

# 项目结项报告

---

项目名称：生物医学可视化开源图鉴 Bizard 拓展

项目主导师：王诗翔

申请人：郑虎

日期：2025.09.28

邮箱：tiger.zheng@foxmail.com



开源之夏



openbiox

---

# 目录

---

- [1 项目信息](#)
  - [1.1 项目名称](#)
  - [1.2 方案描述](#)
    - [1.2.1 新增内容：社区征集图表整理](#)
    - [1.2.2 新增内容：Hiplot图表整理](#)
    - [1.2.3 示例数据集整理](#)
    - [1.2.4 Bizard网站优化](#)
  - [1.3 时间规划](#)
- [2. 项目进度](#)
  - [2.1 已完成工作](#)
    - [2.1.1 社区征集图表整理](#)
    - [2.1.2 Hiplot图表整理](#)
    - [2.1.3 示例数据集整理](#)
    - [2.1.4 Bizard网站优化](#)
  - [2.2 遇到的问题及解决方式](#)
    - [2.2.1 对 GitHub Actions 工作流程理解不深](#)
    - [2.2.2 对一些可视化和网站开发工具不太熟悉](#)
    - [2.2.3 其他细节问题](#)
- [3 项目总结](#)
  - [3.1 项目完成情况总结](#)
  - [3.2 未来计划与展望](#)
  - [3.3 致谢](#)

# 1 项目信息

## 1.1 项目名称

生物医学可视化开源图鉴 Bizard 拓展

## 1.2 方案描述

R 语言凭借其强大的统计分析和可视化功能，在生物医学研究中发挥着重要作用。R 语言的可视化工具能够高效呈现复杂数据，例如通用型可视化工具 [ggplot2](#)、[ggpubr](#) 等能够轻松绘制折线图、散点图、柱状图等基础图形；组学可视化工具 [scCustomize](#)、[scRNAtoolVis](#) 等可以实现单细胞组学测序数据的结果展示；临床可视化工具 [survival](#)、[survminer](#) 等可以实现不同治疗组患者生存曲线的绘制。然而，在将 R 代码应用于实际研究场景时，研究者往往面临诸多挑战。代码实现常涉及复杂的环境配置、算法调试和性能优化，而有用的资料却散落在互联网各处——官方文档、技术博客、论坛讨论和开源项目碎片化分布，导致信息检索效率低下。这种碎片化不仅增加了学习成本，还阻碍了研究的可重复性和协作效率。因此，构建一个全面、整合、开源开放的系统知识库至关重要。

OpenbioX 在 2025 年发起了一个名为 Bizard 的开源项目(<https://github.com/openbioX/Bizard>)，这是一个基于社区协作的全面可视化代码存储库。Bizard 旨在简化生物医学研究中的数据分析，整合了各种可视化代码，促进了针对特定研究需求的最佳可视化方法的选择和定制。为学习、利用和讨论数据可视化提供一体化的平台和协作社区，希望构建国内和国际别树一帜的生物信息学基础设施（图1-1）。

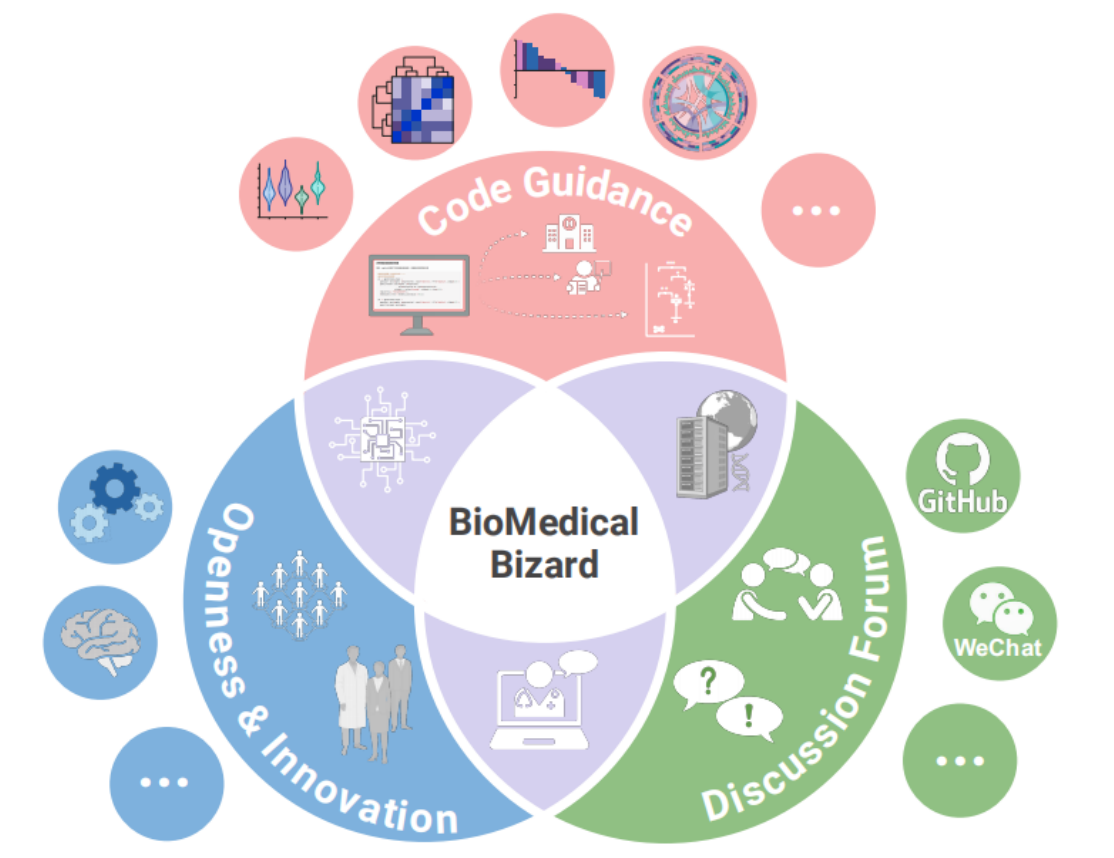


图1-1：Bizard项目预览

Bizard 整合了来自各种来源的可视化代码，包括来自 GraphGallery (<https://r-graph-gallery.com/>) 的传统图表代码以及相关领域的国内外专家的创新可视化方案 (图1-2)。它的目标是为生物医学研究提供全面系统的可视化参考代码与教程库。通过开源开放的模式，促进全球合作，打破技术壁垒，使研究者能聚焦于科学问题本身，而非重复解决技术难题。这样一个基础设施不仅提升研究效率，还将推动生物信息学领域的创新与标准化，为精准医学、基因组学等前沿方向提供坚实支撑。

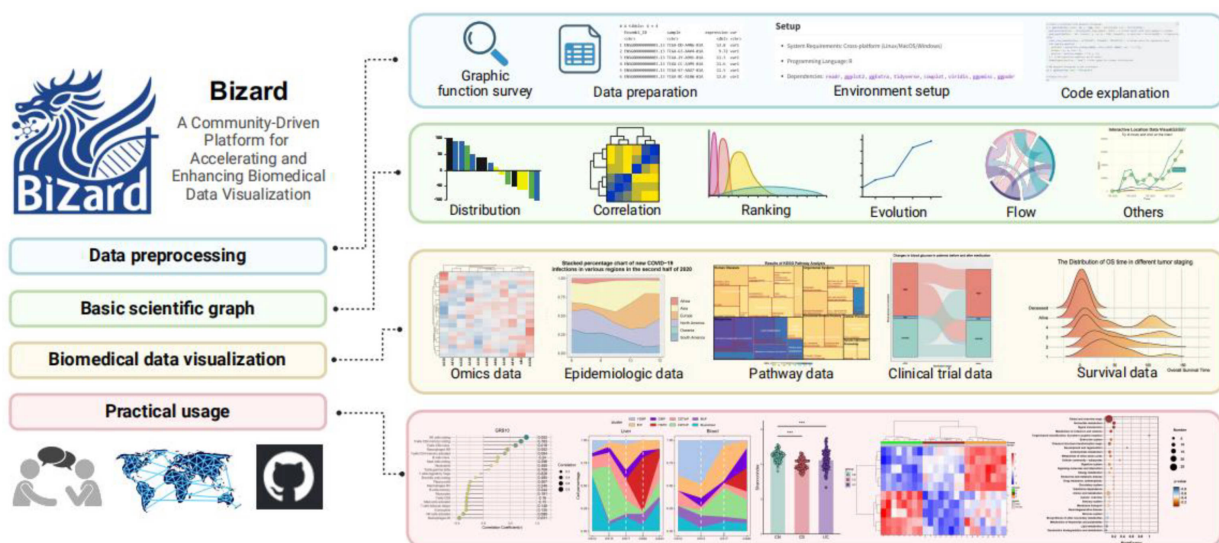


图1-2: Bizard项目工作流程

## 1.2.1 新增内容：社区征集图表整理

为了丰富 Bizard 的图文教程内容，发挥 Bizard 社区化运行模式的优势，项目组向全球各地的社区成员征集生物医学研究相关图表。作为 Bizard 第 1 期任务，经过项目组审核评估，一共有 **31** 个图表考虑立项，涵盖了基础、组学、临床 3 大类型。

具体图表包括：

- (1) **BASICS**: Upset图、韦恩图、断点图、复杂热图、PCA图、载荷图、UMAP图、三元图、部分高亮饼图、径向柱状图
- (2) **OMICS**: 网络图、火山图、组合桑基气泡图、曼哈顿图、染色体结构图、基因结构图、人口图、多组火山图、文字叠加的富集分析柱状图、同线性区块图、KEGG通路图、Motif图、细胞通讯网络图、流图、共线性图
- (3) **CLINICS**: 诺莫图、回归分析表格、元分析森林图、棒棒糖图、Kaplan-Meier生存曲线图、马赛克图

## 1.2.2 新增内容：Hiplot图表整理

[Hiplot](#) 是 OpenbioX 社区开发的生物医学数据可视化交互式应用程序。自 2019 年以来，Hiplot 平台已实现了数百个生物医学可视化插件，涵盖基础统计、多组学分析、回归分析、聚类分析、降维分析、荟萃分析、生存分析、风险模型构建等方向。Bizard 项目组对 Hiplot 平台的可视化插件进行整理，一共收集了 **177** 个图表纳入到 Bizard 教程中。这些图表的纳入不仅丰富了 Bizard 项目的内容，也为 Hiplot 平台提供了可视化插件的源码教程解析，有助于用户理解和使用这些插件。Hiplot 作为 Bizard 第 2 期任务，采取滚动贡献，Github Pull Requests 提交的方式推动。

实现流程如下：

- (1) 用户进入 Bizard 项目主页，将项目 fork 到个人账号；

- (2) 利用 Github 的开发者模式 github.dev (或克隆到本地)，创建图文教程的 qmd 文件；
- (3) 通过 Github 创建对应 Pull Request，根据项目名称修改标题，并将 Pull Request 链接、申领人姓名、日期填写到项目在线表格中；
- (4) Bizard 审核团队对 Pull Request 进行审核，重复修改 qmd 图文教程内容直至满意后即为完成；

### 1.2.3 示例数据集整理

为了方便用户使用绘图代码，Bizard 在每个教程中都提供了示例数据集，并且这些数据集都上传至了腾讯云中通过在线链接的方式提供访问。为了保障后续数据可视化的准确与高效，所有示例数据集都按照统一的规范进行整理。

数据整理的标准包括：

- (1) **格式标准化**：数据整理成易于读取的数据框或者矩阵格式，每一列包含表头；
- (2) **处理缺失值**：尽可能去除含有缺失值的数据；
- (3) **清除冗余**：保证示例数据文件尽可能小，删除无关的数据、重复项及空白行列；
- (4) **命名规范**：数据文件按照绘图进行命名，应当清晰、简洁且具描述性；

### 1.2.4 Bizard网站优化

为了提升 Bizard 网站界面，打造易于使用的数据可视化平台，Bizard 项目组制定了一系列重要的网站优化计划。这些优化旨在全面提升用户体验，使 Bizard 更加易用、强大和开放，更好地服务于每一位数据可视化创作者。

网站优化的内容包括：

- (1) **中英文双语文档**：为服务全球用户，Bizard 将全面实现平台文档的中英文切换功能；
- (2) **Graph Gallery 表格化管理**：针对 Bizard 的上百个图表库，采用表格化管理索引，使用户查找图表更容易；
- (3) **分类目录优化**：引入更精准的多级分类标签和站内搜索引擎，帮助用户更快地定位所需内容；
- (4) **网站自动化部署**：使用 GitHub Actions 实现贡献者提交 Pull Request 后自动化渲染网页并部署到 GitHub Pages；

## 1.3 时间规划

### 第一阶段（2025.7.1 ~ 2025.7.31）

- ☒ 完成对图形可视化的收集和分类工作。
- ☒ 为每个基础分类准备至少 3 个以上可视化图形和应用场景。
- ☒ 完成基础图形分类（BASICS）的可视化教程撰写图表整理，包括：小提琴图、密度图、直方图、箱线图、曲线图等。
- ☒ 完成组学图形分类（OMICS）的可视化教程撰写图表整理，包括：火山图、基因富集分析图、基因结构图、细胞通讯图等。

- ✓ 完成临床图形分类（CLINICS）的可视化教程撰写图表整理，包括：生存曲线图、疾病预测图、医学影像图、风险评分图等。

## 第二阶段（2025.8.1 ~ 2025.9.15）

- ✓ 完成 Hiplot 插件以及补充图形示例数据的整理，整理好的数据集上传至腾讯云以便获取插入教程的链接。
- ✓ 按照 Bizard 项目文档要求，完成 177 个 Hiplot 插件的开源可视化教程撰写图表整理。
- ✓ 完成其他额外补充图形的可视化教程撰写图表整理。

## 第三阶段（2025.9.16 ~ 2025.9.30）

- ✓ 对 Bizard 网站进行优化，增强教程文档的展示效果和易用性。
- ✓ 通过 [babelquarto](#) 包，实现 Bizard 网站的中英文文档切换功能。
- ✓ 通过 [reactable](#) 包，构建 Bizard 网站的 Graph Gallery 页面。方便用户预览所有图形，并支持点击跳转到对应的图文教程。
- ✓ 使用 GitHub Actions 实现网站的自动化渲染和部署

# 2. 项目进度

## 2.1 已完成工作

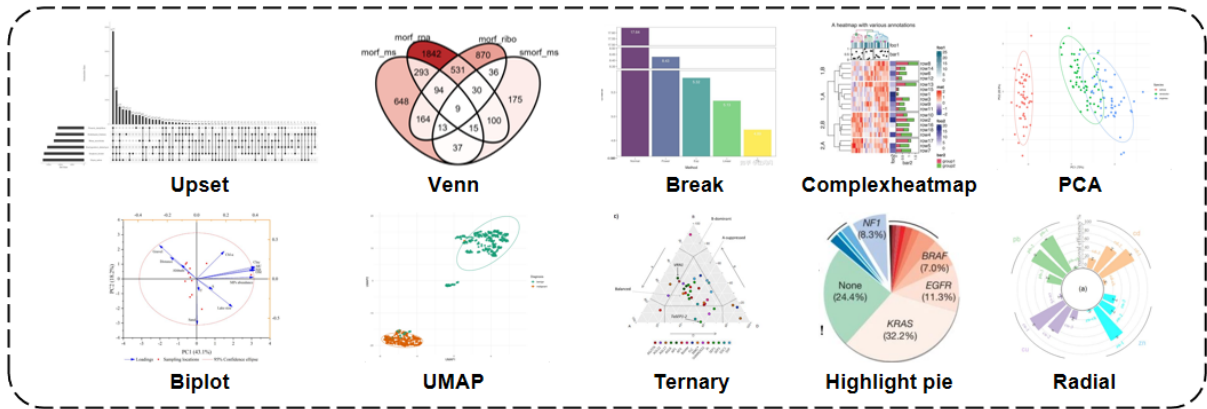
### 2.1.1 社区征集图表整理

相关PR链接: <https://github.com/openbioxBizard/pull/126>, <https://github.com/openbioxBizard/pull/127>, <https://github.com/openbioxBizard/pull/128>, <https://github.com/openbioxBizard/pull/129>, <https://github.com/openbioxBizard/pull/130>, <https://github.com/openbioxBizard/pull/131>, <https://github.com/openbioxBizard/pull/132>, <https://github.com/openbioxBizard/pull/133>, <https://github.com/openbioxBizard/pull/134>, <https://github.com/openbioxBizard/pull/169>, <https://github.com/openbioxBizard/pull/170>, <https://github.com/openbioxBizard/pull/171>

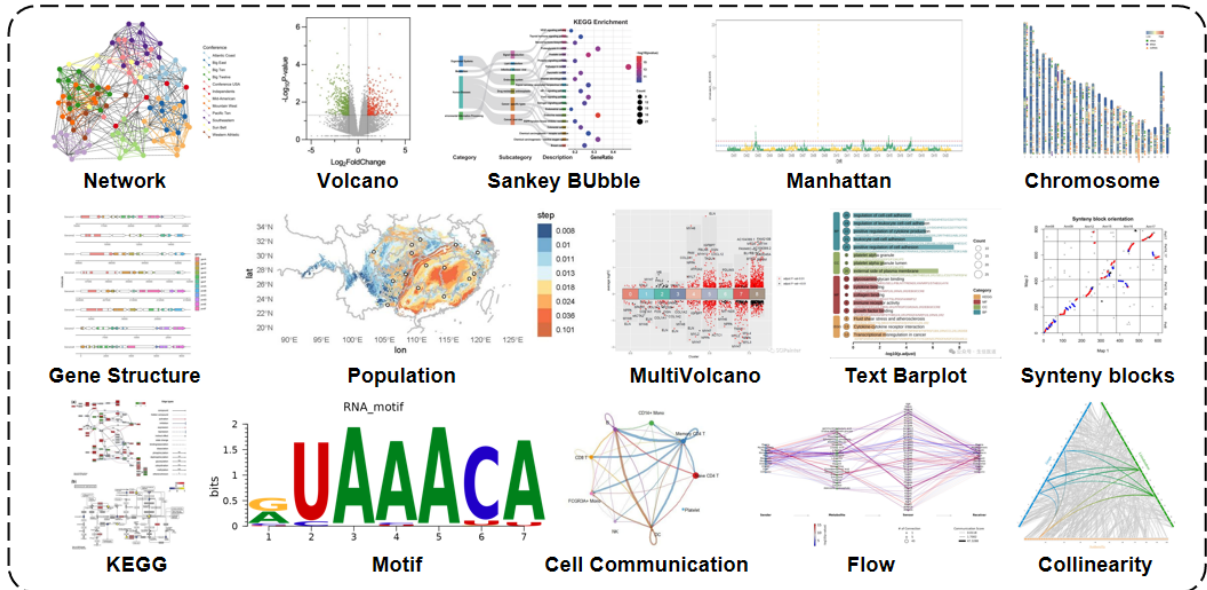
通过社区征集图表，一共纳入了 **31** 个图表，根据图表的内容和用途进行了分类，涵盖了基础、组学、临床 3 大类型（图2-1）。



## BASICS



## OMICS



## CLINICS

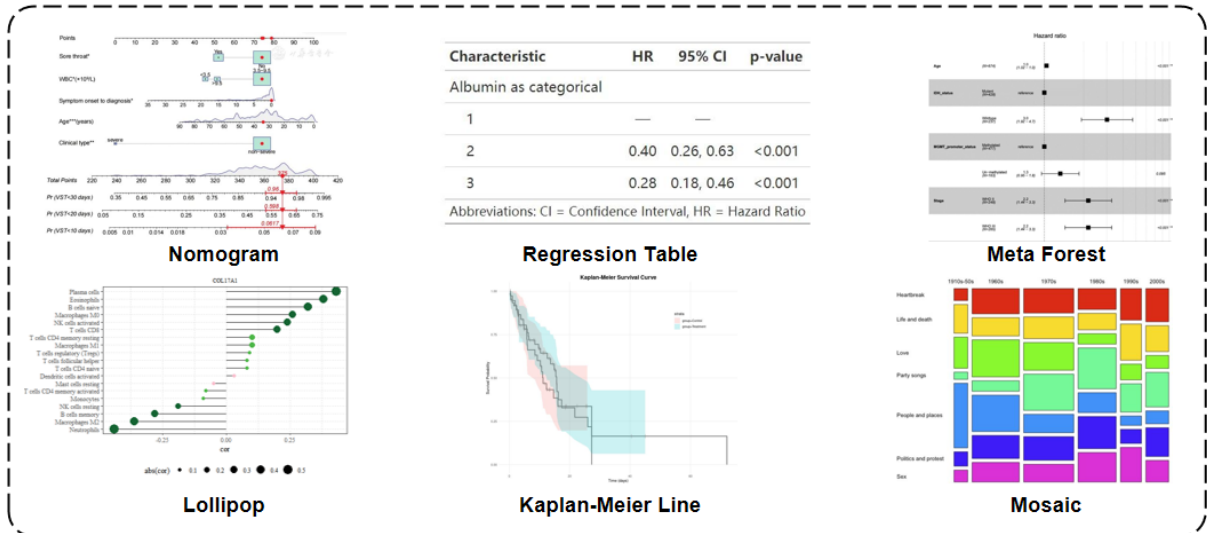


图2-1: Bizard社区征集图表汇总

我们选取其中的基础图形——Upset图，来展示图文教程可视化任务的常规操作流程（图2-2）。用户进入 [Bizard Upset Plot](#) 页面后，可以按照以下五个步骤来使用，具体包括：

- (1) **示例介绍**：展示了图形的基本使用，描述了图形的内容和用途，对图形进行了整体的介绍；
- (2) **环境配置**：安装和加载绘图教程用到的 R 语言程序包；
- (3) **数据准备**：加载 Bizard 准备好的在线示例数据集，并对数据进行整理；

- (4) **可视化**：通过 R 语言绘图代码来绘制图形；
- (5) **应用场景**：介绍了已发表文献中图形的具体应用场景；



图2-2：Bizard Upset 图文教程使用流程

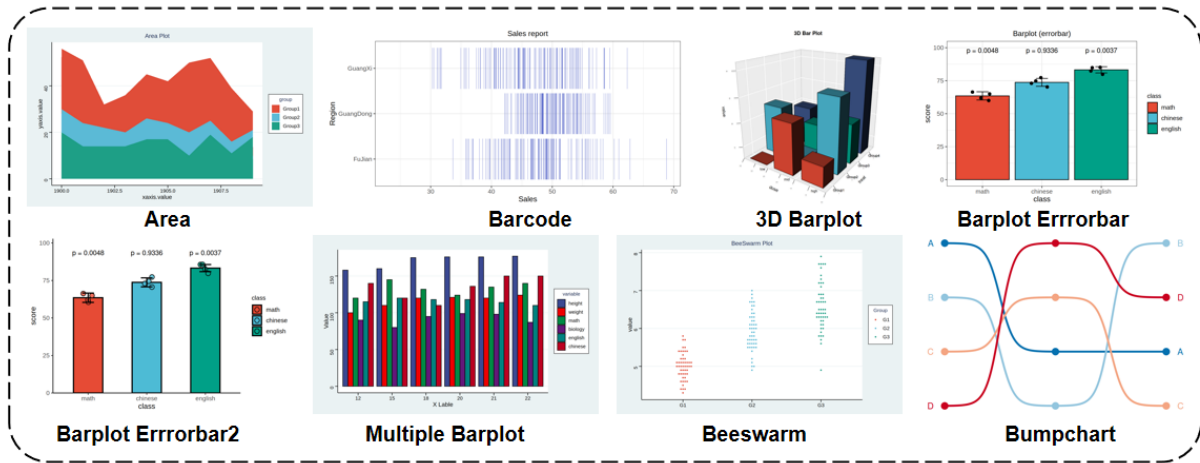
## 2.1.2 Hiplot图表整理

相关PR链接：<https://github.com/openbio/Bizard/pull/135>, <https://github.com/openbio/Bizard/pull/136>, <https://github.com/openbio/Bizard/pull/137>, <https://github.com/openbio/Bizard/pull/138>, <https://github.com/openbio/Bizard/pull/139>, <https://github.com/openbio/Bizard/pull/140>, <https://github.com/openbio/Bizard/pull/141>, <https://github.com/openbio/Bizard/pull/142>, <https://github.com/openbio/Bizard/pull/143>, <https://github.com/openbio/Bizard/pull/144>, <https://github.com/openbio/Bizard/pull/145>, <https://github.com/openbio/Bizard/pull/146>, <https://github.com/openbio/Bizard/pull/147>, <https://github.com/openbio/Bizard/pull/150>, <https://github.com/openbio/Bizard/pull/151>, <https://github.com/openbio/Bizard/pull/153>, <https://github.com/openbio/Bizard/pull/156>, <https://github.com/openbio/Bizard/pull/157>, <https://github.com/openbio/Bizard/pull/158>, <https://github.com/openbio/Bizard/pull/159>

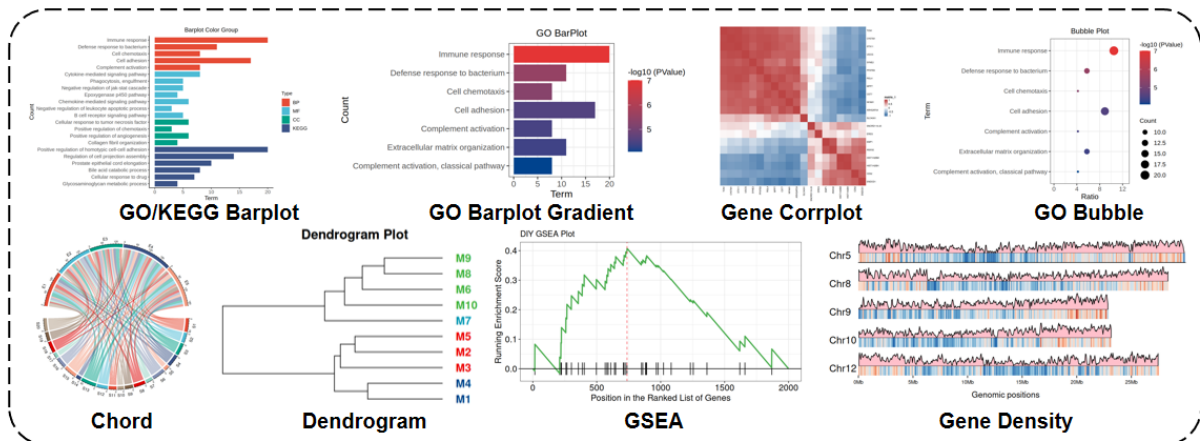
通过对 Hiplot 平台绘图插件进行整理，一共纳入了 177 个图表，根据图表的内容和用途分成了基础、组学、临床、地图 4 大类型。如图2-3所示展示了部分 Hiplot 图表。



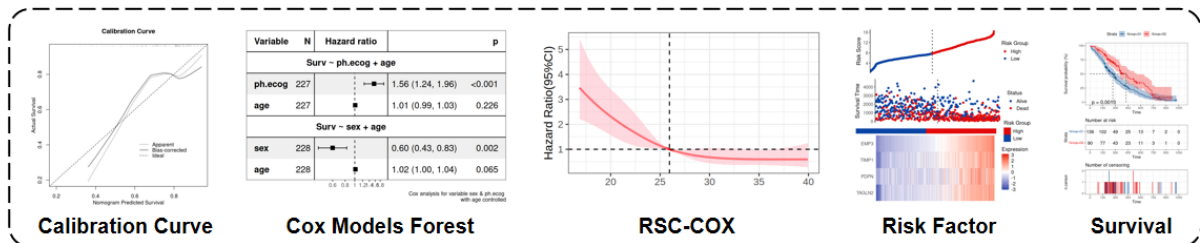
## BASICS



## OMICS



## CLINICS



## MAP

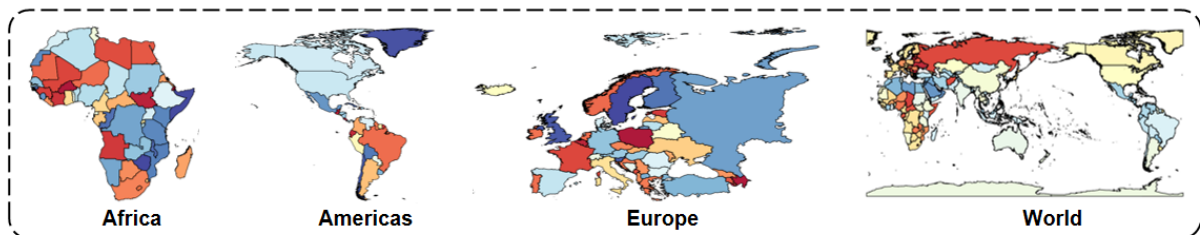


图2-3: Hiplot图表汇总

我们选取其中的基础图形——面积图（Area），来展示 Hiplot 图文教程可视化的常规操作流程（图2-4）。用户进入 [Bizard Area Plot](#) 页面后，可以按照以下五个步骤来使用，具体包括：

- 图形介绍：**提示框显示了 Hiplot 插件的链接，并对图形进行了整体的介绍；
- Hiplot 插件：**点击上一步的链接后会跳转到 Hiplot 平台，使用 Hiplot 插件可以实现无代码绘图，也可以通过下面步骤继续使用源码绘图；
- 环境配置：**安装和加载绘图教程用到的 R 语言程序包；
- 数据准备：**加载 Bizard 准备好的在线示例数据集，并对数据进行整理；

(5) **可视化**：通过 R 语言绘图代码来绘制图形；

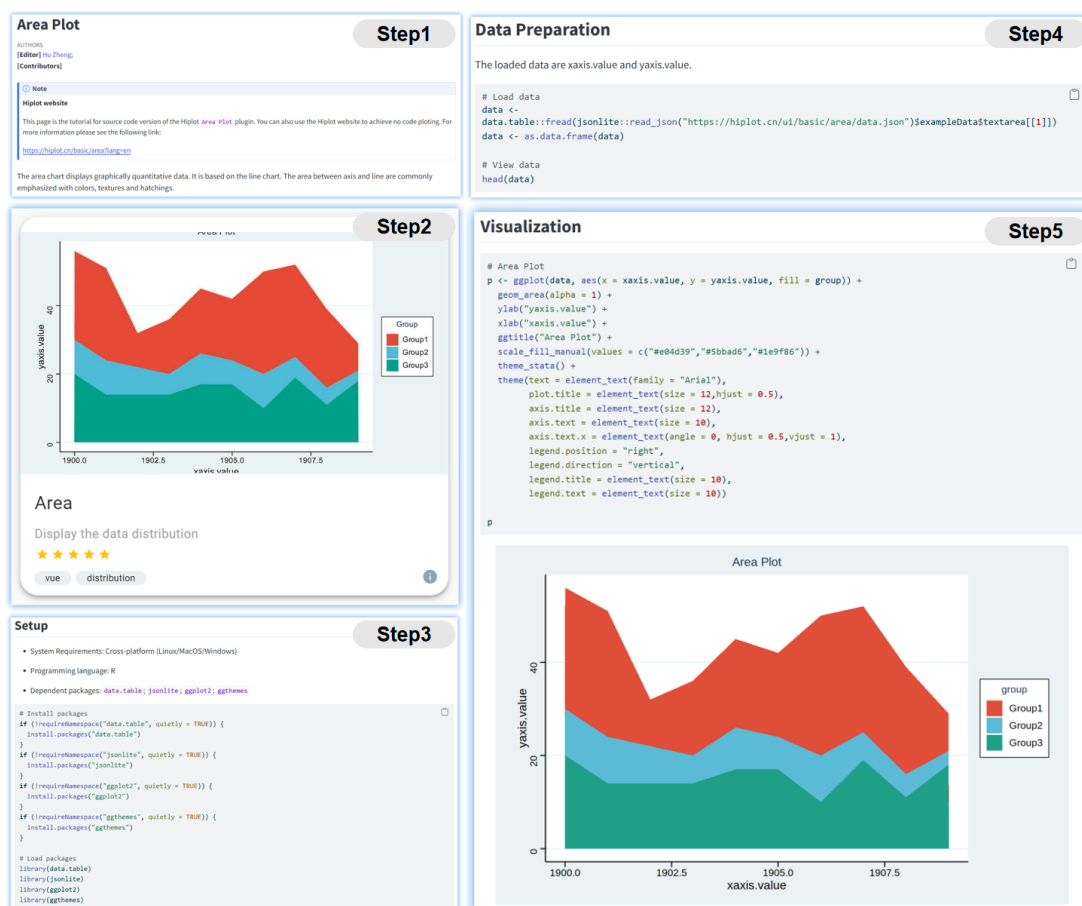


图2-4：Bizard Area 图文教程使用流程

## 2.1.3 示例数据集整理

相关PR链接：<https://github.com/openbio/Bizard/pull/154>, <https://github.com/openbio/Bizard/pull/155>

在数据分析和可视化工作中，高效、规范地整理并托管示例数据集是至关重要的第一步。为了确保数据可被远程安全访问，项目按照特定标准和流程对示例数据集进行整理（图2-5），具体步骤如下：

- (1) **数据准备**：在本地对原始数据集按照标准进行整理，处理缺失值，清除冗余，并按照规定命名；
- (2) **上传至腾讯云**：数据准备完毕后，将其上传至腾讯云对象存储，这样方便用户远程在线读取数据集；
- (3) **使用数据链接**：每个示例数据集都可以通过唯一的互联网访问地址（URL）进行访问，通过 R 语言代码进行读取；

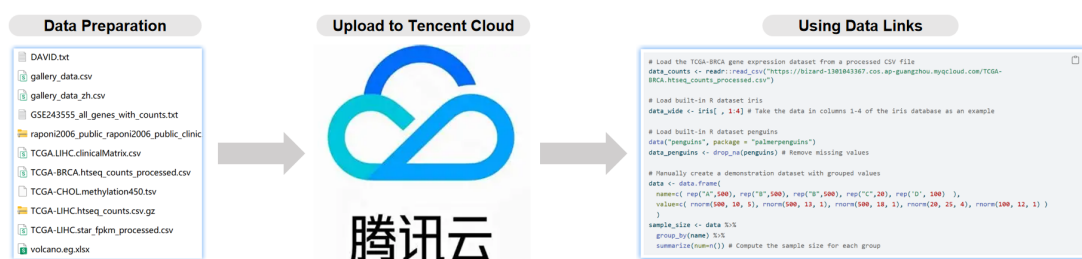


图2-5：示例数据集整理流程

## 2.1.4 Bizard网站优化

### 2.1.4.1 中英文双语文档

相关PR链接: <https://github.com/openbio/Bizard/pull/112>

为了更好地服务全球用户群体, 消除语言壁垒带来的障碍, Bizard 实现了平台所有文档的中英文无缝切换功能。这一举措旨在打造一个真正国际化、包容性强的网站平台。这意味着, 用户只需轻点切换按钮, 即可在英文文档与中文文档之间自由选择, 轻松获取所需信息。为了实现这一目的, Bizard 采用了 [babelquarto](#) 多语言渲染工具, 并按照如下流程进行配置(图2-6)。

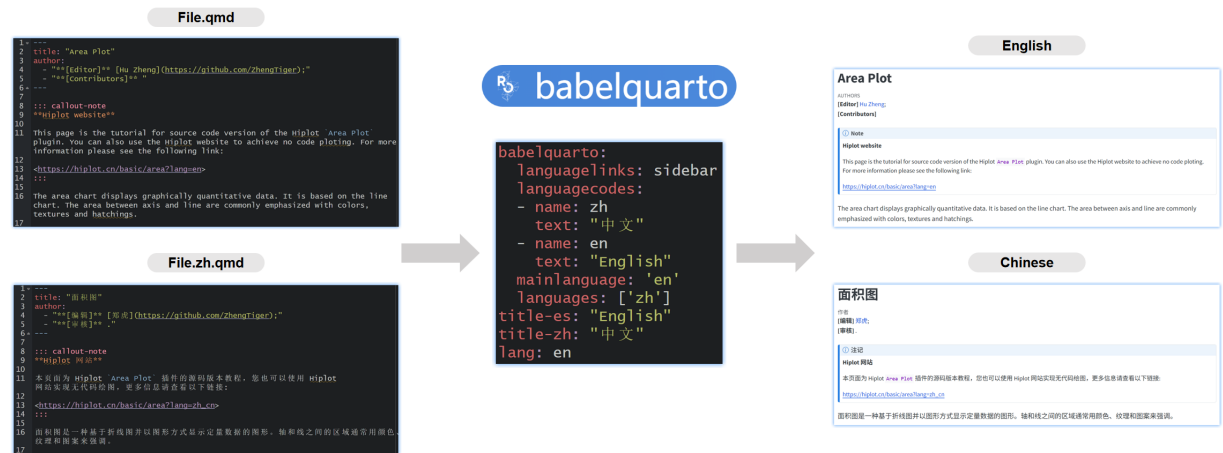


图2-6: 中英文双语文档实现

### 2.1.4.2 Graph Gallery 表格化管理

相关PR链接: <https://github.com/openbio/Bizard/pull/116>, <https://github.com/openbio/Bizard/pull/163>, <https://github.com/openbio/Bizard/pull/164>, <https://github.com/openbio/Bizard/pull/165>

由于 Bizard 平台具有上百张图表及图文教程, 为了彻底解决图表数量激增带来的查找困难、管理混乱与使用效率低下等问题, Bizard 推出了 **Graph Gallery (图表库)** 功能。通过先进的表格化管理工具 [reactable](#) 对平台所有图表进行了整合, 并结合 Javascript、CSS 前端代码渲染表格(图2-7)。在这个表格化界面中, 每张图表都作为一个清晰的条目, 其关键属性如图表ID、图表名称、图表描述、图表类型等一目了然。用户可通过筛选、排序和关键词搜索, 从海量图表中精准定位所需内容, 极大提升了信息检索效率。



图2-7: Graph Gallery 表格化管理实现

### 2.1.4.3 分类目录优化

相关PR链接: <https://github.com/openbio/Bizard/pull/160>, <https://github.com/openbio/Bizard/pull/161>, <https://github.com/openbio/Bizard/pull/162>

为了解决 Bizard 平台上百张图表因数量庞大、维度交叉而导致的分类混乱, Bizard 采用基于**三级分类目录**的图表目录管理体系 (图2-8)。其中**一级目录 (Level1)** 为基础 (BASICS)、组学 (OMICS)、临床 (CLINICS) 和 HIPILOT 几大图形分类; **二级目录 (Level2)** 按照图表分类进行归类, 如: 分布图 (Distribution)、相关性图 (Correlation)、排序图 (Ranking)、组成图 (Composition); **三级目录 (Level3)** 为单个图表类型, 如: 小提琴图 (Violin Plot)、密度图 (Density Plot)、直方图 (Histogram)、箱线图 (Box Plot)。这种树状分类结构的巨大优势在于, 它按照图表的功能和分类进行分组, 将散乱的图表编织成网, 最终构建出一个完整的图表可视化知识库。

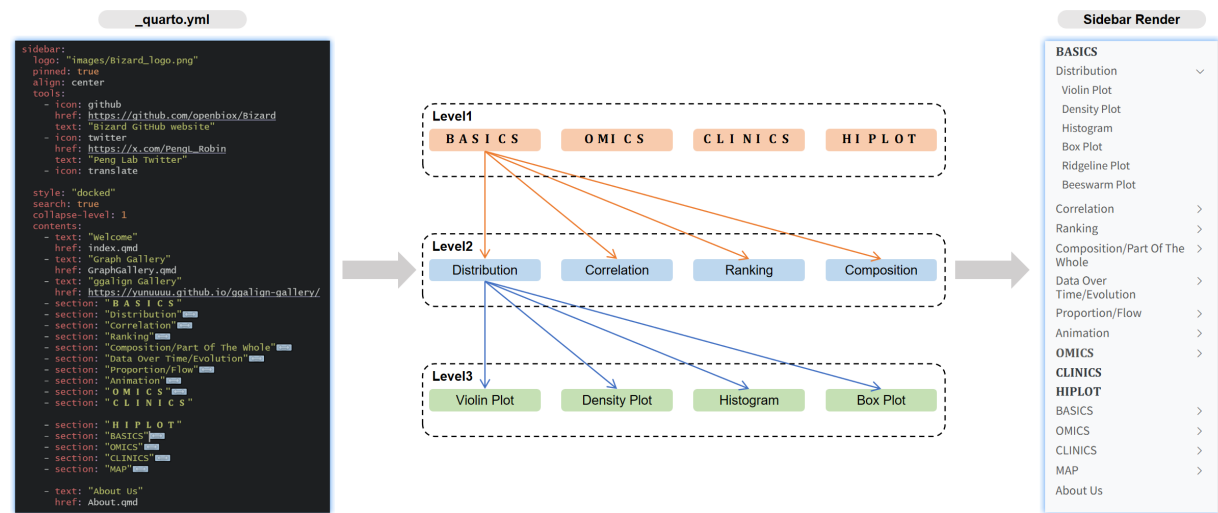


图2-8: 分类目录优化

### 2.1.4.4 网站自动化部署

相关PR链接: <https://github.com/openbio/Bizard/pull/112>

为了保证网站开发的效率与可靠性, 彻底消除手动更新网站所带来的繁琐、延迟及人为错误的风险。Bizard 利用 **GitHub Actions** 构建了一套完备的自动化工作流, 实现了从代码提交到线上部署的全程自动化, 极大地提升了网站迭代的敏捷性与稳定性。其工作流程为: 当开发者将新的网站内容或代码更新推送 (Push) 到指定的主分支 (如 `main` 或 `master`), 或发起一个合并请求 (Pull Request) 时, GitHub Actions 将被自动激活。它会立即在 GitHub 提供的纯净虚拟环境中, 严格按照预设的配置文件 (Workflow File) 执行一系列关键任务, 构建网站并自动化部署到 GitHub Pages 网页 (图2-9)。

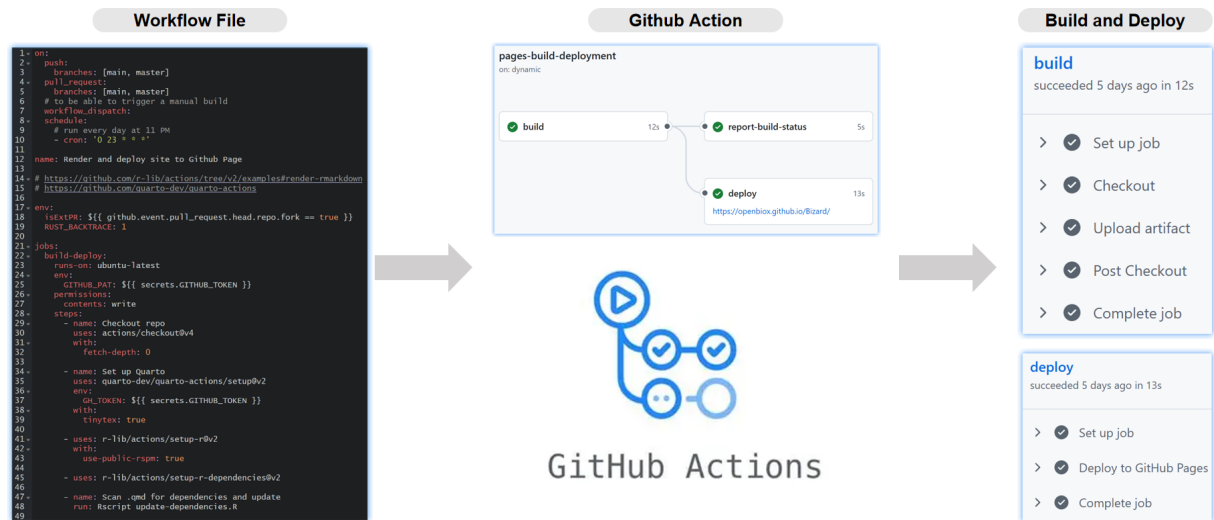


图2-9：网站自动化部署

## 2.2 遇到的问题及解决方式

### 2.2.1 对 GitHub Actions 工作流程理解不深

**遇到问题：**由于之前没有使用过 GitHub Actions 自动化构建工具，对 workflow 文件的字段和代码理解不深，导致构建和部署网站的过程中出现一些报错。

**解决方案：**在开发过程中与导师和项目组其他老师和同学进行讨论，并阅读 [GitHub Actions](#) 文档教程。借助 Deepseek 理解 workflow 文件代码的内容，并实现了添加 [babelquarto](#) 中英文渲染任务到工作流中。

**心得体会：**在进行项目开发时，需要熟悉项目整体的提交流程，理清每一个步骤的作用和触发的 workflow 任务。

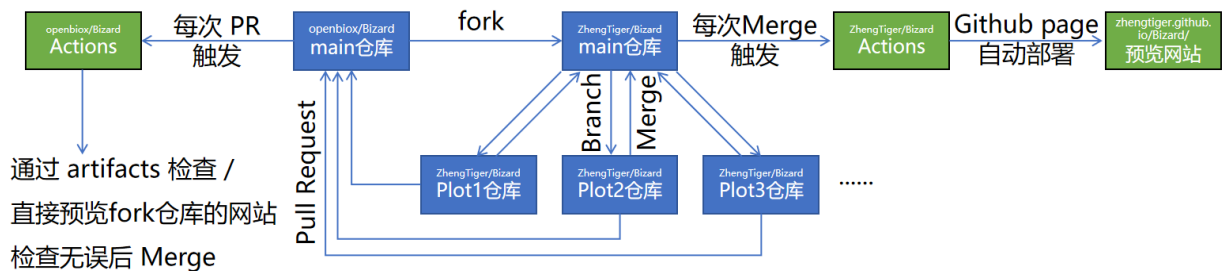


图2-10：Github 提交和自动化构建及部署流程

### 2.2.2 对一些可视化和网站开发工具不太熟悉

**遇到问题：**由于 Bizard 有上百个绘图教程需要编写，其中涉及到了众多不同的绘图软件包。虽然之前已经有一些 R 语言可视化的基础，但是仍然有很多没有接触过的软件包，在开发的过程中经常出现一些报错。

**解决方案：**大部分代码报错都可以借助 Deepseek 进行解决，少部分难以解决的问题通过与导师和项目组成员沟通也得到了解决。另外在开发的过程中我阅读了相关参考书籍和官方文档（[R for Data Science](#), [quarto](#)），系统学习了数据分析和网站开发的流程。

**心得体会：**在项目开发前，进行系统化的学习是十分重要的，它避免了零散知识带来的片面性和漏洞。另外遇到不懂的代码报错不要慌张，Deepseek 等 AI 工具是解决问题的好帮手。





图2-11：参考书籍和文档

### 2.2.3 其他细节问题

**遇到问题：**在开发过程中还遇到了一些其他细节行问题如下表。

**解决方案：**主要通过 Deepseek 辅助，Github issue，以及咨询导师和项目成员等渠道进行解决。

	具体问题	解决方案
1	添加 babelquarto 包到环境报错： Could not solve oackage dependencies	需要在 DESCRIPTION 文件的 Imports: 条目下导入 babelquarto，并在 Remotes: 条目下添加其 Github 来源
2	如何实现对不同的图文教程分开并行开发	借助 git branch 创建分支，在不同分支对各个图文教程分别进行开发，之后将各个分支分别创建 Pull request，审核后合并到主分支
3	Hiplot 一些绘图代码调用了 hiplotlib 中的比较复杂的函数，难以提取出来，如：complex-heatmap	直接将 hiplot/hiplotlib 作为依赖包加入 DESCRIPTION，在绘图时调用 hiplotlib 中的函数
4	一些组学数据文件较大，导致读取慢，而且腾讯云流量消耗也较高	在保证可视化正常运行的情况下，对大文件进行下采样，如：原文件含有 60488 行的基因信息，下采样随机选取 1000 行基因，保证绘图正常运行

**心得体会：**在开发过程中难免会遇到各种细节问题，这是对每位开发者的磨砺。碰到问题时首先要耐心检查，审视错误信息、并使用调试工具逐行分析，这能独立解决大部分问题。另外要善于应用外部资源，借助 Deepseek 等 AI 工具辅助编程，在技术论坛和开源社区查找相似问题的解决方案。必要时及时与导师和项目组成员沟通。

## 3 项目总结

### 3.1 项目完成情况总结

本项目基于开源之夏活动，经过 3 个月（2025.7-2025-9）的开发周期，累计向 [Bizard Github 仓库](#)提交 PR 43 次，均已完成合并。主要成果包括：31 个社区征集图表整理、177 个 Hiplot 图表整理、示例数据集整理、Bizard 网站优化。



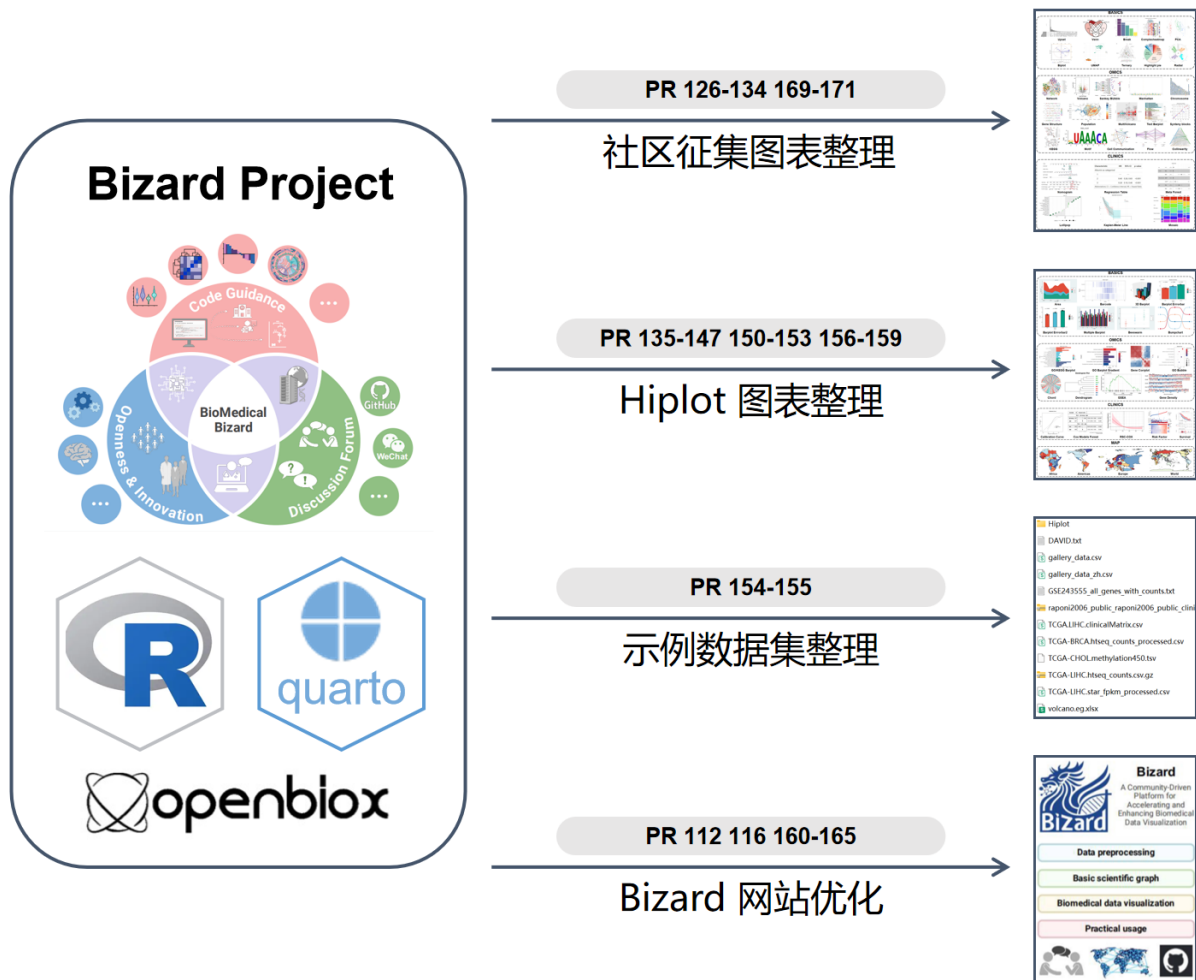


图3-1：项目提交PR汇总

### 不足之处：

由于开发时间的限制，部分项目设想还未得到实现，在后续的开发中会进一步探索，给 Bizard 添加更多对用户有帮助的内容。另外 Bizard 网站的布局也需要进一步优化，以适应未来进一步纳入更多图表。

## 3.2 未来计划与展望

尽管当前版本已实现了 Bizard 的核心功能，但仍然有许多地方值得完善。经过与导师和项目组成员的沟通，Bizard 目前还有如下设想未得到实施：

(1) **引入 AI 大语言模型**：计划引入 AI 绘图助手，通过聊天的方式对图文教程进行讲解，并生成可视化代码。这将极大地降低技术门槛，帮助很多非技术用户快速获得数据分析和可视化的能力。

(2) **搭建基于GitHub Discussion的任务发布与讨论社区**：目前已经初步搭建了环境基础，但后续的社区生态及平台的塑造还需要投入时间运维。创建一个集创意征集、问题答疑、功能需求反馈和协作开发于一体的中心化平台。用户不仅可以在此提交 Bug 报告，更能发起关于新图表类型、新功能特性的讨论，甚至参与 Bizard 项目的制定。

(3) **与 Hiplot 平台进行整合**：Hiplot 在科研绘图领域已经拥有深厚的积累和广泛的用户基础。通过将 Bizard 的新增图表纳入到 Hiplot 无代码分析平台，这将给用户带来更灵活、更强大的交互式分析能力。

后续我们会发布一版 Bizard 的正式版本，并撰写相关科研论文，发表在 SCI 期刊上。

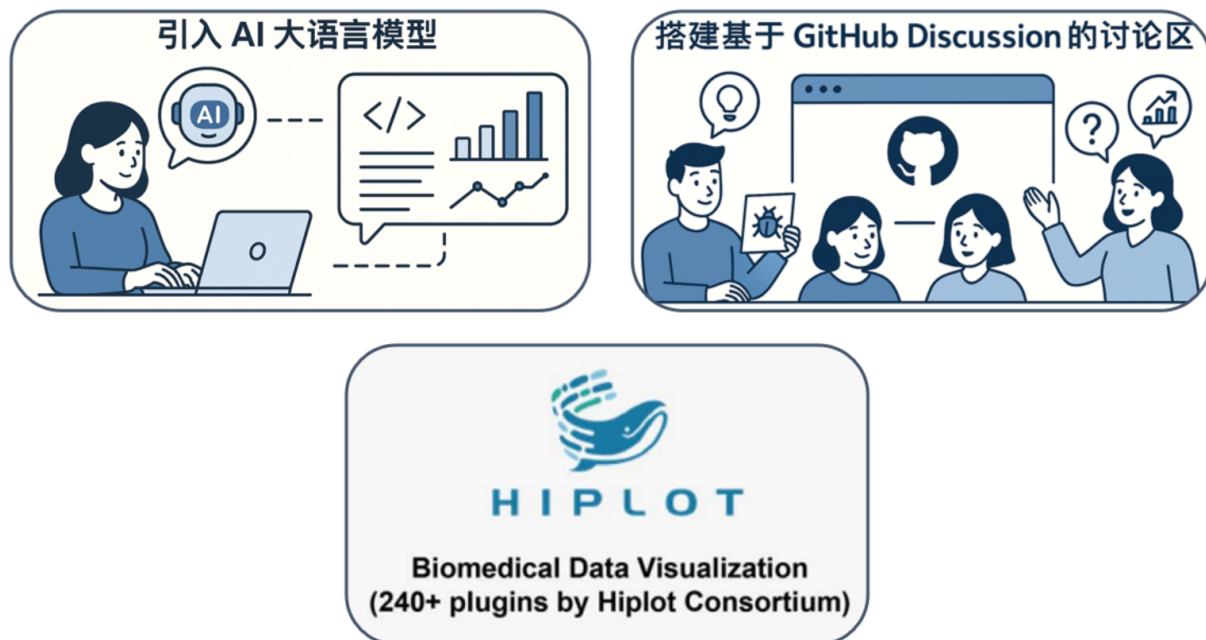


图3-2：未来计划和安排

### 3.3 致谢

感谢开源之夏组委会提供了这一宝贵的平台，让我得以在实践中学习、在协作中成长，让我感受到了开源世界的魅力。

感谢 OpenbioX 社区提供的开源项目支持，为社区成员营造的开放、包容与互助的技术氛围。社区前辈们的无私分享和持续贡献，是我最坚实的后盾和灵感的源泉。

感谢本次项目的指导老师--王诗翔老师，在整个过程中给予的悉心指导。从项目方向的把握到技术细节的斟酌，以及富有洞见的建议和持续的鼓励，是本项目得以顺利推进的关键。同时感谢罗鹏老师、李剑峰老师和项目组的所有同学们在完成项目的过程中给予的帮助。每一次热烈的讨论、每一次无私的帮助，都让我深刻体会到团队协作的魅力与力量。

最后，再次向所有支持、关心本项目的老师、朋友和开源贡献者们致谢。这段经历是我人生中一笔宝贵的财富，我将带着这份收获与感恩，继续在开源技术的道路上探索前行。