

## 《数据加工通讯》—— 数据加工的经验总结

专利检索咨询中心 李小生



### 前言

很荣幸作为数据加工代表参加这次《数据加工通讯》2011年度交流研讨会。我的发言稿题目是“《通讯》-数据加工的经验总结”。《通讯》作为数据加工工作者的良师益友，不仅是我们发表文章的重要载体，也是我们交流业务、相互学习的良好平台。我们可以在这里畅谈经验、表达体会，也可以从这里汲取知识、寻找灵感。

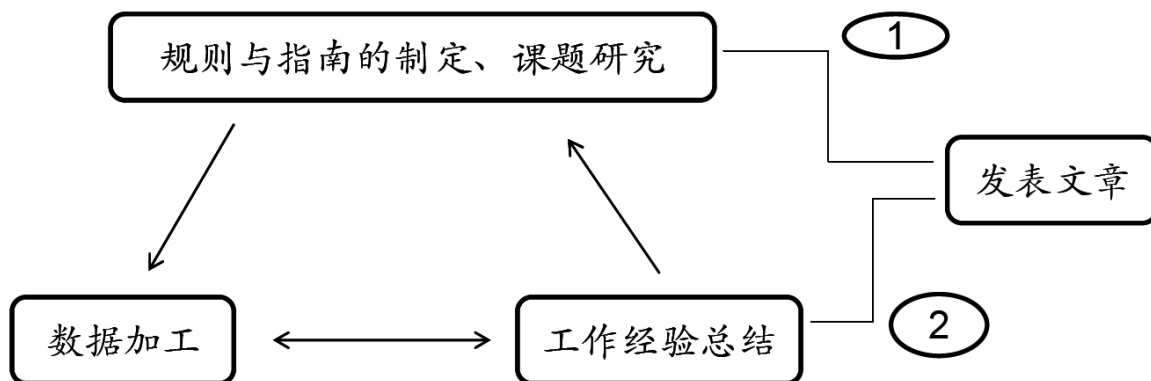
### 数据加工经验总结

下面的图表体现了《通讯》与我们平时的数据加工工作是密切联系

的，有机结合的，并推动着数据加工工作更快更好地发展。首先，在非专利文献数据加工工作中，通过规则和指南的制定以及课题研究，我们获得了第一手的资料，对需要加工的数据进行梳理，发表探索研究类的文章，奠定数据加工工作的基础。利用规则、指南、课题研究成果、以及发表的文章可以多角度地指导数据加工工作，使数据加工工作者能够更加快速准确地把握所要加工的文献。其次，经过一段时间的数据加工实践，数据加工工作者对所加工数据的理解上升到了新的高度，并且不断总结提高，将在

数据加工过程中的经验体会转化成了经验总结性的文章。最后，工作经验总结以及发表的文章作为提升数据加工工作水平的重要理论性资料，指引着我们进一步地完善数据加工规则、

指南，指引着我们更加深入地开展课题研究。所以说，《通讯》是数据加工的经验总结，在我们提高数据加工水平、增强数据加工效率中起到至关重要的作用。



#### 非专利文献数据加工

《通讯》包含的栏目多样，涉及的内容丰富。具体的栏目主要包括业务动态、热点问题、业务探讨、学习与交流、加工书简和数据加工项目进度表，内容主要包括研究探索性文章、经验交流性文章、转载、译文以及信息通讯等。从非专利文献数据加工的角度，《通讯》(1)所选登的文章很多是对标引指南的细化和补充，能够注意到数据加工工作中遇到的细节问题和疑难问题；(2)所选登的文章内容丰富，涉及了非专利文献数据加工中同义词、方剂信息、化合物结构标引等方面的内容；(3)所选登的文章不仅提出了标引工作中遇到的问题，而且提出了解决这些问题的具体方法；(4)所选登的文章对数据加工内部质检工作

具有指导作用。以下从非专利文献数据加工加工条目分类进行举例说明：

1. 化学结构标引：化合物的检索、职能符和构型对化合物标引质量和标引效率有着重要影响。掌握好以上三点，有助于形成正确的标引思路，提高医药类非专利文献数据加工的效率，为专利审查过程中化合物的检索提供一条高效的检索途径。《浅谈医药类非专利文献数据加工中化合物的标引方法》一文中撰写的化合物检索与化合物的构型对标引指南起到较好的补充作用。

2. 化合物的职能符标引：化合物职能符的标引是化合物信息标引中的重要组成部分，通过对化合物的应用和属性的描述，使检索用户可通过结构与职能符结合检索的方式，对技

术方案类型进行区别,提高查准率。《非专利文献数据加工中化合物职能符的含义及标引方法》一文结合具体例子,详细介绍了职能符 N、K、T、P、A、D、E、M 在标引工作中的具体应用。

3. 同义词标引:针对原始文献的关键词未经过规范化处理,以及一词多义或多词一义等原因,导致同一主题的文献,常常由于同义词、近义词、相关词、异形词等而分散于索引的多处,造成标引一致性下降,对于检索来说易造成漏检,影响检索效率。《非专利文献数据加工同义词标引规范化及其在检索中的应用》一文提出了通过改正原文献存在的同义词错误,杜绝标引系统导致的同义词“张冠李戴”和防止标引员误操作、疏忽等导致的同义词错误的途径来提高标引质量。

此外,《医药类中国非专利文献标引的思考》一文结合典型案例,对化合物标引、同义词标引以及方剂信息标引中存在的漏标错标问题进行了详细分析,提出了提高数据深加工质量的若干措施,包括特别注意出现在文章不同位置的同义词、重视典型病例中可能漏标的方剂信息等。

除以上提到文章对非专利文献数据加工工作的帮助以外,《通讯》所选登的文章中有很多与规则、指南制定相关的文章,对规则、指南的完善

具有较大意义。包括《关于非专利文献数据深加工策略的思考》、《中文非专利文献作为专利审查对比文件的情况分析》、《非专利文献数据加工中范例库的建设》、《医药类非专利文献的特点及其加工策略》等。

#### 一点建议

自创刊以来,《通讯》上发表了大量与数据加工有关的文章,我们数据加工员既是作者也是读者,发表自己的看法的同时也聆听的别人的心声。凡是坚持投稿和阅读的数据加工工作者都在伴随着《通讯》一起成长,并一起走向辉煌的未来。一直以来都想为通讯的发展做点什么,仅以几条建议实现自己的心愿:

1. 获取更多的稿源是期刊发展壮大根本所在,宣传自己,增加知名度,通过介绍通讯的办刊方针、方向、所设置栏目、所包括内容等,使全局对本刊有所了解,有利于广泛吸收稿源。这次 2011 年度交流研讨会就有局内很多与数据加工相关的多个部门的同事参与,是一个很好的开端。以后可以经常举办类似的交流会、研讨会,使与数据加工工作相关的同事都能了解这本杂志、阅读这本杂志、进一步的多向这本杂志投稿。

2. 提高稿件质量是期刊发展壮大的前提,可以采用约稿的方式,跟踪各个加工单位的加工动态、提高数

据加工文章的针对性、权威性和实用性。该刊常见的文献部同事与数据加工部门同事共同发表的很多具有指导意义的文章就很好地体现了以上几点。另外，数据加工部门同事发表的文章也对本部门的数据加工工作有较好的指导作用，对其他部门的数据加工工作也有较大的借鉴意义。审查员在该刊发表文章较少，可以向审查员约稿，及时了解用户需求，适时调整

加工策略，这样可以使加工的数据更好地为审查服务。

3. 经常与读者 / 作者保持密切的联系，了解他们的需求，有针对性地刊登一些读者 / 作者迫切需要的内容。

作为数据加工代表，我衷心地希望数据加工工作能越做越好，真诚地祝福《通讯》能越办越好。

(专利文献部 李明 审校)

