

把握信息时代脉搏，全力推进语言文字信息化（代序）*

教育部语言文字信息管理司 李字明

一、我国语言文字信息化工作的简要回顾

1. 政府大力推进语言文字信息化

20世纪80年代中期，政府已经将语言文字信息管理工作提上了工作日程。1986年1月召开的全国语言文字工作会议富有远见地提出，“汉语汉字的信息处理是一门新兴的边缘学科，有广阔的前景，加强这方面的研究，对经济、文化、科学技术的发展具有长远的意义。因此，当前语言文字工作的任务必须包括这项内容。”^①

1994年国务院批准的国家语委“三定方案”中，规定了国家语委在语言文字信息处理领域的职责和任务。1995年初，国家语委设立了专门的职能部门中文信息司。同年，在纪念文字改革和现代汉语规范化工作40周年大会上，李岚清副总理代表党中央、国务院明确指出，中文信息处理的应用研究和宏观管理是国家语言文字工作的重要组成部分。在1997年12月召开的全国语言文字工作会议上，李岚清副总理又指出：语言文字领域的专家学者在加强基础研究的同时，要注重应用研究，要与计算机方面的专家继续密切合作，针对汉语、汉字在信息处理中的问题，组织联合攻关，促进中文信息处理技术的进一步突破和发展。

1998年国务院机构改革，国家语委并入教育部，原中文信息司重组为教育部语言文字信息管理司，专司语言文字规范标准建设和汉语汉字、少数民族语言文字信息处理的宏观管理工作。语言文字信息管理司总结以往经验，并在实践中大胆探索，逐步形成了语言文字信息管理工作新思路，即：以语言文字规范标准制定为核心，以信息化为主线，以评测认证为抓手，以语言工程建设和科学研究为基础，争取经过四五年的努力，建成服务于教育和国家信息化的语言文字坚实平台，满足社会对语言文字规范标准的基本需求，进一步提高语言文字信息处理技术及其应用水平，促进社会语言生活健康发展。

2. 我国语言文字信息化工作的主要成就

经过多年的不懈努力，语言文字信息处理技术取得了斐然的成绩。八十年代以来陆续开发的多种汉字键盘输入方法，解决了汉字进入计算机这一难题。汉字自动识别技术达到了国际先进水平，语音识别和语音合成取得明显进展。汉字编码字符集标准的制定工作取得很大成就，《信息交换用汉字编码字符集》基本集及其辅助集的制定，国际标准编码字符集的工作，为中文信息处理技术的发展奠定了基础。汉字实现了多字体、多字号的输出，汉字字库制作由点阵字库、矢量字库逐步过渡到曲线轮廓字库，在字形的描述和处理上也取得较大进展。语言资源库建设取得很大成绩。语言文字信息处理走向对自然语言、特别是大规模真实文本的自动处理。中文电子出版系统取得巨大成就，中国出版业告别了铅与火，步入光与电的新时代。

少数民族语言文字规范化、信息化工作有新的进展。民族语文术语标准化工作取得了一定成绩，常用的民族文字都实现了计算机处理，有些还实现了计算机编辑、排版、检索，开

* 本文根据在全国语言文字信息化工作会议（2003年11月5日，长沙）上的报告修改而成，将此作为第二届全国学生计算语言学研讨会（SWCL2004）论文集的序，以示对会议的支持。

^① 国家的这一决策，是在著名科学家的支持下得以形成和贯彻的。1986年，我国计算机信息系统首席科学家汪成为提出“电子计算机软件也是语言文字工作”。对此，钱学森非常赞同，进一步指出：“这项工作同语言文字工作的联系最密切，国家有关部门应该在这方面为国家上层决策提供设想和参考意见。”

发了各种应用系统,并得到迅速普及,取得了较好的社会效益。几种传统通用的民族文字都制定了编码字符集的国际标准。

汉语言文字和少数民族语言文字信息处理技术的发展,为我国发展计算机应用技术创造了基本条件,促进了计算机在我国的普及和应用,促进了网络的迅速发展。

二、语言文字信息化重要性的认识

1. 语言文字信息化是国家信息化发展的基础

目前,信息化已成为世界发展的潮流,以信息技术为代表的科技革命迅猛发展。现代信息技术的快速发展,将对世界的发展产生巨大而深刻的影响,社会经济、社会结构、社会文化以及劳动者的素质等,都将发生革命性的变化。同时,信息化也带来了世界发展的更大不平衡,产生了“数字鸿沟”,并由此带来了经济、文化、教育等方面的“社会鸿沟”。

语言文字是信息的主要载体,语言文字信息处理的规范化标准化,直接关系到国家的信息化水平,进而影响到我国的国际竞争力。语言文字信息化是国民经济信息化的重要组成部分。信息化的飞速发展,特别是以网络为核心的计算机时代的到来,激发了人们共享信息的需求,对语言文字的信息处理也提出了更高的要求。语言文字工作必须走在信息时代的前沿,不断提高语言文字的规范化标准化水平,推进语言文字信息处理向更高层次发展,为国家信息化的发展提供坚实的基础平台。

2. 语言文字信息化是文化信息化的需要

信息化进程不仅仅是技术进程,更重要的是文化进程。文化信息化是指用数字化手段通过互联网整合现有文化信息并进行传播的过程。我国是一个多民族国家,各个民族都有着悠久灿烂的文化,各民族的传统文​​化都应实现保存形式的现代化和研究手段的现代化,都要积极参与虚拟空间的竞争,使中华文化在信息时代发扬光大。要建立数字化图书馆、博物馆等中国文化的数据库,使中国文化以数字化的形式长久保存,实现在国际互联网上的无障碍交流。为此,必须将负载中华文化的古今全汉字、古今民族文字和各种有用的符号都信息化。

3. 语言文字信息化是构建现代化学习社会的条件

计算机辅助教学已经成为学校教学过程中的重要手段,教师教学及学生自学都越来越多地借助于计算机,而教学手段的现代化,教学软件、自学软件、测试软件的制作,多媒体系统的自动生成等,都离不开对语言文字的信息处理。另外,以互联网为代表的现代远程教育技术正迅速改变着人类的教育与学习方式,我国也在加紧实施现代远程教育工程,构建终身学习体系。实施现代远程教育、大力加强网络教育,除了有赖于相关信息技术的发展,更有赖于语言文字信息处理的水平和语言文字的规范化标准化程度。语言文字信息化是教育现代化的基础和保障。

4. 语言文字信息化是保护和开发语言资源的需要

语言文字是信息的主要载体,语言资源是最重要的信息资源。它和人力资源、地矿资源、国土资源、森林资源、水资源一样,都是国家的重要资源。在信息社会和数字化时代,必须建立语言资源意识,对语言资源的建设、开发、保护,政府应尽到责任。也只有实现了语言文字信息化,才能更好地保护、开发和有效利用语言资源。

5. 语言文字信息化是保护中华语言文字的需要

目前,数字鸿沟在带来社会、经济、文化等方面鸿沟的同时,也带来了世界语言间更大

的不公平，并威胁着许多语言的生存。信息化社会，不能成为网络语言的语种，其生命力将会逐渐丧失；在网络上处于劣势的语种也会逐渐衰落，甚至消亡。未来社会的经济大国，同时也会是语言大国。目前在联合国6种工作语言中，中文远不是强势语言，中国现在也还不是语言大国。我们要加快语言文字信息化的步伐，抢占网络社会的制高点，让汉语成为强势语言，使中国成为语言大国。

语言文字信息化是国家信息化发展的基础，是经济发展重要的增长点，是文化信息化、教育信息化的重要保障。必须深刻认识语言文字信息化对国家信息化的重要意义，创造性地做好语言文字工作。

三、当前语言文字信息化工作的几项任务

语言文字信息化工作，应在已有的基础上巩固、深化、发展、提高，更好地为现代化教育体系服务，为国家信息化建设服务，为和平崛起的国家发展战略服务。为此，当前要做好以下几个方面的工作：

1. 加快面向信息处理的语言文字规范标准的研制

国家的语言文字工作包括“制定语言文字规范标准”和“推行语言文字规范标准”两个方面。语言文字规范标准是社会语言文字应用和进行管理的依据。前任国家语委主任王湛同志，2001年在全国语委办主任会暨普通话水平测试工作汇报会上指出：“语言文字工作的具体任务归结起来大体可以分为两大类，就是制定规范标准和推行规范标准。制定规范标准具有核心性、基础性的地位，只有根据社会语文生活的发展和需要，制定、完善科学的、切合实际的语言文字的规范和标准，推行标准和规范化建设才有基础。”

信息化要求共享和集成，没有规范标准的支持，共享和集成的目的就难以达到，同时也会在信息化的进程中走弯路。因此，要加快面向信息处理的语言文字规范标准的研制，争取在较短时间内使语言文字规范标准建设有较大进展，基本满足社会需要。

2. 推进语言文字工作的信息化、现代化

在信息化时代，提出“语言文字工程”的概念十分重要。为使语言文字工作适应时代要求，走上现代化的道路，必须抓紧语言文字基础工程建设，加强汉语信息处理的资源平台建设，搭建语言文字政策制定、基础研究、应用研究和规范标准建设的坚实平台。大力促进语言文字工作手段的现代化，利用现代技术手段，促进语言文字工作的数字化、网络化。

3. 加强应用语言学的基础研究

语言文字信息处理越是向高层次发展，就越需要语言文字学及其相关学科的联合攻关，需要语言文字学研究的进一步深入，需要语言文字的更高水平的规范化和标准化。

语言文字工作的主管部门，应成为语言文字规范化、标准化的引导者，语言文字信息处理的科研队伍的组织者，核心技术的扶持者。实施资源战略，集中力量建设各种具有自主知识产权的语言资源和软件资源，如具有语法和语义标记的书面语语料库、语音数据库（包括“地方普通话”的语音数据库）、各种信息词典和知识词典等，并通过明晰知识产权和统一规范标准，增加资源的共享与互补。实行开发战略，及时将科研成果转化为产品占领市场，发挥社会效益和经济效益，并通过产品效益支撑科学研究。

4. 组建全国语言文字标准化队伍，积极开展评测工作

具有“团结、奉献、热情、鼎新”精神的语言文字标准化工作队伍，是语言文字标准化

工作的基础，因此要组建全国语言文字标准化技术委员会，有计划地加快国家通用语言文字规范标准制定、修订工作。语言文字标准化技术委员会，不仅是语言文字规范标准制定、审定的组织者和宣传者，同时，也是社会语言生活热点问题的研究者，国家语言文字规划的智囊团。要建立语言文字应用评测认证中心，加强语言文字规范标准的认证与评测工作。通过有效的评测认证，在市场中择优汰劣，并推进企业标准的建设。

5. 大力加强语言文字规范标准的宣传和普及工作

语言文字的信息化需要规范化程度较高的社会语言文字环境。随着计算机在社会生活各个领域的普及，对社会语言文字规范化的要求也越来越高。要大力加强语言文字规范标准的社会宣传和普及工作，加强国民的语言文字规范意识，提升全社会的语言文字规范水平和语言文字能力。

《国民经济和社会发展第十个五年计划纲要》和党的十六大都明确提出，要“加速发展信息产业，大力推进信息化”，要“以信息化带动工业化”，实现跨越式发展。语言文字工作必须把握着信息时代的脉搏，全力推进语言文字的信息化。