seveso。2021。Skills2Graph：处理数百万条招聘广告，应对工作技能错配问题。载于《第三十届国际人工智能联合会议论文集》（IJCAI-21），Zhi-Hua Zhou（编）。国际人工智能联合会议组织，4984–4987。doi:10. 24963 2021 708 /ijcai. / Marlene

[22] O'Sullivan。2020。产业生命周期：国家市场在能源技术新兴产业发展中的相关性——以风能为例。《演化经济学杂志》30，4（2020年9月1日），1063–1107。doi:10.1007/s00191-020-00677-5[23] 中华人民共和国全国人民代表大会常务委员会。2021。《中华人民共和国高等教育法》。http: // en.npc.gov.cn.cdurl.cn/2021-04/29/c_954855.htm [24] 中华人民共和国国务院。2008。《中华人民共和国劳动合同法》。https://english.www.gov.cn/archive/laws_ regulations/2014/08/23/content _281474983042501.htm[25] 中华人民共和国人力资源和社会保障部（人力资源和社会保障部）。2025。《人力资源和社会保障部关于职工全年月平均工作时间和工资折算问题的通知》。https://www.gov.cn/zhengce/ zhengceku/202501/content _6995777.htm [26] Qiang Chen 和 Zhijie Xiao。2025。离散度回归、偏度回归和峰度回归及其在美国工资结构中的应用。《应用计量经济学杂志》40，3（2025），325–340。arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1002/jae.3105 doi:10.1002/jae.3105[27] Zana Vosough、Dietrich Kammer、Mandy Keck 和 Rainer Groh。2018。用于可视化比较任务的桑基图镜像。349–355。doi:10.5220/0006651203490355[28] Rui Duan、Jiayi Tong、Alex J. Sutton、David A. Asch、Haitao Chu、Christopher H. Schmid 和 Yong Chen。2023。折纸图：一种改进雷达图的新型多变量数据可视化工具。《临床流行病学杂志》156（2023年4月），85–94。doi:10.1016/j.jclinepi.2023.02.020 电子版发布于2023年2月21日。[29] Marcel Hlawatsch、Philipp Leube、Wolfgang Nowak 和 Daniel Weiskopf。2011。Flow Radar Glyphs——带不确定性的非定常流静态可视化。《IEEE可视化与计算机图形学汇刊》17，12（2011），1949–1958。doi:10.1109/TVCG.2011.203[30] Anderson Gregório Marques Soares、Elvis Thermo Carvalho Miranda、Rodrigo Santos do Amor Divino Lima、Carlos Gustavo Resque dos Santos 和 Bianchi Serique Meiguins。2020。利用分层字形在增强型方形化矩形树图中呈现更多信息。《信息》11，2（2020）。doi:10.3390/info11020123