

人工智能阅读工具在引文真实性审查中的应用与分析*

——以 ChatPDF 为例

文 采

《扬州大学学报》编辑部, 225009, 江苏扬州

摘 要 为提升科技论文的引文质量和引文编校的工作效率, 提出采用人工智能阅读工具 ChatPDF 检测引文内容真实性的方法。针对不同引用对象和不同语种参考文献的引文, 测试了 ChatPDF 在引文真实性检测中的实用效果, 探讨了该工具在引文编校工作中的应用优势和局限性, 指出合理使用 ChatPDF 在提高科技期刊引文内容审查工作的高效性和可靠性方面的应用价值, 并强调编辑在运用人工智能技术时应持审慎反思的工作态度。

关键词 引文审查; 人工智能; 大型语言模型; ChatPDF

Application and analysis of artificial intelligence reading tools in authenticity review of citations: a case study of ChatPDF
WEN Cai

Abstract In order to enhance the citation quality and the efficiency of citation editing in scientific papers, we propose a method using the artificial intelligence reading tool ChatPDF to detect the authenticity of citation content. The practical effectiveness of ChatPDF in the detection of citation authenticity is tested with citations referencing different objects and referencing in different languages. The study explores the application advantages and limitations of the tool in citation editing work, pointing out the value of judiciously using ChatPDF to improve the efficiency and reliability of content review in scientific journal citations. It also emphasizes the necessity for editors to maintain a thoughtful and reflective attitude when applying artificial intelligence technology in their work.

Keywords citation review; artificial intelligence (AI); large language models (LLMs); ChatPDF

Author's address Editorial Department of Journal of Yangzhou University, 225009, Yangzhou, Jiangsu, China

DOI: 10.16811/j.cnki.1001-4314.2024.02.017

引文, 是为了显示科学研究的继承性而在论文中引用的相关参考文献, 是论文的重要组成部分。一般来说, 规范性与合理性是引用参考文献的基本要求。常见的参考文献错误包括不规范引用和不合理引用, 前者如著者项、标识符、页码、出版项等形式错误, 后者如无关引用、错误引用、漏引、多引等^[1]。不规范引用可以通过人工判别、参考文献校对软件等方式有效识别, 属于相对显性的错误形式; 而不合理引用则需要认

真比对引文与对应参考文献的关系后才能做出判断^[2], 且其审查可靠性受到编辑个人的专业知识水平、外语水平以及工作精力等影响, 属于相对隐性的错误形式。然而, 随着引文分析法在期刊评价、作者评价、文献评价和科技竞争力评价等诸多领域的广泛应用, 引用越来越受到作者、编辑、机构等人为因素的影响, 不合理引用的现象日益增多^[3], 其中, 引文与参考文献内容不一致的引用, 对读者、期刊和学术领域的不良影响最为严重。如何从数量巨大、种类繁多、形式复杂的参考文献中判断引文的真实性, 加强对引文的审查和监督, 捍卫期刊的学术质量和学术诚信, 是目前科技期刊编辑面临的重要挑战。

随着大型语言模型 (large language models, LLMs) 技术的发展, BERT (bidirectional encoder representations from transformers)、GPT (generative pre-trained transformer) 等人工智能 (artificial intelligence, AI) 模型已具备了强大的文本理解和生成能力^[4-5]。基于 LLMs 技术的 AI 文档阅读工具层出不穷, 如 ChatDOC (www.chatdoc.com), ChatPDF (www.chatpdf.com), Humata (www.humata.ai), Monica 等。其中, ChatPDF 是一款基于 ChatGPT 3.5 面向 PDF 文件的人工智能阅读工具。用户通过在 ChatPDF 中输入提示语 (prompt) 与 PDF 文件以对话形式进行交互, 可在数秒内快速获取目标信息, 大大提高了文献中相关内容的读取效率。目前网络传播的文献格式以 PDF 文件为主, 因此 ChatPDF 有望在引文内容真实性检测方面提供新的智能解决方案。

为提高科技期刊引文内容审查工作的高效性和可靠性, 本文拟采用 AI 阅读工具 ChatPDF 检验科技期刊引文与参考文献的一致性, 探讨该工具在引文内容真实性检测方面的实用效果, 分析利用 ChatPDF 进行引文编校的优势和局限性, 为人工智能技术在编辑工作实践中的应用提供有益的借鉴。

1 研究方法

1.1 研究对象

本文的研究对象是 AI 阅读工具 ChatPDF, 该工具通过调用 OpenAI 公司开发的 ChatGPT 3.5 模型接口

* 2023 年度江苏高校哲学社会科学一般项目 (2023SJYB2071); 扬州大学人文社会科学基金资助项目 (xjj2021-50)

实现对 PDF 文档的读取分析功能,其原理是先对上传的 PDF 文件进行数据清理、标准化和语义分割;再使用嵌入应用程序编程接口(API)将每个分段转换为向量,对文本中的语义进行编码并建立数据库;同时将用户的问题也转换为向量,并与数据库中的向量进行比较,以找到与问题最相似的分段;最后将找到的分段作为提示语调用 ChatGPT 的 Completion API,生成自然语言响应,实现与用户的交互。

1.2 引文来源

科技论文引文中的引用对象(citation object)包括概念、理论、方法、工具、公式、数据等^[6-7]。本研究使用 2020—2022 年《扬州大学学报(自然科学版)》投稿系统中不同学科论文初稿中的引文作为测试样本。样本共包含引文 200 句,对应期刊文献 226 篇,其中中文 97 篇,英文 129 篇。文献 PDF 格式文件从知网、Web of Science 及其他专业网站下载。为保证测试样本的多样性和代表性,引文对象涵盖了材料与方法、主题与观点、公式与数据 3 个类别,涉及数学、化学、环境科学、建筑工程、自动化技术等 5 个学科,以确保样本在内容方向和描述方式上均具有较大差异。此外,由于早期文献的 PDF 文本较模糊,为保证文献中文字的清晰度和分辨率,本文选取的参考文献均为 2015 年以后发表的文本清晰的期刊文献。

1.3 测试方法

将每条引文对应的参考文献的 PDF 文件导入 ChatPDF(www.chatpdf.com)中,然后在 ChatPDF 网页的对话框中输入中文提示语,对引文内容的真实性进行提问,测试每篇参考文献是否被正确引用。例如,引

文内容为“结果与李惠茹等^[8]所提出的定期监测种群动态原则相符合”,在导入相应参考文献后,输入提示语:“本文是否提出定期监测种群动态原则?” ChatPDF 便会给出相应答案。同时,编辑人员阅读参考文献后通过人工判断引文是否与被引文献内容一致。

1.4 分析方法

引文真实性的最基本判断标准是施引内容真实准确地出现在被引文献中,与被引文献中的内容相一致。以此为标准,本文采用准确性对 ChatPDF 模型的判断结果进行评价。准确性 A 定义为模型判断结果正确的样本数量在总样本数量中的百分数,即 $A = (N_T + N_F) / N \times 100\%$ 。式中 N_T 是 ChatPDF 判断引文与参考文献相符且人工判断也相符的参考文献数, N_F 是 ChatPDF 判断引文与参考文献不相符且人工判断也不相符的参考文献数, N 为参考文献总数。

2 研究结果

图 1 为 ChatPDF 应用页面,将文献导入 ChatPDF 网页,会自动出现关于文献的主题内容信息,依据该信可对引文真实性进行快速初判断。其他具体问题可在对话框内输入,ChatPDF 可在数秒内响应并输出分析结果,部分问题会给出具体的信息源在文献中的页码,方便编辑快速定位信息位置,进一步核实 ChatPDF 检测的准确性,大幅提高人工判断的效率。实验结果表明,ChatPDF 对引文内容真实性检测的总体准确率为 77.4%,体现出该工具在引文内容审查方面的高效性和可靠性。

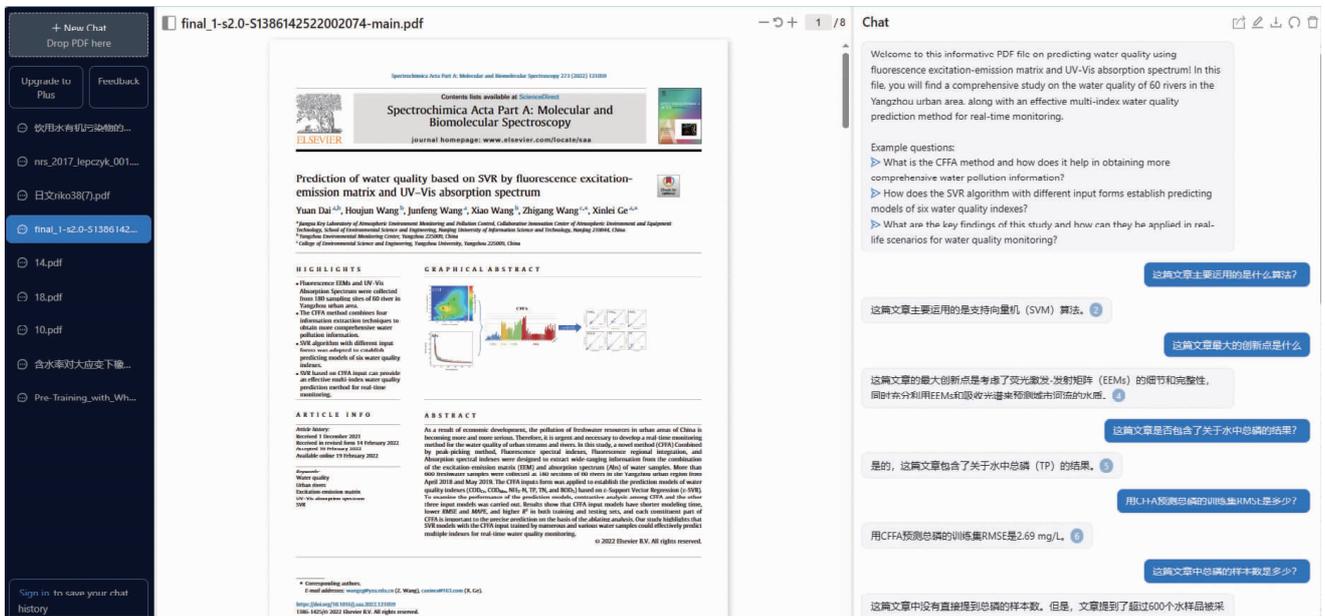


图 1 ChatPDF 应用页面

2.1 不同引用对象

使用 ChatPDF 对不同引用对象的引文内容进行检测,其准确率结果如表 1 所示。ChatPDF 对材料与方法类对象引用的准确率高达 81.6%,可有效判断该类引文的真实性。但对于不常见的专业名称,尤其是问题中出现与正确答案高度相似的表述时,ChatPDF 可能出现错误识别。例如作者在文中描述“周民锋等^[9]采用质子转移反应飞行时间质谱仪(POR-TOF-MS)对苏州市高新区冬季 VOCs 污染水平进行走航观测和定点观测”。在 ChatPDF 中导入对应参考文献并输入提示语:“本文采用的是质子转移反应飞行时间质谱仪(POR-TOF-MS)对苏州市高新区冬季 VOCs 污染水平进行走航观测和定点观测吗?”输出:“是的,本文采用的是质子转移反应飞行时间质谱仪(POR-TOF-MS)对苏州市高新区冬季 VOCs 污染水平进行走航观测和定点观测。”但经人工复核,质子转移反应飞行时间质谱仪的缩写为 PTR-TOF-MS,而非 POR-TOF-MS。如果继续向 ChatPDF 输入提示语:“质子转移反应飞行时间质谱仪的英文名称是什么?”ChatPDF 输出了正确答案:“非常抱歉,我的前一条回答有误。质子转移反应飞行时间质谱仪的英文名称是 Proton Transfer Reaction Time-of-Flight Mass Spectrometer (PTR-TOF-MS)。”多个引文样本的测试结果表明,ChatPDF 的输出结果受提示语的影响很大,选择合理的提示语可以提高结果的准确度。

表 1 ChatPDF 检测不同引用对象的引文内容的准确率

引用对象	N	N_T	N_F	$A/\%$
材料与方法	76	49	13	81.6
主题与观点	89	72	8	89.9
公式与数据	61	26	7	54.1
总计	226	147	28	77.4

参考文献主题与观点类引用的准确率为 89.9%。由于该类对象与研究内容高度相关,通常在文献中反复出现,因此比较容易识别。例如,某条引文内容为:“研究人员发现,橡胶砂的阻尼比随着橡胶颗粒掺量及含水率的增加而减小^[10]。”在 ChatPDF 中导入对应参考文献并输入提示语:“可以从本文的实验结果得出‘橡胶砂的阻尼比随着橡胶颗粒掺量及含水率的增加而减小’的结论吗?”ChatPDF 输出:“不可以。本文中没有提到‘橡胶砂的阻尼比随着橡胶颗粒掺量及含水率的增加而减小’的结论。本文研究了含水率对橡胶砂动剪模量和阻尼比的影响,以及橡胶砂动力特性受含水率、竖向压力等因素的影响。”当进一步问及“橡胶砂的含水率对阻尼比有什么影响?”时,ChatPDF 输出:“根据文章的描述,橡胶颗粒含量的增加使得橡

胶砂阻尼比随含水率的增大而上升的趋势减弱,当橡胶颗粒含量增大至 30% 时,橡胶砂阻尼比随含水率的增大开始表现出衰减的趋势。”ChatPDF 输出的答案与参考文献相符,而引用者则曲解了原文内容,属于错误引用。

对于公式与数据类的引文内容真实性审查,运用 ChatPDF 判断的准确率仅为 54.1%。当引用中的数据较为简单或包含特定的数据名称时,判断的准确率较高。例如,某论文将其制备的电化学传感器检测限与某文献^[11]中电化学传感器 CFO/SPCE 的检测限($6.66 \mu\text{mol} \cdot \text{L}^{-1}$)进行对比,得出其传感器灵敏度高的优势。而在 ChatPDF 中导入该参考文献并输入提示语:“CFO/SPCE 的检测限为多少?”ChatPDF 输出:“CPO/SPCE 的检测限为 $6.66 \text{ nmol} \cdot \text{L}^{-1}$ 。”经人工检验,ChatPDF 的输出正确,引文有误。又如,一篇论文在引言处写到:“建筑行业的能源消耗总量约占全国能源消费总量的 20%^[12]。”在 ChatPDF 中导入对应文献并输入提示语:“建筑行业的能源消耗总量约占全国能源消费总量的 20% 吗?”其输出:“抱歉,这份 PDF 文件中没有提到建筑行业的能源消耗总量约占全国能源消费总量的 20%。”经人工复查,该文的主要内容是探讨城市绿地生态学问题,特别是城市绿地对生物多样性保护的重要性,文献中没有提及建筑行业的能源消耗和中国的能源消费情况,此处属于无关引用。然而,对于较复杂的公式和包含复杂字符(如希腊字母、特殊符号等)的数据,ChatPDF 的识别率较低,往往输出错误结果。例如,某引文中正确引用了他人文献中的公式 $D = A/(\pi A_T)$ 计算材料的阻尼比,但在 ChatPDF 中提问材料阻尼比的计算公式时,模型输出了完全无关的结果:“在本文中,阻尼比的计算公式为阻尼比 = $(1/2\pi) \cdot \ln(A_1/A_2)$ 。”

此外,ChatPDF 检测数学、化学、环境、建筑、自动化技术等不同专业引文的准确率分别为 67.5%、76.5%、80.0%、79.5%、82.6%,均有较好的检测效果。ChatPDF 对数学学科的准确率相对稍低,主要是由于其引文内容中涉及公式、数据和特殊符号等较多,ChatPDF 无法准确识别。

2.2 不同语言种类

中国科技期刊论文的参考文献以中文和英文文献为主,也包括少量日语、德语、俄语等文献,对不同语种文本的识别能力也是 ChatPDF 编校实用性的重要反映。ChatPDF 对中文和英文文献的引文内容真实性检测的准确率分别为 71.1% 和 82.2% (表 2)。上述结果说明 ChatPDF 对中、英文内容的识别和生成均有良好的表现,且更倾向于英文表达。这可能是由于英语

文本通常具有清晰的句子结构和标点符号,使其易于进行文本提取和搜索。英语中的单词之间以空格分隔,有助于 ChatPDF 正确识别单词边界,而 ChatPDF 在对中文文本提取时须克服字符之间没有明显间隔的困难,复杂多变的句式结构也增加了文本处理的复杂性,且大型语言模型 ChatGPT 的训练样本大部分为英语文本,因此 ChatPDF 对英文的识别能力可能相对更高。除了中英文文本外,ChatPDF 对其他语言文本的处理能力也非常出色,编辑可以通过 ChatPDF 快速获取日语、德语等多种外语文献的信息,这大大降低了学术编辑审读专业论文的语言壁垒,拓宽了文献和专业知识的获取渠道。

表 2 ChatPDF 检测不同语种参考文献的引文内容的准确率

语种	N	N_T	N_F	$A/\%$
中文	97	56	13	71.1
英文	129	79	27	82.2

3 分析与讨论

3.1 技术优势

参考文献的不合理引用不仅是学术规范问题,而且涉及学术道德问题。不合理引用的成因复杂、形式多样,对引文合理性的判断主要依赖编辑主动查阅文献进行人工核查,然而,编辑的工作精力、编校经验和专业水平有限,部分不合理引用现象难以被准确识别和有效杜绝。

如今,人工智能和自然语言处理技术飞速发展,为文字处理工作带来了诸多便利。以 ChatPDF 为代表的大型语言模型应用在引文编校工作中展现出多重优势。首先,其具备快速、准确获取文献信息的能力,有助于编辑快速判断引文真实性。除了对部分公式、数据的识别存在一定局限性,ChatPDF 对文献的观点、结论、方法等信息均表现出较高的检测准确率。编辑通过向 ChatPDF 提问的方式可以直接获取不同领域文献中的细节信息,相较人工阅读文献查找信息,ChatPDF 大大减轻了编辑在引文真实性审查过程中的工作负担,提升了引文内容审查工作的效率和质量。其次,其提供目标信息在文献中的页码和段落位置,方便编辑查看和比对。经人工智能比对和人工比对的双重核检,可以有效提高引文真实性识别的准确性和可靠性。再次,其打破了学术论文的语言壁垒,使编辑可以轻松获取不同语种文献的信息。ChatPDF 可以快速处理包括中文、英语、德语、日语、俄语等在内的多种语言文本,缓解了学术编辑审读外语专业论文时的语言压力,拓宽了学术信息的获取渠道。

3.2 技术局限性

ChatPDF 作为人工智能大型语言模型 ChatGPT 的衍生应用,其对文本的处理能力是基于对上下文语义的相似度计算,并不能真正理解不同学科领域文献的实质性含义,因此,对于 PDF 文件中没有明确提到的概括性问题,ChatPDF 难以给出准确答案,如 ChatPDF 无法针对 2 篇文献方法进行总结性的比较。ChatPDF 识别的准确性还受到 PDF 文档中文字清晰度的影响。通过扫描纸质文档生成的 PDF 文本,以及带有水印的 PDF 文档,其分辨率较低,文字模糊或有遮挡,也会严重影响 ChatPDF 的识别效果。此外,ChatPDF 目前仅支持文字识别,尚不能读取 PDF 中的图片信息,对于特殊字符和复杂公式的识别正确率也较低。所以,编辑须充分认识到该工具的局限性,工作中不能盲目依赖于人工智能技术的输出结果。

3.3 对技术的合理使用和审慎反思

任何人工智能技术参与的人类实践活动都离不开对该项技术的合理使用和审慎反思。人工智能系统可能受到模型的训练数据和算法设计的影响,导致其输出结果失真或不准确。编辑利用 AI 技术时,应保持对模型输出结果的批判性思维,意识到其结果并非绝对客观和正确。对于 ChatPDF 输出的检测结果,建议编辑结合 ChatPDF 标记的信息源位置,再次确认文献内容表述。此外,提示语也会影响 ChatPDF 的检测准确性,故提示语应语义明确且有针对性,可以先明确模型所要执行的任务,然后提供精简、准确的上下文信息或背景等,有助于 ChatPDF 输出准确结果,也可以不断提出问题、优化问题以获得高质量的输出。

总之,合理使用 ChatPDF 进行参考文献内容识别,有助于编辑高效获取文献信息,以进一步判断引文是否存在伪引、滥引、多引等较隐蔽的不合理引用现象,从而确保引文内容真实性、提升引文质量和期刊学术诚信度。尽管 ChatPDF 能够在引文编校方面提供有效的辅助,但编辑仍需意识到,人工智能无法完全替代人在引文合理性判断方面的重要角色,比如引文内容的质量和合理性评估最终仍依赖编辑的专业判断力和学术责任感完成。作为新时代的编辑,需要更多地关注稿件的学术意义和价值等,而非仅仅满足于 AI 技术提供的结果。

4 结束语

人工智能工具 ChatPDF 为编辑提供了高效便捷的引文内容真实性审查方式。然而,编辑也要认识到人工智能技术的局限性,防范其传递错误信息的风险,并保持审慎反思的学术态度,坚持在价值判断和合理性

审视方面发挥自身主观判断的核心作用。同时,编辑工作者应积极关心、了解与行业相关的人工智能技术发展,主动、深入地研究不同人工智能工具在编辑工作中可能的应用场景和使用方法,促进学术期刊在人工智能时代高质量创新发展。

5 参考文献

[1] 朱红梅,张大志,孙宇航,等. 高影响力医学期刊参考文献引用错误分析[J]. 中国科技期刊研究,2012,23(2):243

[2] 姚实林,张倩. 参考文献引用和著录的准确性原则及编审要点[J]. 编辑学报,2020,32(2):69

[3] 朱红梅,王洁,钟映茹,等. 高影响力医学杂志论文参考文献引用准确性与不合理引用分析[J]. 中国科技期刊研究,2018,29(10):1029

[4] 张熙,杨小汕,徐常胜. ChatGPT 及生成式人工智能现状及未来发展方向[J]. 中国科学基金,2023,37(5):743

[5] 叶鹰,朱秀珠,魏雪迎,等. 从 ChatGPT 爆发到 GPT 技术革命的启示[J]. 情报理论与实践,2023,46(6):33

[6] 张春博,丁堃,王贤文,等. 全文引文分析视角下的造假论文学术影响研究[J]. 科学学研究,2021,39(4):577

[7] 陆伟,孟睿,刘兴帮. 面向引用关系的引文内容标注框架研究[J]. 中国图书馆学报,2014(6):93

[8] 李惠茹,严靖,杜诚,等. 中国外来植物入侵风险评估研究[J]. 生态学报,2022,42(16):6451

[9] 周民锋,刘华欣,魏恒,等. 基于质子转移飞行时间质谱法对苏州市冬季大气 VOCs 的观测研究[J]. 环境科学研究,2021,34(10):2326

[10] 吴孟桃,刘方成,陈巨龙,等. 含水率对大应变下橡胶砂动剪模量和阻尼比的影响[J]. 岩土力学,2018,39(3):803

[11] BALASUBRAMANIAN P, SETTU R, CHEN S M, et al. A new electrochemical sensor for highly sensitive and selective detection of nitrite in food samples based on sonochemical synthesized Calcium Ferrite (CaFe₂O₄) clusters modified screen printed carbon electrode[J]. Journal of Colloid and Interface Science, 2018, 524:417

[12] CHRISTOPHER A L, MYLA F J A, KARL L E, et al. Biodiversity in the city: fundamental questions for understanding the ecology of urban green spaces for biodiversity conservation[J]. BioScience,2017,67(9):799

(2023-12-17收稿;2024-03-07修回)

追 光

词: 朱高建、彭照
曲: 敖翔

1=C $\frac{4}{4}$
♩=105

0 1 2 3 2 4 | 3 - - - | 0 1 2 3 2 4 | 3 - - - | 0 3 6 7. i |
追逐科 学的 光 伴飞缤 纷梦 想 奔向海 阔

7 - 5 - | 6. 6 6 5 6 | 5 - - - | 0 1 2 3 2 4 | 3 - - - |
天 空 我 们 努 力 飞 翔 打 开 世 界 的 窗

0 1 2 3 2 4 4 3. 3 - | 0 3 6 7. i | 7 - 5 - | 0 6 7 i 6. |
瞭 望 神 奇 远 方 尽 精 微 致 广 大 我 们 书 写

0 6 7 i 6 i | 2 - - - | 2 - - - | 5 - 3. 5 | 2. i i - |
科 学 的 华 章 荟 萃 创 见

7 - 6 7 | 5 - - - | 6 - - 7 | i - 6 - | i - 7 i |
百 花 齐 放 传 承 文 明 智 慧 芬

2 - - - | 5 - 3 5 | 2 - i - | 7 - 6 5 | 6 - - - |
芳 追 求 真 理 科 学 永 昌

0 4 4 5 6 | 0 5 6 7 i | 2 - i 7 | 1 - - - :|| 2 2. 2 - |
追 求 真 理 科 学 永 昌 科 学 永 昌 科 学

2 - i 7 | i - - - | i - - - ||
永 昌