

全国哲学社会科学规划办公室

National Planning Office of Philosophy and Social Science

坚持正确导向 突出国家水准 注重科学管理 服务专家学者

首页

信息 工作要闻 通知公告 社科动态 机构设置 建议 工作 管理规章 课题指南 问题解答 立项信息 项目进展 经费管理 成果管理
成果 成果选介 成果数据库 社科基金专刊 成果要报 交流 学术研讨 国际交流 各地规划管理 资料 专家数据库 常用社科查询

中国共产党新闻 >> 全国哲学社会科学规划办公室 >> 社科动态

中华字库：这里没有你找不到的字

2011年08月24日16:42 来源：《光明日报》

摘要

我国现有的字库只收入了7万多字，已无法满足时代需求。历经5年精心准备的“中华字库”工程进入研发阶段，工程将利用先进的信息技术手段，收集和汇总历代文献资源中出现过的汉字和少数民族文字，辨析源流衍变，确定每个字形的历史地位，建立汉字及少数民族文字的编码和主要字体字符库。字库建成后将是开放性的，会及时吸收新的字或字形。

你是否知道？计算机屏幕上显现的每个字符，背后都有一个对应的被国际标准化组织正式认证的区位编码。如果字符没有编码，它将无法在电脑中显现和存储。而这个字符编码汇聚的地方，就是字库。

我国现有的字库只收入了7万多字，除了遗漏的许多现代汉字外，古汉字和许多少数民族文字也没被纳入进来。而随着电脑应用的日渐广泛，现有的字库已远远不能满足人们各方面的需求。

一条喜讯传来：历经5年精心准备的“中华字库”工程日前正式进入研发阶段。

“一套权威全面的字表”

“中华字库”工程总体组专职副组长张力伟介绍，该工程将利用先进的信息技术手段，收集和汇总历代文献资源中出现过的汉字和少数民族文字，辨析源流衍变，确定每个字形的历史地位，建立汉字及少数民族文字的编码和主要字体字符库，重点研发汉字输入、输出、存储、传输以及兼容等关键技术。

完成后的“中华字库”预计可编码字符数在50万左右，包括汉字古文字约10万、楷书汉字约30万、各少数民族文字约10万。

张力伟表示：“‘中华字库’最终的呈现状态，就是做出一套权威全面的字表来，然后提交给国际标准化组织，在电脑区位中给每个字编码，厂商再据此做出字形，从而满足现实的各种需求。”

重要新闻

>>>>

- 王晓晖：《从怎么看到怎么办》与热点面…
- 2011年度国家社会科学基金重大项目…
- 辛勤耕耘 推出鸿编

图片新闻

>>>>



《中国诗歌通史》鉴定 湖南社科基金项目会议



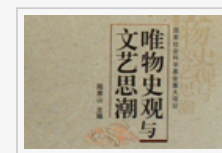
学术研讨

>>>>

- 构建中国民法学体系 掌握世界法学话语…
- 宗教研究关注现代化进程
- 西部地区社会管理关乎国家全局

成果选介

>>>>



《唯物史观与文艺思潮》
中国人民大学陆贵山主
持重大项目“历史唯物史
观与当代文艺思潮”……

- 21世纪中国人口发展战略研究
- 始终不渝奉行互利共赢的开放战略
- 中国西部少数民族地区发展失衡、…

资料下载区

- ▣ 申报资料下载 >>>>
- ▣ 管理资料下载 >>>>
- ▣ 结项资料下载 >>>>

常用社科查询库

- 教育部人文社会科学重点研究基地
- 省部共建人文社会科学重点研究基地
- 国家人类学重点学科名单
- 哲学社会科学类学会
- 人文社科类博士后流动站

“永久的保存和无限的传播”

上古至今的各种传世文献，是中华各族人民的精神财富，具有不可估量的价值。“让我国的传统文献实现‘永久的保存和无限的传播’，是‘中华字库’建成后最重要的意义。”张力伟指出。

张力伟说：“网上已有的一些文献资料，文字的准确度不够，还常常出现拼字、空格、替代等，都是因为字库不完善。古文字中的甲骨文、铜器铭文等，是文化的瑰宝。但由于目前我国的古文字都没有编码，导致现在无法对其实现数字化，录入时只能转写成现代的通行字，转写中字的意义损失很大。因此，从长远来看，要对传统文化典籍进行数字化的存储和传播，就需要建立一个容量更大的‘中华字库’。”

“‘中华字库’工程完成后，不仅对文字学和文献学领域贡献巨大，还将会对我国的整个社会科学研究产生巨大的推动作用。”北京师范大学教授、“中华字库”工程楷书召集组召集人李国英这样说。

“工程实施起来不容易”

日常生活中，有人因为名字中使用了生僻字，在上户口、入学、就医、出国等各种需要电脑录入个人信息的场合出现麻烦。据统计，我国不少的行政村名包含生僻字，而为其村民录入二代身份证信息就会遇到障碍。

“‘中华字库’的研发，本质上虽然是为了传世文献的数字出版，但将连带地解决这些现实应用问题。”张力伟表示。

“字库工程实施起来并不容易！”张力伟感慨地说：“这需要我们做大量工作。某些字在历史上无依据，就不能收录。因此，‘中华字库’工程不仅仅是一个简单的文字整理，更要对我国传统文献进行一番细致的整理。明清之前的文献，我们要统统过一遍，明清之后则将选取有代表性的版本，涉及文字量将达二三十亿之巨。”

“字库将是开放性的，建成后并非一劳永逸”

复旦大学教授、“中华字库”工程首席专家裘锡圭认为，“中华字库”工程性质比较特殊，其研究对象不仅仅是古今汉字，还包括少数民族文字，其本质上又有很浓厚的技术工程性质。因此，技术和文字这两部分工作之间如何配合好，是一个关键问题。

据了解，在承担字库工程的单位中，既有复旦大学、清华大学、中国科学院、中国社会科学院等科研力量强大的高校和科研院所，又有北大方正、汉王科技等技术研发领先的企业。参与字库工程的研发人员中，既有文字、文献学的学者，也有汉字信息处理、计算机技术等方面的专家。

裘锡圭说：“按照原计划，‘中华字库’工程要在5年内完成，时间紧迫。在文字方面，前三四年要做的主要是各种文字的整理工作，后一两年主要是字库集成的工作。”

“过去研究文字，是先抄卡片，然后在书斋里一个字一个字地考订。然而由于文献数量巨大，仅仅用传统手段，‘中华字库’工程是不可能完成的。所幸的是，这几年，图像处理技术、数据库技术和网络通信技术等计算机高科技手段的出现，为‘中华字库’工程的完成提供了可能性。”从工程论证之初就开始参与这项工作，张力伟对此感受深刻。

对这项工程的未来，裘锡圭先生有着更为深远的考虑，“‘中华字库’将是开放性的，

建成后并非一劳永逸。如果在新发现的文字资料里出现了新的字或字形（其实，即使是已有资料里的字，也不可能一次就全部收齐），在相关的研究领域出现了能够补充纠正以往研究的新成果，我们就应该及时加以吸收，使字库得以不断改进、不断完善。理想的状态是，字库工程完成后，还应有经常性的机构为此提供长期的机制保障。这是很有必要的。”

（本报记者 吴 娜）

（责任编辑：秦华）

您是第 35555555555555 位访问者

© 1999-2011 全国哲学社会科学规划办公室 版权所有

地址：北京市西长安街5号5号楼5层 邮政编码：100806 电话：010-66032841 ， 010-66011398

Email : qgskghb@vip.163.com