

中国互联网网页历史数据使用承诺书

承诺人（单位）：

承诺对象：中国计算机学会（CCF）

首先承诺人对以下事项表示知情：

- 1) 中国计算机学会决定开放共享70多亿篇中国互联网网页文本数据（以下简称CWP200T），以更好地发挥其潜在的研究价值。
- 2) CWP200T来自北京大学网络实验室对CCF的无偿捐赠，是该实验室出于对网页信息保存和教学科研的目的，在2002-2015年间通过Web爬虫搜集的。
- 3) CWP200T所涉及的网页内容跨度约15年（因为2002年也搜到许多2001年的网页），但不保证包含任何特定的内容。
- 4) CWP200T通过向承诺人提供拷贝的形式实现开放共享。CCF不从中谋求任何商业利益，除收取必要的拷贝服务费外，数据本身完全免费。

在此基础上承诺人做出如下承诺：

- 1) 不管是以其原始形态，还是经过了简单的处理，无论是否有牟利动机，承诺人都不得把CWP200T的全部或者部分拷贝或泄露给第三方。如遇第三方有相关请求，请他们联系CCF。
- 2) 尊重国家法律并承担相关责任。如遇国家有关部门要求删除全部或部分拷贝数据，承诺人将会诚实配合执行。
- 3) 尊重公民隐私并承担相关责任。承诺人公开发表的从CWP200T中得到的研究成果只应讨论抽象概念、统计现象和客观规律，而不得直接指出具体个人或其他实体，也不能出现可能与个人和实体相关联的电话号码、电子邮件、地址等信息。上述不应出现的具体信息，承诺人也不得提供或泄露给第三方。
- 4) 除了发表研究得到的抽象概念、统计现象和客观规律之外，CWP200T中的个人、公司企业、和各机构信息，不得用于各类推广、推销和商业运作。

本人（单位）已经完全理解上述内容及条款并承诺遵守。

承诺人（单位）： （盖章）
承诺单位责任人： （签字）
 年 月 日

附：CWP200T拷贝单位信息

编号：

申请单位			
联系人		手机	
通讯地址		邮箱	
数据计划用途：			