

# 数据管理中的文件档案与内容管理 白皮书

—  
2020年

# 《数据管理中的文件档案与内容管理白皮书》(2020年)

## 编写组

中国人民大学电子文件管理研究中心：刘越男 杨建梁 祁天娇 何思源

上海鸿翼软件技术股份有限公司：龙凌云 罗永秀 张中 姚宝敬

# 目 录

1	背景.....	1
2	文件、档案与内容管理的相关概念.....	1
2.1	文件、档案与文档.....	1
2.2	内容与数据.....	2
2.3	文档管理与内容管理.....	2
2.4	文档生命周期与数据生命周期.....	3
3	数据管理框架下的文件档案与内容管理.....	3
4	文档与内容管理价值与效用.....	5
4.1	安全合规.....	5
4.2	提效降本.....	6
4.3	业务连续.....	6
4.4	决策支持.....	6
4.5	洞察创新.....	6
4.6	权益保障.....	7
4.7	资产增值.....	7
4.8	记忆（历史）留存.....	7
5	内容管理软件的应用及类型.....	7
5.1	应用架构视角下的内容管理软件应用.....	7
5.1.1	企业内容管理系统.....	8
5.1.2	内容协作系统.....	8
5.1.3	网页内容管理系统.....	9
5.1.4	文档管理系统.....	9
5.1.5	影像管理系统.....	10
5.1.6	质量体系文件管理系统.....	10
5.1.7	知识管理系统.....	10
5.2	不同行业领域中的内容管理应用类型.....	10
5.2.1	制造行业内容管理应用.....	11
5.2.2	制药行业内容管理应用.....	12

5.2.3	军工行业内容管理应用.....	13
5.2.4	金融行业内容管理应用.....	13
6	内容管理系统与文件档案管理系统的.....	14
6.1	内容管理系统的发展过程.....	14
6.2	内容管理系统的发展阶段.....	15
6.3	文件档案管理系统的.....	17
7	文件档案与内容管理的应用问题.....	17
7.1	IT 视角下的应用问题.....	17
7.1.1	业务系统核心文件“烟囱割裂”.....	17
7.1.2	离散过程文件难以统一存储.....	18
7.1.3	海量数据带来存储扩展和.....	18
7.1.4	多分支机构对部署架构挑战.....	18
7.2	业务视角下的应用问题.....	18
7.2.1	用户很难发现想要的文件档案.....	18
7.2.2	业务文件不能及时、完整归档.....	19
7.2.3	文件档案未形成体系化知识.....	19
7.2.4	文件档案与业务融合不足.....	19
7.2.5	管理人员综合能力依然欠缺.....	19
7.3	数据视角下的文档档案应用问题.....	20
7.3.1	文件档案核心数字资产地位未被认可.....	20
7.3.2	非结构化数据管理战略规划缺失.....	20
7.3.3	内容数据化能力不足.....	20
7.3.4	文件档案数据价值释放不够.....	20
8	数据管理中的文件档案管理框架.....	21
8.1	价值实现层.....	22
8.2	文档治理层.....	22
8.3	管理活动层.....	23
8.4	技术实现层.....	24
9	数据管理中文件档案与内容管理的发展.....	25

9.1	结构化数据和非结构化数据融合发展 .....	25
9.2	文档管理和内容管理融合发展 .....	26
9.3	文档管理和数字化业务融合发展 .....	26
9.4	文档管理基础建设和智能开发融合发展 .....	27

# 1 背景

随着大数据、人工智能等信息技术的快速发展，数据的价值愈发凸显，人们越来越意识到数据管理对于机构的重要性。然而，根据国内外多个研究机构的调查，大多数机构中非结构化数据占全部数据的 80%以上。数字环境下，大多机构的文件档案管理依赖基于非结构化数据（即本白皮书所称内容）的解决方案，文件档案与内容有着异常紧密的关系。如果说数据是一座煤矿，谁能掌握这座煤矿，谁就能为机构发展提供动能，那么文件档案与内容数据就是潜藏在煤矿深处的钻石，谁能够率先发现这些钻石，谁就能先人一步洞察和刻画机构内外的种种变化，进一步将数据的价值和潜能释放出来。国际数据管理协会在其标志性出版物《数据管理知识体系指南（DAMA-DMBOK2）》（第二版）中将文件档案与内容管理放在了数据管理框架中显著的位置，凸显了文件档案与内容管理对于数据管理的重要性。然而，现阶段我国大多数机构管理者和相关人员对文件档案与内容管理的认识尚有不足，文件档案与内容管理在实践中并没有受到应有的重视。文件档案与内容管理的管理范式与一般的数据管理相比也存在较大差异，相关方法和技术工具尚不丰富。面对现状与问题，编写组特别编制了《数据管理中的文件档案与内容管理白皮书》，为行业从业人员、用户及潜在用户和社会相关人员提供文件档案与内容管理涉及的基本概念、效用价值、主要技术、系统功能等通识性知识，以期加强文件档案管理领域和技术领域的对话和合作，推动文件档案事业与内容管理行业的共同发展。

本白皮书的编制主要参考了《中华人民共和国档案法》（2020）、ISO 15489-1:2016《信息与文献 文件管理-第 1 部分：概念与原则》、GB/T 18894-2016《电子文件归档与电子档案管理规范》、GB/T 36073-2018《数据管理能力成熟度评估模型》、《数据管理知识体系指南（DAMA-DMBOK2）》（第二版）等法律法规、标准规范和文献材料。

## 2 文件、档案与内容管理的相关概念

### 2.1 文件、档案与文档

从文件档案管理视角来看，文件（records）是指机构或个人在履行其法定义

务或开展业务活动过程中形成、接收并维护的作为凭证和具有查考作用的信息，可简称为“业务凭证”。<sup>1</sup>档案（records/archives）是指形成文件的业务活动结束后仍然具有保存价值的文件。文档是文件和档案的合称。

从计算机技术视角来看，文件（file）是作为一个单元存储或处理的命名的记录集。<sup>2</sup>文档（document）主要是指非结构化数据。

除非特别说明，本白皮书使用的文件、档案、文档概念均来自文件档案管理领域。

## 2.2 内容与数据

从信息的角度来看，内容是指以任何形式或载体存在的有含义和背景的信息。在计算机技术领域，内容（content）一般指非结构化数据，是非结构化数据厂商提出的特定概念。在本白皮书中，数据与数字数据同义，是指以数字形式存在的信息记录。按照数据单元被定义的程度，数据一般包括结构化数据、非结构化数据以及介于两者之中的半结构化数据。其中，结构化数据是指存储在数据库里，可以用二维表结构来逻辑表达实现的数据。非结构化数据是指未通过数据模型预先定义的数据。<sup>3</sup>

## 2.3 文档管理与内容管理

文档管理是一项专业工作，指对文件从生成、处理、流转、归档、移交、鉴定、保存到利用、处置等全生命周期进行管控的领域。手工环境下文件形成过程中的管理和档案管理相对独立，数字环境下文档一体化是文件档案管理的基本要求。

内容管理是一个 IT 产品分支，支持对非结构化数据进行采集、处理、组织、存储、查询和共享，实现从内容采集、创建、传递到内容分析等整个内容价值链

---

<sup>1</sup> 中华人民共和国国家标准 GB/T 26162.1-2010/ISO 15489-1:2001 信息与文献 文件管理 第 1 部分：通则 [S]

<sup>2</sup> 中华人民共和国国家标准 GB/T 5271.1-2000/ISO/IEC 2382-1:1993 信息技术 词汇 第 1 部分：基本术语 [S]

<sup>3</sup> DAMA 国际著，DAMA 中国分会翻译组 译. DAMA 数据管理知识体系指南[M]. 北京：机械工业出版社，2020：229

的整合。

从概念来看，文件档案既可能是结构化数据，也可能是非结构化数据。但是在实践中，文件档案管理系统通常采用非结构化数据管理方案，以结构化数据存在的文件或者在归档环节被转化为非结构化数据（比如版式电子发票），以便固化其内容、背景与结构并进行长期保存；或者以数据文件的方式保存在文档系统中，该数据文件可能是多个业务记录的集合体。

虽然内容管理方案是非结构化数据的管理方案，但对内容的管理，需要借助于描述内容的结构化数据（元数据）展开；对内容的分析、挖掘与开发，则需要将非结构化数据转化为结构化数据，这是电子文件数据化的重要任务。结构化数据和非结构化数据，存在互相转化、互相支持、相互协同的关系。

## 2.4 文档生命周期与数据生命周期

生命周期是看待和管理文档、内容和数据的重要视角，它倡导将文档、内容和数据看成一个有始有终、不断发展变化的对象，并根据其不同阶段的特点实施区别化的管理。其中，文档生命周期是指文件从产生到最终销毁或作为档案永久保存的整个运动过程。从文档价值来看，文档生命周期包括现行文件、半现行文件和非现行文件等不同阶段；从管理环节来看，文档生命周期主要包括生成、鉴定、归档、分类、存储、利用、分析、保存和处置等环节。根据 DAMA 的界定，数据生命周期包括计划、设计、创建/接收、存储或处置、利用或复用、增值等环节，一方面，该定义将生命周期提前到数据实际产生之前的规划设计阶段，这与 1997 年国际档案理事会把电子文件划分为设计/概念阶段、形成阶段和维护阶段异曲同工。另一方面，该定义更为强调数据的增值利用，这提醒文档管理人员也要重视文档的复用和增值。

## 3 数据管理框架下的文件档案与内容管理

随着文件档案和内容的属性逐渐明确，价值逐渐得到认可，文件档案与内容管理已经成为数据管理的重要组成部分。在 DAMA 的数据管理功能框架中，文件档案与内容管理被认为是数据生命周期管理中数据“使用与增强”阶段的组

成部分，如图 1 所示，在文字描述部分，综合了非结构化数据和文件档案专业管理的内容。在数据管理的框架下，文件档案与内容管理已经成为机构数字治理和数据资产化的关键环节，是在数据流转过程中数据价值洞察和挖掘的基础性工作。



图 1 DAMA 数据管理功能框架

尽管 DAMA 将文件档案与内容管理放到了一个比较高的位置，但是其提出的数据管理功能框架对于文件档案与内容管理的描述层次比较单薄，对文件档案与内容管理的价值认知有限。DAMA 的数据管理功能框架将文件档案与内容管理描述为数据生命周期管理中一个特定阶段，即在数据管理中的某个阶段，数据管理会表现为文件档案与内容管理。但是，文件档案与内容有着更丰富的内涵，从横向来看，文件档案与内容管理同样存在其生命周期。文件档案与内容的管理过程与其生命周期相对应，包括文件档案与内容数据的产生、捕获、分类索引、访问控制、存储、利用和重用、长期保存以及处置等。文件档案与内容管理并非是数据管理的一个特定阶段，而是贯穿于整个数据生命周期。从纵向来看，对于文件档案和内容的管理也需要着眼数据治理开展规划与监管，并提供元数据管理、风险管理等基础数据管理活动的支撑。

本白皮书认为：文件档案与内容管理并不只是数据管理的一个独立分支，而是渗透和融入数据管理的所有方面。文件档案管理工作因内容管理方案的应用而延展，内容管理方案因文件档案管理的应用而专业。

## 4 文档与内容管理价值与效用

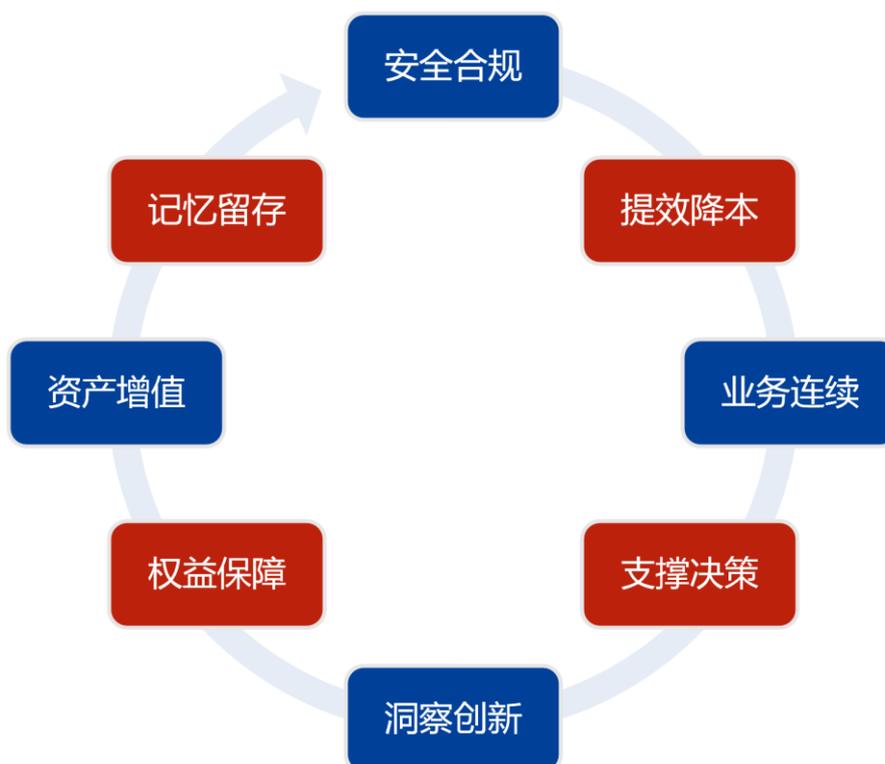


图 2 文件档案与内容管理的价值

### 4.1 安全合规

文件档案与内容管理有利于帮助机构规避法律风险。有效的文件档案与内容管理能够保障机构依法维护机构业务记录的真实性和可靠性，保障证据性记录的长期保管，避免重要记录因管理不善而遗失、泄露或被篡改所带来法律风险，保障机构的合法利益。

## 4.2 提效降本

文件档案与内容管理有利于帮助机构优化业务流程，促进机构高效透明运转，提高机构运行效率。相对于结构化数据，文件档案与内容等非结构化数据更加细致的刻画了业务流程中的事物运行情况，是对业务流程清晰的逻辑映射。有效的文件档案与内容管理，能够帮助管理者更好的洞察机构业务情况，从而提高机构运行效率，实现机构高效透明运转。

## 4.3 业务连续

文件档案与内容管理有利于帮助机构应对突发事件，迅速恢复异常业务。文件档案与内容管理能够将机构运转的重要业务规则和事件予以留存和保管，当机构面对包括自然灾害、业务事故等突发事件而导致业务中断时，有效的文件档案与内容管理能够将数据取出以还原业务，从而帮助机构迅速恢复异常业务，降低突发事件所带来的经济损失。

## 4.4 决策支持

文件档案与内容管理有利于促进信息资源共享整合，提升机构决策水平。文件档案与内容管理强调数据的资源性和业务性，有效的文件档案与内容管理能够帮助机构实现跨部门的信息资源共享和整合，减少部门之间因系统互操作差异等因素所带来的信息壁垒，降低跨部门合作的障碍，从而提升机构的整体决策水平。

## 4.5 洞察创新

文件档案与内容管理有利于提升机构知识发现和创新的能力，帮助机构洞察创新。有效的文件档案与内容管理能够实现对文件档案与内容的有效开发和利用，能够帮助机构充分挖掘和管理文件档案与内容中蕴涵的知识，从而提高机构的知识管理水平，提升机构的知识发现能力，实现机构创新发展。

## 4.6 权益保障

文件档案与内容管理有利于帮助机构留存合法凭证，保管客户法律证据。文件档案与内容管理一方面能够留存和妥善保管机构自身的证据性记录，另一方面还可以帮助机构客户留存相关合法凭证，保管客户法律证据，在机构业务范围内为客户提供可靠的数据管理服务。有效的文件档案与内容管理能够一定程度上提升客户对机构的评价，有利于增强客户关系管理。

## 4.7 资产增值

文件档案与内容管理有利于强化机构信息管控能力，实现机构的信息资产增值。大数据与人工智能时代，信息的资产性越来越被社会所认可。文件档案与内容更加贴近业务流程，在信息流转过程中往往有多个经手人，容易造成信息资产的流失。有效的文件档案与内容管理能够有效避免相关数据泄露和遗失，强化机构对信息资产的整体管控能力，从而实现机构信息资产的保值与增值。

## 4.8 记忆（历史）留存

文件档案与内容管理有利于留存机构记忆，促进文化建设。从时间维度来看，短期来看，文件档案与内容是业务流程和事物的真实、可靠的记录，留存了机构的业务和流程记忆。长期来看，文件档案与内容是机构精神与文化的积淀，是机构文化记忆的载体，有助于促进机构文化的建设，丰厚机构的历史底蕴。

# 5 内容管理软件的应用及类型

## 5.1 应用架构视角下的内容管理软件应用

内容管理（Content Management，简称 CM）是一种以各种类型内容为主体，实现内容创建、编辑、共享流转、利用、保护、管理、业务融合、洞察分析、归档销毁的应用软件集合。内容管理涵盖了企业内容管理、文档管理系统、内容协

作系统、网页内容管理、影像管理系统、数字资产管理、质量体系文件管理、知识管理系统、涉密电子文档管理、内容归档系统、内容安全保护和内容智能应用等软件和系统（如图 3 所示）。接下来针对部分典型应用系统进行介绍。



图 3 内容管理应用全景图

### 5.1.1 企业内容管理系统

企业内容管理（Enterprise Content Management, ECM）是一种战略、方法和工具，通常为平台型产品，通过内容获取、管理、存储、保护、利用等方式挖掘和释放组织流程相关的内容与文档的价值，最终促进机构数字化转型，提升运营效率，并获得业务洞察能力与长远竞争优势。ECM 已有二十多年的发展历史，在国内虽然起步较晚，但近年越来越受到各行各业的重视，在吸取国外发展经验，正朝更成熟更智能的方向发展。

ECM 的核心能力包括：各种形态内容的全生命周期管理；提供统一的非结构化数据平台；全方位的内容数据保护技术；基于内容服务平台与低代码平台，实现与业务的深度融合；内容洞察能力用于知识创新和辅助决策等。

### 5.1.2 内容协作系统

内容协作系统是由内容创作工具、协同编辑组件和分享功能等组成的围绕内

容创作、生成、互动、传递和利用等环节的基础内容管理软件。内容创作工具包括 Office、网页超文本 Wiki、在线笔记、在线流程图、思维导图、Office/CAD 创作集成组件等。协同编辑组件主要是对 Office 等文档实现基于版本控制的协同编辑或在线多人编辑。分享功能解决组织内部成员间、以及组织内外部成员间的内容传递和协作。

内容协作系统强调用户在进行内容创作和使用过程中的体验和交互，不受硬件设备和访问终端的限制，可以随时随地进行创作工作，并且通过全场景的团队协作方式来提升整体内容创作和利用效率。

### 5.1.3 网页内容管理系统

网页内容管理（Web Content Management，WCM）有两个主要应用分支：WCM 和 CMS。WCM 侧重于通过提供应用组件实现面向组织内部员工的信息内容聚合和呈现（在国内也称作：Portal）。CMS 是组织进行面向公众网站的内容创作、发布及管理的系统。随着两个应用分支的发展，两者之间的界限越来越模糊，并呈现出合二为一的趋势。两个分支虽有不同但都是针对 Web 内容的创作、发布和管理。网页内容管理强调创作方便、布局灵活、内容时效、表现丰富、扩展集成、宣传利用等。通过便捷的图文音视频编排即可形成具有视觉冲击和宣传效果的信息页面供用户浏览和利用。

### 5.1.4 文档管理系统

文档管理系统是指对非结构化文档进行捕获和管理并提供长期利用的信息系统，能够实现对文档的集中存储、权限管理、查询索引、文档审计、版本控制、编目编号、安全保护、规则应用、存储加密、数据备份、开发利用、统计编研等基本功能，强调对文档进行全生命周期管理。

文档管理系统可以实现对电子邮件、文件、表格、图片、账单、文本记录、会议纪要、设计图纸、合同方案、单据报告、音视频素材等与经营活动相关的各类文件档案进行统一存储、统一管理和统一利用。文档管理系统主要提供了文档收集、文档整理、文档价值鉴定、文档保管、文档编目、文档检索、文档统计和文档编辑研究等功能应用。

### 5.1.5 影像管理系统

影像管理系统（Imaging Management System, IMS）是在业务活动中将纸质原始凭证进行扫描生成电子影像文件，进行统一管理和保护，并通过文字识别和提取技术获取影像文件中的文字信息进而规范管理和利用的系统。影像管理系统主要包含三大模块：扫描仪、文字识别模块和文档管理系统。扫描仪将纸质单据凭证进行电子化和影像化；文字识别模块通过 OCR 技术提取影像文字进行价值信息获取；文档管理系统则对电子影像文件进行安全存储、共享分发、查询搜索和电子化流程管理等后续操作。影像管理系统在金融投资、保险信托、财务会计、通信、政府机关等产生大量纸质文档、票据的行业广泛的应用。

### 5.1.6 质量体系文件管理系统

质量体系文件管理（ISO）是对体系文件的新增、发布、签收、培训、复审修订、废止等进行全生命周期管理。质量体系文件是建立并保持企业开展质量管理和质量保证的重要基础，是质量体系审核和质量体系认证的主要依据。主要管理四类文件：质量手册、程序文件、指导书和表单记录。质量体系文件管理在制造业和医药制造方面具有普遍的需求和广泛的应用。

### 5.1.7 知识管理系统

知识管理（Knowledge Management, KM）是对知识、知识创造过程和知识的应用进行规划和管理的活动。知识管理系统是在组织中构建的知识系统，让组织中的内容与知识，在获得、创造、分享、整合、记录、存取、更新、创新等过程中形成永不间断的累积，成为管理与应用的智慧资本。知识管理系统的规划与建设工作包括：建立知识统一存储平台、建立多维知识呈现体系、构建知识管理蓝图、建立社区化网络知识管理体系等。

## 5.2 不同行业领域中的内容管理应用类型

内容管理软件在各行各业中被广泛应用，其中以制造业、医疗制药、工程建设、建筑设计、军工、汽车、银行金融、房地产、教育、航天航空、IT 互联网、

媒体出版、零售电商、运输、通讯、政府机关和公益组织等更为突出（如图 4 所示）。不同行业的内容管理应用涉及到的软件系统组合各有侧重和特色，接下来针对部分典型行业进行介绍。



图 4 内容管理行业全景图

### 5.2.1 制造行业内容管理应用

中国是世界工厂，需要实现全球化接轨并通过数字化转型以支撑制造业的深层次发展。在这一进程中，制造行业必须解决海量业务中非结构化数据的高速互通、全球供应链的文档数据统一存储及一致性要求、快速扩展的员工协作要求、全生命周期的质量体系管控、技术及体系文件的外发协作、跨区域及体系的研发协作、数据安全和合规性保证、统一数据保证业务连贯性、随需的业务访问请求及多系统整合等问题，才能有效应对企业内容管理带来的挑战。因此，制造行业内容管理在底层架构层面需要以 PB 级存储为支撑，并且进行多数据中心部署、集团化架构、多站点部署、跨国部署、多系统集成、全球加速和实时数据同步。应用层面涵盖：文档管理系统、ISO 质量体系文件管理、文件档案一体化管理、研发文档安全防扩散管理、文档外发交换管理、知识管理、纸质文件捕获管理、

图文档管理等。通过这些内容管理应用连接起企业各个部门，实现制造升级和智能制造，如图 5 所示。

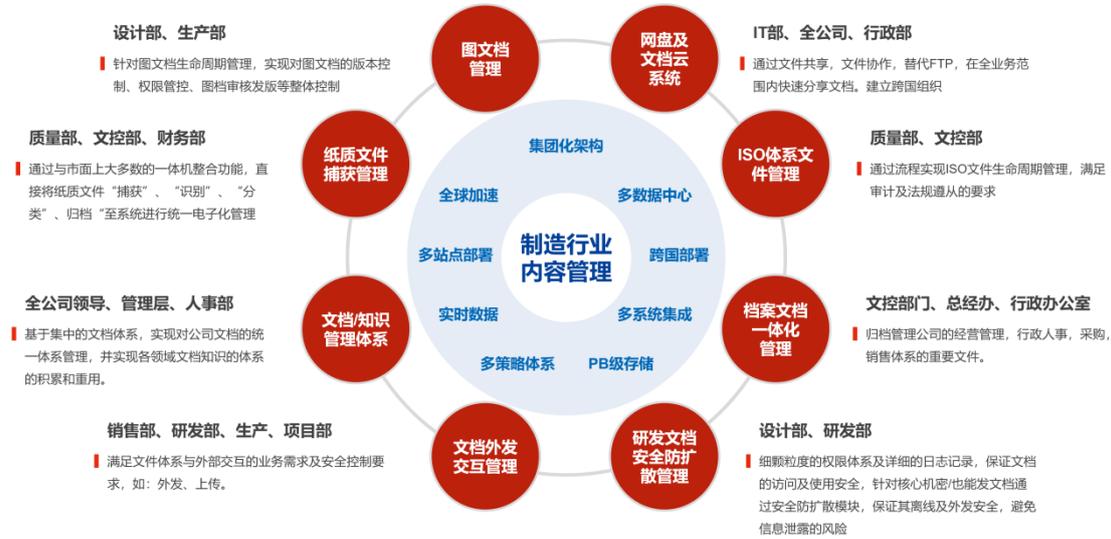


图 5 制造行业内容管理应用

## 5.2.2 制药行业内容管理应用

制药行业内容管理整合了文档管理系统、质量管理平台、资源计划管理、实验室信息管理平台、生产执行系统、记录管理系统、培训管理、档案管理系统等，共同构成全面的质量管理体系。规范体系文件的无纸化生命周期过程，提供涵盖文档、记录、档案及培训等全业务过程的制药管理应用，建立集中、安全的文件管理以及可追溯体系，同时借助系统实现业务数据的采集和分析，如图 6 所示。



图 6 制药行业内容管理应用

### 5.2.3 军工行业内容管理应用

军工行业负责国家武器装备的科研、生产和配套等工作，对数据保密性、安全性具有非常高的要求。军工行业承接大量的以国防建设、军事工业发展为目的的各类项目，项目过程中的项目数据、办公数据、终端数据、业务数据和测试数据等内容需要加以管理和利用。完整的军工行业内容管理应用体系是以电子文档管理为基础，并在电子文档管理系统中通过分权管理体系、密级管控模块、智能定密模块、文件不落地模块等保障数据的安全和保密；在电子文档管理基础上结合网间文件安全交换，保障数据流转和使用过程中的安全。在上层通过项目文档管理应用进行项目过程及内容管理；通过质量体系文件管理应用对武器装备制造环节的质量进行严格规范地管控；通过文件档案一体化应用和知识管理应用将数据结构化和体系化，挖掘数据价值促进产品改良和组织创新，如图 7 所示。

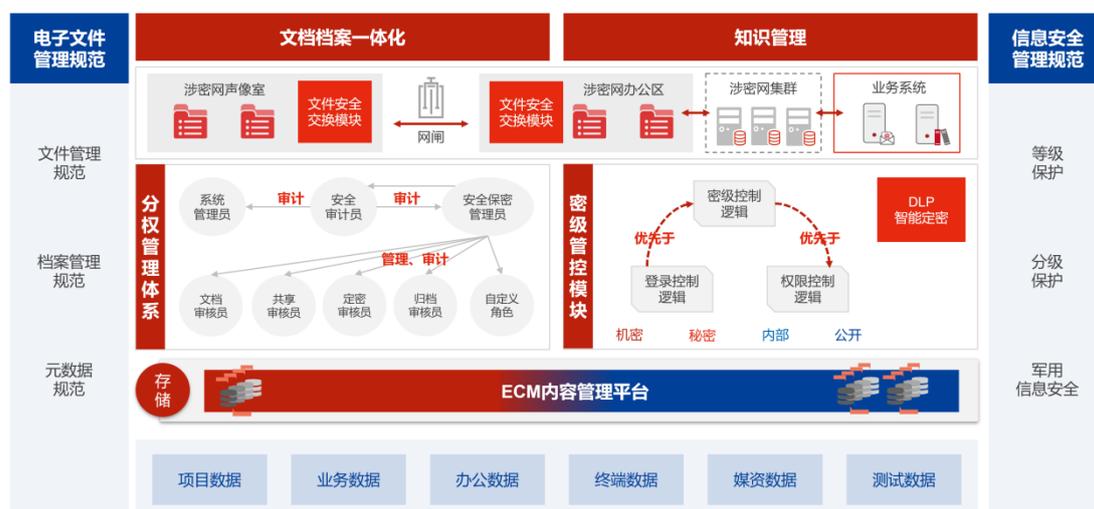


图 7 军工行业内容管理应用

### 5.2.4 金融行业内容管理应用

金融行业拥有大量的产品信息、CIF 信息、CRM 信息、合同资料、员工资料、财务报告等非结构化数据，这些数据作为企业资产需要进行体系化管理。金融行业的网络架构以生产网、办公网、互联网分级网络进行管理，同时根据国家档案局 2015 年通过并施行的《金融企业业务档案管理规定》要求，金融行业需要对自身业务办理及活动中形成的各种电子文件和档案记录进行统一规范的管理和应用。较为典型的金融行业内容管理底层以非结构化数据中台作为数据存储中心和基础平台，在非结构化数据中台的基础上搭载安全网盘应用、文件安全交换应用、影像管理应用和知识管理应用等。其中，文件安全交换应用解决金融企

业多级隔离网络间的数据交换和摆渡需求；影像管理应用通过扫描仪捕获业务单据凭证等影像资料并基于文字识别模块进行数据抽取和业务流转；知识管理应用通过挖掘企业存量数据资产和分析外部市场环境促进金融体制或金融工具的创新。办公网或互联网用户通过 VDI 安全云桌面接入方式进行网盘数据资源访问，保障云端数据不外泄。同时从风险评估、风险控制、运行监控和应急恢复方面进行全方位的风险预警和处理，如图 8 所示。

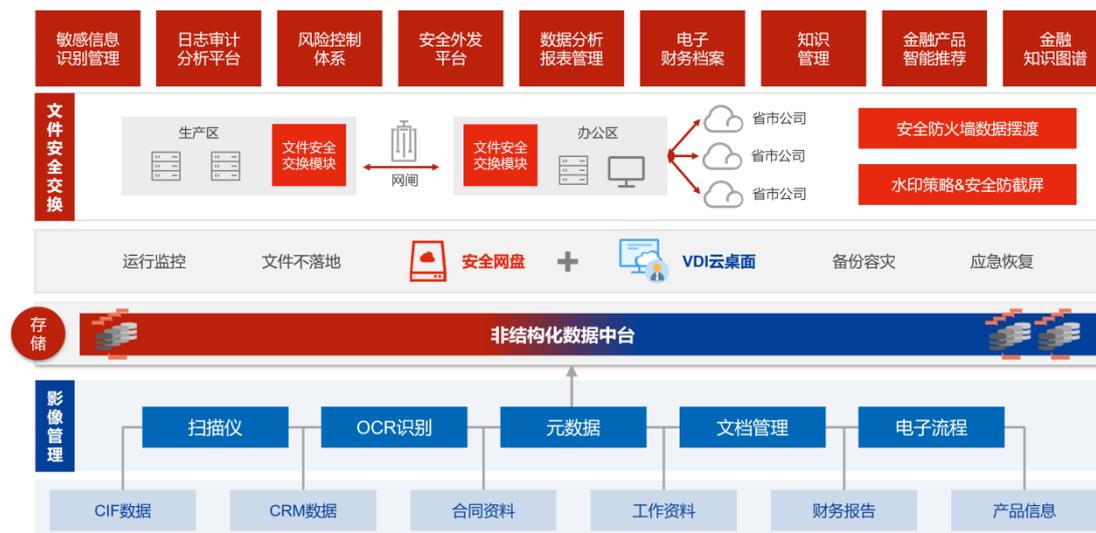


图 8 金融行业内容管理应用

## 6 内容管理系统与文件档案管理系统的发展

### 6.1 内容管理系统的发展过程

随着计算机技术的普及，办公自动化的推行，非结构化数据日益增加，国内外内容管理加速发展。国际上，1990 年，Documentum 公司成立，成为第一家利用标准关系型数据库技术以及面向对象方法提供企业级文档管理解决方案的公司；2000 年左右，以电子商务和电子政务为代表的门户网站的发展带来了网页内容的指数级增长，促进了网页内容管理的成熟与发展；2002 年 Documentum 正式发布 ECM 产品；2006 年，Sharepoint Portal Server 发布；2010 年，OpenText 发布；2010 年后，云计算、移动互联网、大数据等新技术改变了 ECM 的形式与内容，ECM 的内涵与外延不断更新。国内来看，2002 年起，航空、核电和工程领域的国家 ECM 一线厂商开始进入我国，首先在这些行业内掀起 ECM 潮流；2008 年，上海鸿翼软件技术股份有限公司发布国内首款完整 ECM 产品“鸿翼

edoc2 ECM”；2009年拓尔思信息技术股份在金融和政府领域推出 WCM 产品；2010年信雅达系统工程股份有限公司在金融领域推出 ECM 影像管理产品；2016年之后，应用层的网盘和功能更全面的 ECM 出现，ECM 系统中的文档协同和服务能力不断提升；2017年开始，人工智能（AI）在 ECM 系统中的逐渐显现，例如利用 AI 实现元数据自动化提取等（如图 9 所示）。

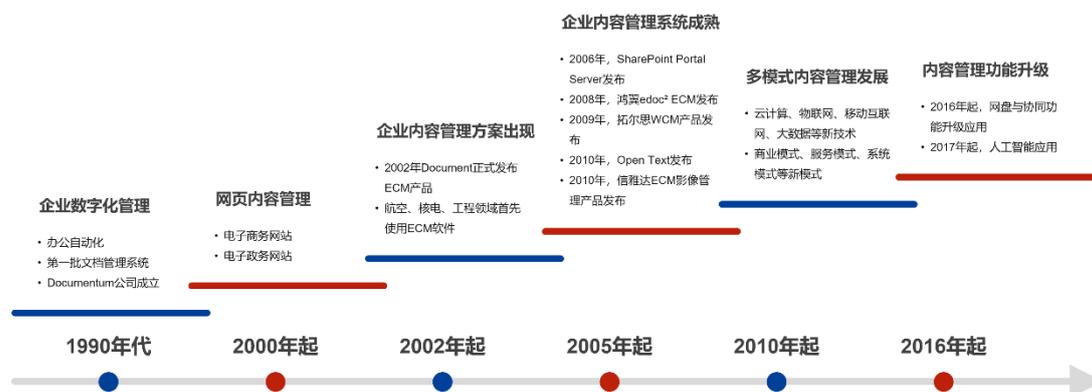


图 9 内容管理系统发展时间线

## 6.2 内容管理系统的发展阶段

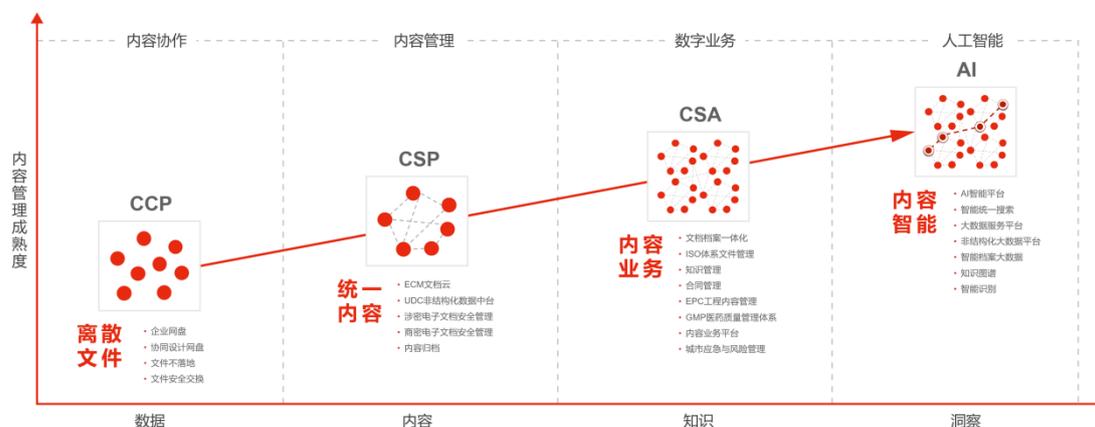


图 10 内容管理系统四阶段进阶模型

按照其与机构业务的结合程度，内容管理系统的发展可分为四个阶段：

(1) 内容协作阶段。此阶段文档离散分布在机构内 PC 端，仅能通过系统完成文档之间的协作，不能实现内容与机构业务的融合。

(2) 内容服务阶段。此阶段机构数据以业务系统文件（比如 SAP 里票据影像）、体系文件（比如 SOP 作业标准文档）等形式存在，通过元数据进行网状式广泛关联，数据来源于业务，又输出服务于业务。数据经汇聚、整理、处理后，以全内容服务形式开放，构建起一个统一的机构非结构化数据管理平台。

(3) 数字业务阶段。此阶段，行业数据经过不同维度的整理、提炼，围绕业务的垂直领域性、体系性进行立体式关联与聚合，形成了行业性的知识体系，对企业迅速应对市场变化和进行业务创新进行支持。

(4) 智能服务阶段。此阶段通过深度学习，自然语言处理（NLP）、知识图谱等技术对海量数据进行深度整理，能够立足业务场景提供内容智能化服务，为机构降本增效、决策制定提供支持。

表1 内容管理系统四阶段特征

	一阶	二阶	三阶	四阶
数据形态	用户电脑上文件，过程性、离散文件	业务系统文件（比如SAP里发票影像），体系文件（比如SOP作业标准）	经过整理、提炼后的行业领域性数据，从不同维度视角归纳后的数据	海量数据，标注数据
数据颗粒度	文件级	内容级	知识级	语义级
数据间关系	离散的，不关联	通过内容元数据进行网状式广泛关联	围绕业务的垂直领域性、体系性关联，是一种能立体式关联与聚合	语义层关联，主题图谱和实体知识图谱
业务融合	满足文件协作，无业务融合	数据来源于业务，又输出服务于业务	深度融合业务，甚至形成业务应用	立足于业务场景上内容智能
数据安全	协作过程中的数据流动，需要存储、访问、流转等安全保护能力，属于企业安全防护初级阶段	全形态数据防护覆盖；内容级颗粒度安全措施更精准；基于统一存储的全生命周期全方位安全防护体系	业务场景化渗透性安全防控，安全紧随业务，为业务中数据安全保驾护航	多层面图谱结合用户画像和操作日志，深度追溯数据安全相关问题，洞察预测安全风险
数字化能力	文件协作层面的数字化	内容中台数字化，数据经汇聚、整理、处理后，以全内容服务形式开放	业务数字化，基于低代码平台形成垂直领域的业务体系化应用，支撑企业迅速应对市场变化和进行业务创新探索	智能数字化，利用NLP、AI、图谱与大数据技术为企业降本增效、辅助决策

## 6.3 文件档案管理系统的发展

文件档案管理系统既可能是一个单独的内容管理应用系统，也可能是企业内容管理平台的应用，还可能是某些内容管理软件的部分功能。

20 世纪 80 年代，很多机构开发了辅助纸质文档管理的早期文档管理系统，主要用于档案目录信息的集成、检索与统计等；90 年代，办公自动化进程加快驱动电子文件管理系统的出现与发展，功能聚焦在电子文件的捕获归档、存储、长期保存和检索等。到 21 世纪初，文件档案管理系统已经成为信息化程度较高或知识密集型企业的必备系统，数字档案馆概念兴起，传统纸质文档的数字化、元数据著录以及长期保存等问题成为这一阶段旨在解决的主要问题。2010 年之后信息技术在业务领域的深入应用将电子文件、档案管理系统建设推向纵深化，面向电子文件单轨运行和电子档案单套保存的试点系统不断出现，以民生档案跨馆出证为代表的跨机构、跨区域的档案数据整合服务应用不断普及。2020 年《新档案法》发布，电子文件单轨制管理走向合法合规，电子文件管理系统与业务系统的无缝挂接成为必然趋势。同时，随着文档资源管理与开发能力的提升，以用户为中心、面向内容和知识的文档管理系统成为新的趋势，更加强调对文档全生命周期的管理以及基于文档管理的内容共享与知识服务。

# 7 文件档案与内容管理的应用问题

## 7.1 IT 视角下的应用问题

### 7.1.1 业务系统核心文件“烟囱割裂”

为了高效完成业务推进，机构业务会不断细分，伴随业务细分而来的是每个环节的业务单元都在逐步数字化和系统化。在系统的使用过程中，核心的业务系统文件被分散在各个业务系统中的，每个业务系统的数据体系也会愈发复杂和深化，形成一个个高耸林立的数据“烟囱”，各个“烟囱”之间没有搭建互通的桥梁，机构数据体系在这种架构下，呈现出割裂的“烟囱群”情形。“烟囱割裂”式的业务系统给文档管理带来巨大挑战。每个业务系统的数据体系都异常复杂和庞大，离散在各个细分业务系统中的文档数据变得不可知、不可控、不可取、不可联和不可信。

## 7.1.2 离散过程文件难以统一存储

组织的过程文件一般都离散存储于在各种文件服务器、邮件系统、个人电脑、运行设备和文件柜中，其涵盖了围绕机构业务活动产生的各种往来邮件、Office 文档、会议纪要、设计图纸、音视频文件等。由于其形成环节众多，过程文件往往分布广泛且脱离控制，加之体量庞大，造成了离散过程文件的采集汇聚极为困难。

## 7.1.3 海量数据带来存储扩展和安全问题

据 IDC 预测每年全球数据的增长量约为 40%。随着互联网、移动互联网、物联网技术的发展，数据产生途径已延伸至更广泛的用户群体、各种机器设备和多样化的现实世界。海量数据在不间断产生，业务连续性对存储动态扩展的要求，加上勒索蠕虫病毒的越发猖狂等，都给文档管理带来了严峻的存储和安全挑战。

## 7.1.4 多分支机构对部署架构挑战

多分支机构对文件档案管理体系的部署架构带来多方面挑战。第一个挑战就是多分支机构和庞大的用户群给系统架构带来的高并发压力。第二个挑战是在不同安全等级的分支机构间如何快速实现数据摆渡和高效地传递利用。第三个挑战是如何满足全球各地的分支机构都可以快速、低延迟地进行文件档案资源访问和利用。

# 7.2 业务视角下的应用问题

## 7.2.1 用户很难发现想要的文件档案

用户获取文件档案资源一般路径是从分类目录逐层进入发现、元数据检索、内容检索等。然而，文档分类规则往往并不严格和规范，欠缺业务传导途径。元数据也只是涉及非常有限的业务信息，且元数据质量经常不可控，这些都大大限制了通过元数据进行内容发现的效果。而对于内容检索，普遍性存在搜不到、搜不准、搜不完整等问题。

## 7.2.2 业务文件不能及时、完整归档

各种业务系统会产生大量的文件需要统一管理。但是在业务系统与文件档案管理系统对接的过程中，往往由于档案策略没有前置、与业务系统整合难度大、整合技术能力有限、业务人员与档案管理人员缺乏有效沟通等问题，造成业务文件不能及时、完整地归档。

## 7.2.3 文件档案未形成体系化知识

文件档案是机构在业务运营过程中形成和沉淀的显性知识，是数据中最宝贵、最有价值的部分。然而，在机构开展知识组织、知识挖掘过程中，往往过度重视人与隐性知识挖掘的重要性，而忽视了文件档案显性知识的价值创造。另外，文件档案管理系统与业务系统进行集成整合程度较低，文件数据难汇聚一起；加上非结构化数据开发与挖掘的技术手段有限，导致文件档案知识化效果较差。

## 7.2.4 文件档案与业务融合不足

文件档案数据及其服务未能嵌入至具体应用场景中，未能成为有效而全面的业务支持数据，更没有形成文件档案驱动业务的管理理念。文件档案与业务融合的不足导致业务无法从文件档案数据中汲取知识资源以进行快速响应和业务决策。

## 7.2.5 管理人员综合能力依然欠缺

在文件档案管理过程中，相关人员需要非常高的综合能力，包括熟悉信息对象、业务管理、档案管理、数据管理、信息化技术等。综合能力的欠缺会给文件档案管理造成几种常见困难：其一是在建设者欠缺综合理解情况下，无法基于行业产品和技术特色来准确描述自身需求，从而造成系统实现与应用场景产生较大偏差；其二是建设者更多地关注 IT 技术和数据管理而忽视业务场景的需求，导致文件档案管理系统无法完全满足业务使用；其三是文档管理者欠缺技术层面的考虑而导致系统架构支撑不力、安全性不足、难以扩展等问题。

## 7.3 数据视角下的文档档案应用问题

### 7.3.1 文件档案核心数字资产地位未被认可

当下，文件档案作为资产所发挥出的效用和价值并未得到明显的关注和认可。然而随着各种非结构化数据处理和分析技术的发展，文件档案将渐渐发挥出非常关键的重要价值。不过，这种转变还需要些时间，文件档案的核心数字资产地位在大多数的组织机构中仍未被认可，给文件档案管理带来诸多不利影响。

### 7.3.2 非结构化数据管理战略规划缺失

机构数据管理实践中，非结构化数据意识相对淡薄。随着中央文件将“数据”纳入为核心生产要素，各界日益关注与重视非结构数据管理。完整的非结构化数据管理战略包含现状评估、业务效率、安全合规、数据文化和洞察创新等方面内容，机构若缺少非结构化数据管理战略顶层规划，也必然导致文件档案在组织机构内的不被重视及其开发利用不足。

### 7.3.3 内容数据化能力不足

非结构化的文件是一种组织无序、难分析的大颗粒度数据，且不同类型文件其数据组织、数据形成、数据呈现等都各不一样，从而造成了内容管理系统中的文件档案没有类似结构化数据的统一数据模型，也无法在其上层进行丰富的算术运算、逻辑分析和数据挖掘，更没有形成完善的技术生态体系。

虽然基于元数据的著录、标引等都是内容数据化的基本手段，但都属于元数据应用的“浅层”状态；目前通过 AI 与图谱技术对文件档案可进行一定程度的内容结构化与内容关联，但仍处于培育与早期发展阶段，无法实现全面而深层的结构化。内容数据化能力不足是当前文件档案的利用开发不足、分析探索不够、安全管控不力的最关键原因。

### 7.3.4 文件档案数据价值释放不够

本应是机构最具有价值的文件档案，却面临着“档案只是用于法规遵从的档案”、“数据中心建设与档案无关”、“档案人员与 IT、数据人员的语言不通”等基

础而又普遍存在的问题，造成多年来文件档案开发利用手段有限、能力不足，致使本属于高含金量的数据未得到应有的价值传递与释放。

## 8 数据管理中的文件档案管理框架



图 11 数据管理中的文件档案管理框架

根据国内外文件档案管理标准以及 DAMA 数据管理框架，结合文件档案管理领域的实践进展，可总结提炼出文件档案管理框架，如图 11 所示。该框架自上而下包括四个部分，分别是价值实现层、文档治理层、管理活动层和技术实现层。价值实现层是文档管理的价值目标和整体定位，是方向性的，处于最顶层；文档治理层涉及战略、职责、制度、沟通和监督等，是宏观层面的统筹考量；管理活动层既涉及文件档案生命周期中的流程性管理活动，也涉及贯穿全生命周期的管理活动，是微观层面的操作实施；技术实现层是支撑其他层次的基础和保障。图 11 中的箭头代表文档管理的动力，其中，左侧的箭头表示文档管理是需求引导、价值驱动的，右侧的箭头表示文档管理同时也是技术赋能、系统驱动的。

## 8.1 价值实现层

在数据管理框架下，文档管理的核心价值包括合规与风控、协同与高效、业务连续性、创新与决策、文化记忆等五个方面。**合规与风控**是合规管理和风险控制的简称，即满足机构的法规遵从要求，并能以文档的形式证明自身的合规性，规避违反安全、隐私相关规定带来的风险；**协同与高效**是指文档的高效协作和管理，能够减轻机构运作的压力，提高业务工作效率；**业务连续性**是指通过沉淀机构核心业务信息资产（即文档），在发生突发事故时能够迅速恢复业务，防止业务中断；**创新与决策**是指基于人工智能和大数据等技术实现文档的智能挖掘、分析、利用、展示、呈现和洞察，提高机构的创新水平和决策能力；**文化记忆**是指文档管理不仅具有业务方面的价值，还有文化传承方面的贡献，基于文档的开发利用能有效盘活机构记忆、传承组织机构的精神和文化内核。

## 8.2 文档治理层

文档治理层是指在组织与文化变革的背景下，重点关注战略管理、组织与职责、制度与流程、沟通与协调、监督与评估等方面。其中，**战略管理**包括规划、实施和评估。战略规划应当明确文档管理的理念、价值观和愿景，组织内外部的法律和业务环境，目标、使命陈述和评价手段，工作内容和时间安排以及制定规划的路径等；战略实施应当根据现状和发展目标之间的差距，确定工作内容的优先级，提供资源和资金保障，推动战略实施；战略评估是指在战略实施过程中跟踪进度，根据评估的反馈结果适时修订和更新规划。

**组织与职责**包括组织架构、岗位设置、文档管理责任和团队建设等内容。其中，组织架构是指建立权责明确且内部沟通顺畅的文档管理组织；岗位设置是指建立文档管理相关的岗位，明确岗位职责及任职要求；文档管理责任不仅是文档管理专业人员的责任，也是组织机构内所有员工的职责，应当明确文档管理专业人员、高层领导、管理者、系统开发与设计人员、全体员工在文档全生命周期中的责任；团队建设包括建立文档管理能力体系，制定系统且持续的团队培训和能力提升计划，确保相关人员具备履职所需的技能。

**制度与流程**是指组织机构应建立完善的文档管理制度规范体系，并按照制定、发布、宣贯、实施、评估和反馈的管理流程进行制度规范的检查和更新，以确保各项文档管理活动的规范有序开展。一般来说，制度规范体系应包括政策、办法和细则三个层级，规定文档生命周期各项活动的目标、遵循的行动原则、完成的明确任务、实行的工作方式和责任主体、采取的一般步骤和具体措施，以及与文档管理系统建设运维、信息安全、人员与岗位职责等相关的管理要求。

**沟通与协调**是指精准识别不同利益相关者的差异化诉求，确保组织机构内全部利益相关者都能及时了解相关政策、标准、流程、角色、职责和计划的最新情况，通过引入高层管理者等方式建立文档管理相关事宜跨部门的协商机制，提供多元的问题沟通和解决渠道，旨在提升跨部门和部门内部的文档管理能力，形成良好的文档管理文化。

**监督与评估**是指组织机构应当根据相关规范性文件构建符合业务需求的文档管理成熟度模型，建立具体的评价指标体系，定期开展成熟度等级评价。通过评估组织机构文档管理的政策、系统、流程和方法，发现存在的问题并解决问题，实现文档管理能力的持续提升。

上述各项治理活动都需充分考虑组织机构内外部的法律、社会、技术与文化环境对文档管理工作的影响。当环境发生变化时，文档治理活动需要进行适应性的灵活调整和变更。

## 8.3 管理活动层

管理活动层涉及两种类型的业务活动，一种是流程性的管理活动，包括生成、鉴定、捕获/归档、分类与索引、存储、使用和复用、长期保存、处置等。科学合理和高效有序的文档管理活动需要分类方案、归档范围、元数据方案、访问许可规则、保管期限与处置表等配套业务管理规则的支撑。

其中，**生成**是文档生命周期的第一个阶段，在此阶段中，文档通过创建或接收的方式得以积累，主要包括协作、创建、接收、流转、更改、分发、传输等。**鉴定**是指综合考虑法律、业务、社会、历史、文化等因素对文档的价值进行综合评估，判断其是否属于归档范围并划分保管期限的行为。**归档（系统中的捕获）**是指将业务活动中生成或接收到的数据对象作为文档及其元数据一起保存到文

档管理系统中的过程，包括为文档分配唯一标识符、捕获或生成关于文档的元数据、建立文件与其他文件以及业务活动之间的关联。**分类与索引**是指依据分类方案中逻辑上的结构化规定、方法和程序规则，对文档进行的系统标识并整理成类目，同时建立相应的索引元数据。**存储**是指确保文档能够不受未经授权的访问、更改、丢失或破坏（包括盗窃和灾难）的方式进行存储，确保文档的实体安全和信息安全。**使用和复用**是指用户对文档的查找、使用和检索以及基于文档内容创建新的文档并开展业务活动的行为。**长期保存**是指一系列对文档信息进行持续管理和维护的活动，包括更新、复制、迁移、仿真、封装等策略，以确保文档信息的长期有效性、真实可信，能够被未来的使用者理解和应用。**处置**是指按照鉴定得出的保管期限与处置表对保管到期的文档实施移交、销毁或续存的一系列过程。

其次是贯穿全流程的管理活动，包括安全合规、元数据管理和质量管理等，这些活动都是持续性的工作。**安全管理**是指文档管理需要遵从相关法律、法规，以及合规、隐私等方面的安全保护要求，构建完备的文档合规体系，从政策法规、数据资产、利益相关者和基础设施角度构建基于数据全生命周期的合规控制机制。**元数据管理**是有效开展文档管理工作、确保文档质量的关键。元数据包括文档元数据和文档管理元数据，涉及文档、业务、责任者、关系等不同类型的实体。管理的内容涉及元数据定义、元数据策略、元数据权限、元数据搜索、元数据视图和元数据分析等关键环节。**质量管理**需要从文档质量要求、文档质量策略、文档质量政策、文档质量标准、文档质量程序、文档质量实施规划等方面展开，围绕“计划-执行-检查-处理”的质量管理生命周期确保文档质量持续满足要求。

## 8.4 技术实现层

技术实现层包括采集技术、传输技术、交换技术、处理技术、存储技术、归档技术、转档索引、洞察技术以及面向文档“四性”的真实性保障技术、完整性保障技术、可用性保障技术和安全性保障技术。

其中，**采集技术**包括上传、在线创建、模板创建、一体机、同步、强制入盘、强制采集、页面抓取、表单采集、爬虫工具、Connector、Adapter、SDK、API 等；**传输技术**包括大文件切片、小文件聚合、直连存储、端点续传、去重技术、增量计算、block 文件块智能调整、限速策略等高并发快速传输技术；**交换技术**包括

共享、摆渡、数据外发等；**处理技术**需要将 AI 关键技术机器学习、深度学习、NLP（自然语言处理）与大数据技术深度融合，实现对文档的智能分类、智能标签、智能 OCR 识别、智能抽取和生成等；**存储技术**包括对象存储、分布式存储、分层存储、混合存储、策略存储等技术；**归档技术**包括基于 Web Service 的归档、基于中间库的归档、基于 ESB 的归档、基于 Hessian 协议的归档、离线归档等，涉及归档合规化、多级归档、存根归档、转非结构化归档等方面；**转档索引技术**包括格式转换与预览技术、大规模搜索引擎等技术；**洞察技术**通过集成人工智能、知识图谱和大数据等新一代信息技术，实现非结构化数据商业智能功能，包括统一搜索、智能搜索、智能推荐、知识图谱、数据挖掘、可视化呈现、数据洞察等。

**真实性保障技术**包括安全网络、用户身份识别/认证、权限控制、操作日志、数字签名、数字摘要、可信时间戳、区块链等；**完整性保障技术**包括定义归档范围、完整性检测技术等；**可读性保障技术**包括复制、更新、迁移、仿真、封装、载体维护、技术保存等；**安全性保障技术**包括容灾备份等存储安全技术，细颗粒度访问权限控制、身份认证、密级权限验证和安全域边界权限等使用安全技术，水印、基于智能 DLP（数据泄露防护）敏感检测的传输安全技术，基于大数据和知识图谱技术的审计安全技术等。

## 9 数据管理中文件档案与内容管理的发展趋势

### 9.1 结构化数据和非结构化数据融合发展

结构化数据和非结构化数据都是组织机构的重要业务资产和关键生产要素。其中，非结构化数据已经成为数据治理体系中不可忽视的组成部分，在组织机构运转过程中发挥着不可替代的重要作用，亟待进行资产化管理。同时结构化数据和非结构化数据之间存在相互转化和互为补充的关系。一方面，结构化数据可以通过输出版式文档的方式，转化为非结构化数据，从而满足合规管理和长期保存的需求。另一方面，非结构化数据可以通过 OCR 识别、单据识别、物体识别、基于语义内容的智能标签/智能分类、实体抽取、业务元数据等手段和方式转化为结构化数据，从而满足业务和应用的需求。其中，实体关联图谱的有效抽取和抽象主题图谱的构建与持续完善是未来文件档案内容管理的重要发展方向之一。组

织机构设置的数据治理委员会、数据管理部门等的关注重点将从结构化数据逐渐扩展至非结构化数据及其与结构化数据的融合，最终实现数据之间的关联、展示、分析和洞察，全面发挥数据的价值。

## 9.2 文档管理和内容管理融合发展

文件档案管理是内容管理软件的重要功能，文档管理行业和内容管理行业密不可分，相互支撑、互为依靠，逐渐形成联动发展、互利共赢的良好生态。我国文档管理行业和内容管理行业之间的联系愈加紧密，文档管理行业逐渐成为内容管理行业的重要支撑。两者融合发展主要表现在以下三个方面：

一是目标和理念的融合。以往的文档管理侧重安全合规、业务连续、权益保障、记忆留存等，内容管理侧重支持决策、提效降本、创新发展、资产维护等，现在两者都被纳入数据管理体系中，将互相利用对方的优势，共同支撑组织机构的运作以及业务活动的高效有序开展。

二是组织和职能的融合。文档管理部门不再是居于后端的保管者，而是重新思考自身的职能和定位，一方面，参与机构整体的数据治理工作；另一方面，成为非结构化数据治理的责任部门，文档工作者成为“内容管事”，承担非结构化数据的统筹规划、全生命周期管控、数据质量保障、安全合规等职责。

三是技能和工具的融合。一方面，内容管理行业需要来自文档管理行业的支撑以及与之配套的管理理念和管理规则；另一方面，文档管理行业也需要与其需求相匹配的系统、平台、技术、工具，内容管理行业能够为其提供更为丰富多样的产品及信息化的集成化解决方案。

## 9.3 文档管理和数字化业务融合发展

首先，从文档管理与业务活动的关联程度来看，国际标准化组织发布的文档管理核心标准（如 ISO 15489、23081、21946、26122 等）确立了面向组织机构业务需求的以职能活动和业务流程分析为核心的方法体系，两者的融合程度逐步提高。组织机构实施文档管理的目标不再局限于文档本身的规范化和有序化，而是要更多地支撑、反馈和驱动业务活动。一方面，文档管理“从业务中来”，即业务

数据化。通过分析业务可形成数据资产清单，准确识别文档及其元数据，实现数据资产的实时登记和动态监测，将文档管理需求嵌入业务流程，实施嵌入式档案化管理，满足合规管理要求。另一方面，文档管理要“回业务中去”，即数据业务化。基于业务流程分析，可精准识别各项业务的数据需求，确保合适的人能够在合适的时间以合适的方式获取合适的文档，实现数据赋能业务。

其次，从数字转型的角度来看，文档管理是实现组织机构数字转型的抓手和推力，文档管理产品是优化资源配置、促进高效协同的重要工具。近年来，美国、英国、澳大利亚、新西兰、加拿大等国的文档管理主管机构相继出台政策法规，实施战略部署，推动文档管理的数字转型。我国相关主管部门也在大力推进电子文件的单套归档和单轨管理。新修订的档案法明确规定“电子档案与传统载体档案具有同等效力，可以以电子形式作为凭证使用”。文档管理数字转型意味着全流程的电子化和数字连续性的实现，是打通组织机构数字化“最后一公里”的关键。例如，就数字政府建设而言，电子证照是实现在线政务服务的关键；就数字经济发展而言，推行电子发票、规范电子会计凭证报销入账归档是促进企业运作无纸化、实现企业高效运作的重要举措；对智慧法院而言，电子卷宗随案生成和归档是实现案件全流程在线办理的重要环节。

## 9.4 文档管理基础建设和智能开发融合发展

文档和内容管理呈现出从简单的文档集中、协作办公、安全交换走向基于人工智能的智能开发的趋势，从数字化、数据化走向智能化，但需要指出的是，各阶段之间并不是直接替代的关系，而可能长期并存。未来的文档管理工作应该做到“两手抓”，一方面是文档管理基础建设，包括文档协作、文件归档、安全管理等方面。近年来，基于文档协作和共享的知识高效协同和文档全生命周期安全防控体系的市场需求呈上升趋势，成为文档管理的重要基础工作。另一方面是基于文档的深度挖掘和分析，呈现出智能化乃至智慧化的趋势。文档管理应是语义层面的管理，基于人工智能，尤其是深度学习和知识图谱，实现洞察分析，同时融合应用云计算、大数据、区块链、5G等其他信息技术，共同赋能文档管理工作。具体来说，文档管理的智能应用场景将重点包括智能推荐、图谱探索、智能搜索、

安全分析。总之，前面的文档管理基础建设是实现智能开发的前提和保障，后面的智能开发技术也可以反向推动基础工作的完善，实现文档管理基础建设和智能开发的融合发展。