

数据挖掘技术在电子文件管理中的应用

黄爱玲

(南昌工学院 江西 南昌 330000)

[摘要]电子文件不是直观的文件形式, 相较传统文件形式, 电子文件具有容易复制和快速多文件拷贝等优势, 它需要借助信息工具来对文件进行读取和管理。然而当前普遍存在管理系统有缺陷、开发利用程度较低以及管理方式不合理等问题, 影响着电子文件的管理效率和水平。而数据挖掘技术作为一种新型的信息技术, 在电子文件管理过程中应用数据挖掘技术可以有效完善电子文件管理流程, 对提升电子文件管理效率具有重要作用。

[关键词] 电子文件; 数据挖掘技术; 文件管理; 应用

[DOI] 10.12252/j.issn.2096-627X.2021.07.278

数据挖掘技术是中高科技信息处理技术, 可以从庞大的数据库中选出有价值、可靠的数据信息, 它有效地解决了大数据背景下的数据爆炸、有价值信息匮乏的问题, 数据挖掘技术包括统计学、人工智能、数据库等方面内容, 通过发现数据与数据之间以及数学与模式之间的关联来进行预测, 方便企业更好地做出决策, 被广泛应用于企业电子文件管理中。本文通过分析当前电子文件管理存在的问题, 重点探讨在数据挖掘技术在电子文件管理中的具体应用。

一、电子文件与数据挖掘技术

电子文件与传统的纸质文件具有很大的差异, 它对系统的依赖性更高, 通过计算机等工具进行文件的传递与处理, 具有很强的可操作性, 方便数据信息的修改和文件共享。因电子文件是数据信息, 故对文件的保存环境要求比较低, 然而却对保存环境的防磁性以及湿度要求较高。近些年来, 在办公自动化背景下, 电子文件数量呈现出几何式增长, 成为当前重要文件形式。

而数据挖掘技术作为新型的数据管理技术, 融合统计学、人工智能以及数据库等领域内容, 改变着人们使用数据的方式, 推进了企业智能化和自动化进程。数据挖掘技术从大量的数据中挖掘出隐藏在数据中的规律, 为人们提供潜在且有价值的信息。因此, 在企业电子文件管理过程中, 数据挖掘技术可有效完善文件管理流程, 解决了文件管理服务质量和效率不高的现状, 有效提升电子文件管理水平。

二、企业电子文件管理存在的问题

随着科学技术与信息技术地不断发展, 电子文件已经被广泛普及应用, 为了适应社会发展的需求, 档案管理相关部门需要思考的问题是如何充分做好电子文件管理工作以提升档案管理效率和质量。虽然我国已经实现了自动化办公, 然而电子文件管理却未形成完整的管理体系, 经常出现文件损坏甚至丢失的情况, 缺少对电子文件的维护管理, 影响电子文件的正常使用, 影响企业的正常决策与管理。同时, 企业常常忽视草稿类电子文件, 尤其是在办公自动化的初期阶段, 许多企业利用计算机编辑文件, 然后打印出纸质文件后没有妥善保管好这些电子草稿文件, 致使未能充分发挥电子草稿文件的作用。另外,

办公系统还有许多有待完善的地方, 电子文件归档程序也不够合理, 主要体现在归档管理有待提高。许多计算机系统尽管有文件归档的功能, 但不能将文件归档到脱机工作的载体, 而计算机系统一旦遇到问题, 所存档的电子文件就会被损坏, 甚至会丢失。因此, 当前电子文件的存储安全性不高, 大多数电子文件存储在计算机硬盘上, 如果计算机系统遇到故障, 电子文件容易出现损坏或丢失现象。并且许多电子文件的存储位置也不够合理, 有的重要文件只存储在U盘中, 而U盘是移动电子类存储器, 出现故障的概率也比较大, 如果移动U盘遇到了故障, 也直接导致文件丢失。

三、电子文件管理中数据挖掘应用可行性分析

企业在长期的业务工作中积累了大量的电子文件, 为数据挖掘技术应用提供了可能, 尤其是在办公自动化背景下, 数据挖掘技术通过整合和转换大量不规范的数据, 提取出有价值的文件信息。而随着我国信息化技术水平不断提升, 人们对档案信息的需求不再局限于简单检索和统计的事务需要层面, 因此企业要充分重视电子文件的使用价值, 利用档案数据信息, 充分挖掘出数据中隐含信息, 全面提升电子文件管理水平, 数据挖掘技术可以深层次分析数据内部结构。而在信息时代, 档案管理部门要基于用户的需求, 拓展各项服务功能, 并提供个性化和多样化服务。故数据挖掘技术作为一种新型的知识发现工具, 同时还是协调各种信息技术、分析各种数据信息的工具和方法, 它能为知识管理提供途径和方法。

自数据挖掘技术提出以来, 通过多年的发展, 已经形成相对成熟的一套技术体系。数据挖掘系统设计从最初的单机独立应用, 直到今天的预言模型和数据管理无缝集成具有高性能的接口, 已经具有非常好的扩展性, 可以在网络背景下多算法操作高度异构数据, 而数据挖掘技术在应用上, 已经在零售业、保险业、银行等多行业广泛应用, 尤其是在图书馆应用上, 通过数据挖掘技术分析用户的借阅习惯, 为图书馆的采购与分类提供参考, 同时还为用户提供个性化定制服务。而在情报学方面应用数据挖掘技术则通过自动收集和分析情报信息, 获得有价值的情报, 最终设计出最佳的解决方案, 为决策服务。这些数据挖掘技术在电子文件相关领域的实践应用, 为数据挖掘技

术更深入、全方位的应用提供参考和借鉴。人们一直在探索和
研究数据挖掘技术的应用,通过长期实践,我国在电子文件管
理已经取得了一定进步,然而仍限于理论探究和尝试时期,我
国在电子文件管理方面的实践探索是不断探索和应用各种理论
和技术的过程,而数据挖掘技术是一种新型的、先进的数据分
析和处理技术,它已被相关研究人员所关注,并被应用在电子
文件管理过程中,用来解决电子文件管理遇到的一些技术性问
题,其中包括数字化信息资源建设、存储等。借助数据挖掘技
术实现数字资源体系构建分析,显著提升电子文件管理效率,
通过深入分析电子文件的有效性及其利用率,建立科学、合理
的电子文件信息体系,而在存储方面,数据挖掘技术是数据处
理分析技术,通过抽取、转换和分析数据库的数据,故那里关键
数据,确保电子文件管理的有效性,其中数据挖掘技术在电子
文件管理上的比较重要的应用就是元数据管理。

从上述分析我们可以发现,在日常工作中,电子文件发
挥着非常重要的作用,而数据挖掘技术有效服务于电子文件管
理,为其提供了技术上的支撑,解决电子文件管理中遇到的一
些技术难题,档案部门要积极探究电子文件管理中数据挖掘技
术的应用,不断提升电子文件管理的有效性和规范性。

四、数据挖掘技术在电子文件管理中的应用措施

1、优化电子文件档案管理

档案管理部分的重要工作内容就是信息服务,而信息服务
的基础是信息资源,信息服务的基本保障则取决于信息技术,
数据挖掘技术通过用户的信息反馈通过关联分析用户提出的
要求、用户遇到资源问题、用户的评价及意见、查询时被拒绝情
况,找到用户的兴趣以及相关资源情况,在所挖掘结果的
基础上,扩大档案资源收集范围,优化档案存储结构及存储质
量,以满足不同受众的个性需求。因此,利用数据挖掘技术可
以有效缓解电子文件管理中的收不到文件或文件收集不全面和
不真实等问题,为优化电子文件档案管理提供结构上的支持。

2、挖掘文件之间的联系

在电子文件整理上,数据挖掘技术可以充分发挥分类、标
引以及联系等技术优势,还可以调整所接收的数据,挖掘电子
文件之间的联系。首先数据挖掘技术可以预处理数据,对其进
行消除噪音处理,将重复的文件去除,并对文件的内容异常、
不一致甚至缺失等问题进行修正与补充,保障电子文件信息
质量。在电子档案管理中,通过数据挖掘技术找出文件之间的
历史联系以及事由联系,确定关联标识,并将文件编上号码。
再通过已建的检索模型来描述文件间的关系,最终建立检索数
据库。通过应用数据挖掘技术大大节省了人工时间,还能挖掘
出文件之间的联系,通过对电子文件进行聚类分割、等距分割
等方式的离散化处理,使数据可以直观呈现在用户面前,为后

续的数据分析奠定基础。

3、精准鉴定电子文件

在档案管理过程中,需要鉴定文件的使用价值,确定文件
的有效性。而鉴定电子文件包括技术鉴定和文件内容鉴定两
方面内容,在当前的电子环境下,要想鉴定文件内容的完整
性和可读性,如果按照传统的鉴定办法则是逐份对文件进行
审查,这样鉴定办法操作比较麻烦,鉴定过程耗时比较长,
同时因文件数量比较大,还伴有文件丢失的风险。而数据挖
掘技术可以通过属性提取和文件的业务活动与单位智能等直
观图表来初步判断电子文件的价值;数据挖掘技术在进一步
提取文件题目和摘要,更加准确地判定文件的使用价值。而
在电子文件技术鉴定过程中,在文件整理时需要提取、修改
以及补充文件的属性,然后就可以精准判定电子文件的完整
性和可读性。

4、提供个性化服务

在档案管理过程中,许多文件都是多年的累积,因此,
档案工作者借助数据挖掘技术在获取用户档案查询记录、浏
览习惯、客户背景等资料,通过对比分析,找到相关背景用
户的浏览习惯、喜好以及浏览内容,然后创建用户兴趣模型。
同时档案工作者还通过数据挖掘技术对用户查询的资料进行
分类和聚类分析,按兴趣大小进行排序,并找到用户感兴趣
的内容及文件。然后按照排序先后推送相关信息,为用户提
供针对性信息服务,同时也给用户提供了个性化的电子文件,
让用户在家中就可以享受优质的电子文件资源服务,为企业
争取更多的客户。

结语

综上所述,在人们的日常工作中电子文件扮演非常重要
的角色,企业在电子文件管理中存在着诸多问题,在电子文
件管理中应用数据挖掘技术,可以给文件管理提供必要的技
术支持,让数据挖掘技术有效地服务于电子文件管理,切实
解决电子文件管理中遇到的各种问题。因此企业应该不断
探索数据挖掘技术在电子文件管理中的应用,逐步提升电子
文件管理的安全性和规范性。

参考文献

- [1]高宇,王金虹,马斌.数据挖掘技术在电子商务管理中的应用[J].集成电路应用,2019,36(04):65-66.
- [2]刘思皖.数据挖掘技术在电子商务中的应用[J].信息记录材料,2021,22(04):207-208.
- [3]刘玉娥.数据挖掘技术在电子商务管理中的应用探讨[J].信息化建设,2016(05):76-77.
- [4]崔素丽.数据挖掘技术在电子商务管理中的应用[J].时代农机,2019,46(03):75-76.