

ZX (Zotero - arXiv) 学术日报 产品文档

1. 产品概述

在学术研究领域，科研人员、研究生和学术爱好者需要不断追踪所在领域的最新研究动态，以保持对前沿知识的了解。然而，arXiv 等预印本平台每日发布的论文数量众多，手动筛选相关论文既耗时又容易遗漏重要信息。**ZX (Zotero - arXiv) 学术日报**（下称：本产品）正是为解决这一问题而设计的自动化论文推荐工具。

本产品将与用户的 Zotero 图书馆深度集成，基于用户在 Zotero 中收藏的文献内容，运用先进的算法和自然语言处理技术，每日精准筛选出 arXiv 上与用户研究兴趣高度相关的预印本论文。并且，它将以邮件的形式自动推送给用户，让用户无需手动检索，即可及时获取最新的研究成果。

2. 价值主张

2.1 目标用户

1. **科研人员**：高校、科研机构的专业研究人员，需及时掌握所在领域前沿研究动态。
2. **研究生**：从事学术研究的学生，需阅读大量文献以确定研究方向和获取灵感。
3. **学术爱好者**：对特定学术领域感兴趣，希望及时了解相关领域最新研究成果的人群。

2.2 解决痛点

1. **信息过载问题**：arXiv 作为一个预印本论文平台，每日发布的论文数量众多，涵盖了各个学科领域。对于用户来说，手动从海量的论文中筛选出与自己研究方向相关的论文，不仅耗时耗力，而且容易遗漏重要的研究成果。本产品能够根据用户 Zotero 图书馆中的文献内容，自动筛选出相关度高的论文，大大节省了用户的时间和精力。
2. **跟踪研究动态难题**：科研工作者通常工作繁忙，很难每天都抽出时间关注 arXiv 的更新情况。这就导致他们可能会错过一些重要的研究成果，影响研究工作的进展。本产品能够每日自动推送推荐的论文，确保用户不会错过任何有价值的研究动态。
3. **快速了解论文内容需求**：在面对大量的论文时，用户需要快速了解每篇论文的核心内容，以决定是否值得深入阅读。本产品为每篇推荐论文生成简洁明了的摘要，帮助用户快速掌握论文的核心要点，提高阅读效率。

2.3 相比现有解决方案的优势

1. **个性化推荐**：本产品依据用户 Zotero 图书馆的文献内容**近期阅读记录**进行个性化推荐，能够更准确地反映用户的研究兴趣和需求。相比一些基于通用算法或热门话题的推荐工具，本产品的推荐结果更加贴合用户的实际情况，提高了推荐的准确性和实用性。
2. **文章分类功能**：本产品能够根据论文的主题和内容，自动分类推荐论文，让用户可以更方便地找到与自己研究方向相关的论文。这种分类功能不仅提高了推荐的效率，还让用户能够更直观地了解自己的研究动态。
3. **AI 辅助**：为每篇推荐论文生成摘要，帮助用户快速了解论文的核心内容。这种 AI 辅助的方式，不仅提高了用户的阅读效率，还能让用户在短时间内对多篇论文进行初步评估，筛选出值得深入阅读的论文。同时，通过集成现有的翻译API，本产品还能为用户翻译论文摘要，进一步扩大了用户获取学术信息的能力，简化了学术交流的障碍。
4. **多元化推荐来源**：产品主要从 arXiv 平台获取论文，可以进一步拓展推荐来源，接入更多的学术资源，如 Google Scholar、dblp等数据库，以保证用户获得最新文献信息。
5. **灵活配置选项**：用户可以根据自己的需求，灵活配置各项功能。例如，用户可以设置最大推荐论文数量，以控制邮件中展示的论文数量；可以选择是否发送空邮件，即使当天没有新论文，也能根据自己的喜好决定是否收到通知。

3. 核心功能

3.1 论文推荐

1. **数据收集**：本产品通过与 Zotero 平台的对接，获取用户 Zotero 图书馆中收藏的所有论文信息，涵盖论文标题、摘要、作者、发表时间等内容。同时，会从 arXiv 平台获取前一日发布的所有预印本论文。
2. **兴趣匹配**：本产品会分析 arXiv 上的论文与用户 Zotero 图书馆中论文的相似程度，并且会着重考虑用户近期收藏的论文，以更精准地契合用户当下的研究兴趣。
3. **排序推荐**：依据相似度计算结果，为每篇 arXiv 论文赋予一个综合得分。得分越高，表明该论文与用户研究兴趣的相关性越强。最后，本产品会按照得分高低对论文进行排序，将最符合用户兴趣的论文推荐给用户。

3.2 摘要生成

1. **内容准备**：对于每一篇推荐的论文，本产品会整合其标题、摘要以及引言和结论（若有）等内容，作为生成简短摘要的基础素材。
2. **摘要生成**：本产品提供两种摘要生成方式供用户选择。一种是使用云端的大语言模型服务，这种方式生成的摘要质量较高，但可能会产生一定的费用；另一种是使用本地轻量级模型，无需联网，成本较低，但可能会受到本地设备硬件性能的限制。用户可以根据自身需求选择合适的方式，并指定摘要生成的语言。

3.3 邮件推送

1. **内容呈现**：本产品会将推荐的论文以美观、易读的 HTML 格式整理成邮件内容。每篇论文在邮件中会包含标题、作者、相关性评分（以星星表示）、arXiv ID、简短摘要、PDF 链接，若有代码实现，还会提供相应的代码链接。若当天没有新论文，邮件中会给出相应提示。
2. **发送设置**：用户可以自行设置邮件的发送者和接收者邮箱，确保邮件能够准确无误地发送到指定邮箱。此外，用户还可以决定，即使当天没有新论文，是否接收通知邮件。

3.4 个性化配置

1. **文献过滤**：用户可以通过设置类似推荐的规则，指定不想在推荐中出现的 Zotero 文献集合，从而提高推荐结果的精准度。
2. **数量调整**：用户可以设定邮件中展示的论文最大数量，考虑到生成摘要的时间成本，用户可根据自身阅读习惯和时间安排合理调整该数量。
3. **模型选择**：用户可以根据性能和成本需求，选择使用云端的大语言模型 API 或者本地轻量级模型来生成论文摘要。
4. **语言设定**：用户可以设置摘要生成的语言，确保生成的摘要以自己期望的语言呈现。
5. **要素勾选**：用户可以勾选除摘要外想要详细了解的论文内容，如实验设计、实验结果、数据集描述等。
6. **深度设置**：用户可以根据需求设置文献总结的深度，提供"简洁版/学术版/评审级"三级总结深度，简洁版仅列核心结论，评审级包含更详细的总结。
7. **话题追问**：大语言模型自动生成3个可选延伸问题（如"该模型在跨领域应用的可能性？"），支持用户自定义问题关键词。
8. **多维对比**：用户可以选择最近阅读的论文与推荐论文对比，自动创建对比表格，支持维度选择：创新性、复杂度、实验规模。