



电子查询(eDiscovery)和企业数据管理 — 您需要了解的技术



目录

- 2 电子查询前景大好**
- 3 公司的诉讼费用高达数百万美元?**
- 7 哪些数据管理挑战会影响电子查询?**
- 8 为何 IT 的早期参与必不可少?**
- 9 企业数据管理如何支持电子查询计划?**
- 14 企业数据管理有哪些益处?**
- 15 对企业数据管理解决方案应该抱有期望?**
- 18 IBM Optim 支持您的电子查询计划**

电子查询(eDiscovery)前景大好

新的产品，新的技术，新的法规，以及新的负债！随着法律（如萨班斯-奥克斯利法案、SEC-17a 和 Basel II 等）对数据保持和责任的规定更加严格，各行业的公司都面对着更大的审计和潜在诉讼风险。不仅风险增大，相关的成本也可能增加到上百万美元。

要满足审计需求、防止诉讼或者至少避免高昂罚金，一个关键因素就是要能够检索适当的信息和支持证据来满足审计、法律和业务需求。然而，随着“新式数据中心”和“无纸化”应用环境的兴起，几乎有 95% 的业务信息是通过电子方式创建和保存的，因此电子查询得到了发展。

Forrester Research 称，电子查询就是收集、准备、检查、格式化和生成纸面或数字格式的企业内容以响应内部、法律或规章的要求的过程。¹ 内容包括文件、数据库、电子邮件和任意其他能够在诉讼中充当相关证据的电子存储信息（Electronically Stored Information, ESI）。

尽管公司知道通过电子方式管理业务关键信息有很多好处，但真正实施起来还是面临着很多 IT 难题。IT 企业需要经得起考验的技术来管理和定位特定信息，从而快速响应任何电子查询或审计请求。如果不具备有效的企业数据管理策略和 IT 的早期参与，无法解决数据保持和审计的难题，那么公司对电子查询请求的响应方式就成了成功和失败之间的分水岭。

公司的诉讼费用高达数百万美元？

Fulbright & Jaworski L. L. P 国际律师事务所进行的第三次年度诉讼趋势调查提供了一些惊人的统计信息。根据来自全世界 422 家专业律师事务所的资料，调查发现“美国公司平均面临着 305 例未决的国际诉讼。对于大型的美国公司，也就是年收入总额在十亿美元以上的美国公司而言，诉讼的数量上升到 566 例，而在这些大公司当中有接近一半每年还会出现新的争端。”²

没有行业能够免于潜在诉讼的风险。调查研究了 13 个“不同的行业，包括能源、制造、金融服务、零售/批发、技术/通信、工程/建筑、卫生保健与制药、房地产、保险和教育，以及非盈利性组织和贸易协会。”³而保险行业中平均每家公司有 1,696 例诉讼，无可争辩地排在榜首，紧随其后的是零售和能源公司（超过 330 例）。

另一项重要的统计数据是成本。调查报告称“大型美国公司的诉讼费用平均为 19,000,000 美元”，而“美国小公司的律师称他们用于解决争端的平均费用总计只有 178,000 美元”。另外，“尽管这些案例的绝大部分发生在美国法庭中，国际争端的潮流也在愈演愈烈——超过三分之一的公司说，他们的诉讼事件中有 20% 发生在外国，这证明美国式的诉讼正在变得全球化。”⁴

为电子查询做准备的业务理由。显然，发生诉讼的可能性正在变大；不仅是在美国，在全球范围内亦是如此。相关的法律成本和处罚可能十分高昂。另外，因为大部分案例都在开庭审讯之前得到了解决，所以发现阶段通常是诉讼中最为关键和昂贵的部分。

例如，在出现电子查询请求时，将信息汇总到一起需要耗费时间和资源。首先，必须找出正确的数据，这些数据可能是在企业内不同平台上运行的不同应用数据库中进行管理的。这些数据的维护方式可能会让它们更易于访问。其次，必须保持数据和元数据的完整性，这样才能在业务上下文中准确检索到未曾改动过的数据。

接下来，检查数据以确保将必须产生的信息与拒绝公开的数据或不相关的数据抑或必须阻止的数据分离开。最后，必须在可传递的介质中生成数据。以每小时或每 GB 数据的成本为基数，很容易计算出发现阶段的费用是多少美元。

以下指标出自一份最近的 Gartner 研究报告（参见表 1），用于指示成本增加的速度：⁵

表 1. 电子查询(eDiscovery)的成本指标

任务	时间
找出正确的数据	200 美元/ 小时
保持数据	100 到 300 美元/ 小时
收集数据	200 到 300 美元/ 小时
检查数据	120 到 350 美元/ 小时
生成数据	1000 到 2,100 美元/ GB

可是，即使您的企业从未收到过审计或电子发现请求，做好这方面的准备对于业务也有着积极意义。很多公司开始检查他们的 IT 基础设施系统和应用，以便找出可能存在争论的领域。除了电子邮件之外，还要考虑支持定制与打包应用的数据库、文件服务器、网络访问控件、桌面工作站、笔记本电脑、黑莓手机和企业内的其他移动设备。

为电子查询做准备的法律理由。为电子查询做好准备有几条法律理由。第一条也是最重要的一条是 2006 年 12 月生效的联邦民事诉讼规则(Federal Rules of Civil Procedure,FRCP)。这一新规则负责管理在联邦法庭备案的诉讼案件，并规定了所有企业的公共或私人的权利和义务。

基本上，这些规则会影响公司在记录保持和信息管理计划上的投资一些值得注意的修正有：⁶

- 规则 26(b) 定义了分层存储环境中电子信息的可发现性和可访问性。这条规则还触及到了被遮蔽(redacted)文件中信息的私密性和对发现过程中生成的拒绝公开信息的无意使用。遮蔽(redaction)是指将信息定为“拒绝公开”的过程。公司必须证明其信息系统和过程的完整性，并表明记录与信息保持策略是一致地实施的。
- 规则 26(f) 引导各方开发一个发现计划，并解决与保持可发现信息相关的问题。各方必须就产生所需格式的数据的方法达成一致。这条规则迫使企业找出要保留信息时，备份与归档技术之间的战略差异。

- 规则 34(b) 将电子存储信息划作单独的一类，包括电子邮件、视频、音频等。至于比如应用数据等以电子方式生成和管理的信息，元数据通常包含在发现响应中。元数据用于“描述”数据，对于理解信息的业务上下文是必不可少的，必须加以保留。
- 规则 37(f) 提供了“安全港”，以回避日常业务活动和信息管理系统的例行操作中不可恢复（损坏）或丢失的数据所引起的惩罚。然而，应该注意保持相关信息。

FRCP 规则建议公司实施过程用以捕捉和保留与现行或未决诉讼直接相关的信息。然而要提出警告的是，如果保留的数据比所需的多，多出的数据也可用于电子查询。⁷另外，现行或未决的诉讼允许“诉讼保留”。这意味着，公司必须对过程进行监控，以确保相关数据不会被修改或损毁。有效的数据保持策略对于法规遵从性至关重要。

FRCP 规则责令公司负责采取“合理步骤”以落实过程和技术来保护数据。为了对电子查询做好充分准备，公司必须准备好对业务数据进行评估和分类的适当过程，并建立正确的数据保持、备份、恢复和诉讼保留策略。好消息是，已经有技术可用于实施有效的保持策略和支持电子查询计划，但仍存在着大量挑战。

哪些数据管理挑战会影响电子查询？

Gartner 最近的一份研究报告指出，“由于法规遵从性和发现要求的提高，管理和保持历史信息的复杂需求不断增多，企业正在经历这一痛苦阶段。电子保存内容的数量继续疯狂增长，需要管理的内容类型也越来越多……企业将需要借助信息存档归档技术来实现他们的内容管理与保持策略，从而有效地兑现他们的承诺。”⁸

报告还指出，“很少有企业拥有同时针对存在于面向事务的应用中的结构化文件和非结构化内容的存档策略。活动数据存储通常能够增长，定义的内容生命周期管理方案中不会包含存档，这对于支持业务连续性以及这些关键应用的备份和恢复的 IT 能力提出了很高的要求。” Gartner 又指出，“一般保存在数据库中的结构化数据将需要特殊的数据库存档解决方案。”⁹原因何在？因为驱动业务运作的 ERP、CRM 和定制应用收集了大量数据，必须对这些数据进行管理、维护和保护。应用数据或“结构化数据”对电子查询提出了更多挑战。当数据保存在一个或多个复杂的关系数据库中时，这种挑战更加严峻。

例如，如果引用两个不同数据库中的数据，那么这些数据必须一起存档，同时还要保持数据关系。合格的数据库存档解决方案必须管理复杂的关系数据，并保持数据的参考完整性。如果数据模型很直观，那么创建具有参考完整性的数据子集就会相对简单。然而，当今企业应用使用的数据库多数都依赖于复杂的数据模型。

为了支持电子查询要求，理想的数据库存档解决方案必须捕捉业务对象细节层面上的数据，并保持数据的参考完整性。例如，完整的“工资单”档案要包括时间和劳动力事务细节以及员工的主数据，这样便可不依赖生产数据就满足报表编制要求。

为何IT的早期参与必不可少？

IT、业务用户、遵从性专家和法律顾问之间的交流是必要的。组建跨功能领域的团队来定义数据管理战略，可以帮助确定何时存档哪些数据、应该存储在什么位置、谁可以访问数据以及应该保留的期限。落实企业数据管理战略时，这个团队将得到发展而管理企业内的结构化和非结构化数据，这样不仅可以支持电子查询和审计的要求，而且还可以改善日常运营。

记录保持程序必须清晰定义、良好执行并统一实施，以防止有意篡改，并遵从 **FRCP** 对于电子查询的要求。IT 人员实行的数据库存档和存储战略需要满足可访问性要求，以符合遵从性同时降低成本。

“IT 专业人士扮演着第一响应者和把关者的双重角色，保证了正确协议的就绪和实施。跟上电子查询相关法律和技术变化的脚步，无疑可以帮助公司在遇到诉讼时保持策略上的优势。”¹⁰

企业数据管理如何支持电子查询计划？

有效的企业数据管理战略允许您管理应用数据的整个生命周期。借助全面的数据库存档功能，可以主动分析企业内的不同数据类型，并且定义每个数据集在其生命周期各个时间点的价值。通过根据业务规则而自动化存档过程，可以将当前活动的数据与历史事务隔离开来，并且更好地控制基于数据业务价值和访问要求而实现保持策略的过程。

理想情况下，已存档的报告和参考数据必须保持在它们原有的业务上下文中。您还需要能够在各种存储介质上以不变格式管理存档，从而保证遵从性和轻松的按需应变访问。企业数据管理存档功能提供了很多选项，可以帮助您挖掘出存储器资源的最大价值。

通过基于其业务价值来存档历史数据，以及实现分层存储策略，您可以优化存储器利用率并推迟高速资源的购买，同时针对电子查询和其他报告要求提供授权的数据访问。已存档数据无论位置如何，始终可以访问。

具有存档功能的企业数据管理战略可以改进管理应用数据和支持电子查询请求的方式。下面说明了此种战略的一些优势。

管理数据增长。业务应用将持续收集收据。然而，如果放任当前数据和历史数据不断累计，巨大的数据量会损害性能、限制可用性和增加成本。作为公认的最佳实践，例行数据库存档能够将应用数据库维持在一个可管理的规模，可以降低容量要求和运营成本。

隔离历史数据与当前数据的功能让公司能够管理持续增长，同时保持当前数据的可访问性，以满足报告、审计或电子查询的要求。可以为管理不同类型的历史数据定义保持战略，从而满足遵从性和电子查询的要求。

当前的应用数据依然保留在良好管理的生产环境中，可发挥最高性能。历史数据可以存档、编制索引、压缩和保存到各种存储介质上。如果产生发现请求，就可以快速定位和检索到特定的存档数据，从而控制成本。

访问和退役数据。通常，电子查询请求对所需信息的交付有着严格的期限要求。公司需要知道存档数据所在的位置，以及谁有权访问这些数据。快速索引、定位、访问和检索数据的功能确保能够及时响应，并避免无谓的成本和罚金。

给档案编制索引可以提高搜索的准确性，这对于找出所需数据很重要。用于监控谁有权访问和何时最后一次访问存档数据的其他审计功能可以提供更高水平的控制。企业数据管理选项支持对存档数据的授权访问。除了本地访问和独立于应用的访问之外，还有一种开放式数据管理访问功能，在确保快速检索数据的同时还减轻了风险。

保护数据隐私。公司必须保护所有敏感信息，无论这些信息位于生产系统、开发数据库还是企业内任何别的地方。遮蔽，或者说阻挡信息的过程，是电子查询的重要一环。公司需要此项功能来屏蔽电子查询请求中不必要的敏感信息。

行业分析家们认识到，数据去标识化提供了另一种保护数据隐私的途径。通过实现支持屏蔽或转换敏感信息的功能，您已经获得了所需的数据管理战略来保护数据隐私和保持数据完整性。

定义数据保持策略。根据 Forrester Research 的说法，“如今的保持管理有一个迫切的需要——更低的电子查询成本和更高效的遵从性过程。没有准备相应计划的企业现在不仅丧失了眼前的业务优势，而且还失去了从总体上更好地管理信息的机会。”¹¹

首先必须确定使用哪个应用来管理电子查询最可能需要的数据，并为应用分类，以便根据业务要求进行存档。接下来，由于对数据进行了分类，定义和记录用于管理活动数据、不活动数据和依照遵从性管理的数据的业务规则和数据保持策略变得更加轻松。

如果历史数据不再需要用于遵从性或业务目的，那么删除它们可以防止信息资产变成信息负债。例如，银行对账单、已取消的支票和会计记录必须保留 7 年。如果您的公司面临着一个支持破产处理审计或电子查询请求，而这些记录在保持期限到期后仍未删除，那么可以要求将它们作为证据。要提出警告的是，如果这些记录确实适于电子查询请求，此时需要能够申请诉讼保留来保护这些记录免于“损毁”。

允许申请诉讼保留或者为数据的保持和删除而定义自动化策略，这种功能提供了额外的保护。您可以确保在面临审计或电子查询请求时，如有必要则保留数据，或者在再也无需保留时删除数据。

区分备份和档案。数据库档案和数据库备份都是公司日常运营必不可少的最佳实践的重要组成部分。备份使恢复数据成为可能，从而确保业务的连续性。相比之下，存档是为了长期保持数据，对于电子查询特别重要。然而，二者在企业数据管理中扮演着同等重要的不同角色。

一致的备份处理保证了在发生灾难或由于疏忽导致数据丢失的情况下，可以从最后一次的备份点快速恢复操作数据。成功的恢复取决于执行备份过程的频率和效果。然而，因为备份通常会被覆盖，它无法长期保持数据，因此也不能支持电子查询请求。

相比之下，持续进行的数据库存档允许您出于报告和遵从性目的而管理和维护历史数据及其相关的元数据。始终可以按照需要访问已归档数据，以支持电子查询工作。实现例行数据库存档功能可以显著缩短备份与恢复的时间，这是另外一个益处。将数据库的规模控制在易于管理的范围内，可以将性能和可用性维持在可以接受的水平。

保存数据直到其生命周期结束。 公司需要各种高性价比的存储方案和功能，以便通过统一格式管理数据，从而满足电子查询的要求。分层的存储策略支持根据数据的业务价值和访问要求来管理数据，直到保持期限到期为止。

例如，将活动数据保留在高性能的存储环境中以实现快速的事务吞吐，这很有意义。报告数据能够被迁移到中层存储器中，因此您可以控制成本，同时仍能满足服务需求。

最后，能够将参考数据以不变格式存储在安全的 WORM（一次写入，多次读取）设备中，并在该 WORM 设备中使用已定义的策略管理对这些数据的访问和最终处置。企业数据管理的灵活性允许 IT 组织利用性价比最高的平台来保存企业数据，直到其生命周期结束。

退役应用。定期检查应用产品组合，有助于在操作和维护应用的成本与应用提供的业务价值之前取得平衡。您的目标可能是消除冗余、将应用迁移到性价比更高的平台上或者降低对于厂商不再支持的系统的依赖性，从而简化 IT 基础设施。

包含数据库存档功能的企业数据管理策略保证了对应用数据的访问，即便是在原始应用停止服务之后。站点可以使用行业标准方法来查询和报告已存档的业务记录，从而满足电子查询的要求。

企业数据管理有哪些益处？

最终，贯彻包含数据库存档功能的企业数据管理战略，可以提供切实的益处来支持审计和电子查询的要求。

- 通过存档以隔离当前数据与历史数据，并按照 FRCP 的要求以统一格式保存历史的应用数据、管理持续的数据增长。另外，将数据库的规模控制在易于管理的范围内，可以提高服务水平性能、增强可用性并降低操作成本。

- 定义和实施自动的数据保持、诉讼保留和处理策略，以保证保留正确的信息在必要时使用，并且在不再需要这些信息来满足遵从性要求时删除它们。
- 使用索引编制技术和各种访问方法（比如本地访问或独立于应用的访问），以保证能够按需应变地访问当前数据和存档数据。
- 隔离电子查询请求不需要的敏感数据。应用各种数据屏蔽技术，以防止无意泄露拒绝公开的信息，同时支持遮蔽的要求。
- 实现分层存储战略，以便根据业务价值和访问需求来管理历史数据。从而促进在数据的整个生命周期内经济有效地管理数据。
- 让无法再交付业务价值的应用退役。在应用停止服务很久之后，依然可以访问存档的应用数据。

理想的解决方案应该要解决可支持审计和电子查询要求的关键问题。

对企业数据管理解决方案应该抱有期望？

理想的企业数据管理解决方案应该提供一致的技术和方法，并且可以将这些技术和方法推广到各种应用、数据库、操作系统和硬件平台上。此解决方案提供的久经考验的功能可以解决由电子查询和审计要求而引发的关键 IT 问题。

存档和检索完整的业务对象和元数据。能够确保所有数据的完整性的存档功能是重点。理想的企业数据管理解决方案必须能够捕捉完整业务对象及相关元数据。还应该能够有选择性地检索存档的事务记录同时不必进行还原，并且能够将存档数据恢复到生产环境或其他目标环境中。

整合的存档和索引编制功能可以加快搜索存档的速度。理想的解决方案应该支持为每份存档建立一到多个索引，从而减少准确检索到所需数据的时间。

提供各种兼容的存储方案。确定将档案保存在数据库还是文件中，又或许在不同的时间点使用不同格式来保存。这些选择各有千秋。例如，当您刚开始一个存档项目时，您的员工很可能可以轻松地管理档案数据库。但随着时间流逝，档案数据库可能会膨胀到初始生产环境的大小，迫使您也要存档这个数据库。所有可能的存储费用节省都将失去，而项目管理的工作也很快变得极为沉重。

相比之下，压缩文件格式所占空间较小，因此可以最大程度地节省存储空间。还可以为压缩文件编制索引，从而实现已存档数据的快速检索。最重要的是，以文件格式管理存档，就可在*档案的生命期*内可以提供最广泛的访问方法。

提供长期数据访问选项。访问存档数据的典型方法包括本地（基于应用的）访问和独立于应用的访问。有些存档解决方案只提供其中一种访问方法，但极少数同时提供两种。您需要衡量每种方法的优势，从而满足您特定的访问需要。例如，本地访问支持通过原始的应用接口与存档数据交互。这种方法允许用户使用熟悉的格式与过程访问他们所需的信息。然而，如果计划退役一个应用，但仍然想要保留对存档的访问，那么独立于应用的访问方法可能是更好的选择。

在档案的有效期内，独立于应用的访问方法是最灵活的。这种方法使授权用户能够使用行业标准方法和报表编写器与档案交互。您可以使用如 ODBC/JDBC、XML 和 SQL 等各种行业标准的访问方法，以及如 Business Objects™、Cognos® 甚至是 Microsoft® Excel 等报表工具。利用这种访问途径，您的档案就绝不会过时，因为在原始应用经过升级、修改甚至退役很久后都仍然可以访问它。

满足数据保持和诉讼保留要求。接下来，在数据的整个有效期内经济有效且一致地管理存档和保持期限的功能是重点。在各种存储介质上保存档案的功能可让基于业务价值和访问要求的数据管理方法永不过时。解决方案应该提供诉讼保留功能。这样可以保持业务记录的可访问性，直到合法的保持期限到期为止，之后可以手动删除或者自动处理到期的事务记录。

IBM Optim 支持您的电子查询计划

为电子查询做准备对于业务有着积极的意义。企业必须认识到，对其信息和企业应用数据的组织具有业务价值。为了电子查询而管理企业应用数据面临着很多 IT 挑战。为了遵守 FRCP 规则，公司必须采取一些“合理步骤”来实现支持电子查询要求的过程和技术。

IBM® Optim™ 数据增长解决方案提供了企业数据管理和经过验证的数据库存档功能，可以帮助您满足审计和电子查询的要求：

- 存档和检索完整的业务对象与元数据
- 提供各种兼容的存储方案
- 提供长期数据访问选项
- 满足数据保持和诉讼保留要求

包含数据库存档功能的企业数据管理战略提供了有效方法来管理、保存、访问和检索数据，从而支持电子查询的要求。作为公认的最佳实践，数据库存档还能提高应用的性能和可用性，同时降低风险和运营成本。

关于 IBM Optim

IBM® Optim™ 企业数据管理解决方案关注关键的业务问题，比如数据增长管理、数据隐私遵从性、测试数据管理、电子查询、应用升级、迁移和退役。

Optim 使应用数据管理适应业务目标，以优化性能、降低风险并控制成本，同时交付在不同的企业应用、数据库和平台中伸缩的能力。如今，Optim 帮助全球行业范围内的企业利用在其生命周期的每个阶段管理企业应用数据的能力，发挥其企业应用和数据库的商业价值。

更多信息

要了解 IBM Optim 企业数据管理解决方案的更多信息，请联系 IBM 销售代表，或者访问：<http://www-01.ibm.com/software/cn/data/data-management/optim-solutions/>



© 版权所有 IBM Corporation 2008

IBM Software Group
111 Campus Drive
Princeton, NJ
USA,
08540-6400
800.457.7060
609.627.5500
传真 609.627.7799
www.optimsolution.com

在美国印刷

2008 年 3 月
保留所有权利。

¹ Barry Murphy 和 Robert Markham, “eDiscovery Bursts Onto the Scene,” Forrester Research, 2006 年 3 月 1 日。

² “Fulbright Launches its Third Annual Litigation Trends Findings – U.S.,” Press Release, 2006 年 10 月 10 日。

³ Ibid.

⁴ Ibid.

⁵ Debra Logan, “Mapping Technology and Vendors to the Electronic Discovery Reference Model,” Gartner Research, ID Number: G00153110, 2007 年 1 月 9 日。

⁶ Vivian Tero, “Worldwide Legal Discovery and Litigation Support Infrastructure 2007 – 2011 Forecast: Legal Matter and Compliance Records Repositories Are the Early use Cases of the Active Archiving Architecture,” IDC Market Analysis, IDC#207155, 2007 年 6 月, 28-29 页。

⁷ Alan Brill and Michele C.S. Lange, “The New E-discovery Rules: What You Don’t Know Can Definitely Hurt You and Your Company,” Computerworld – Storage, 2006 年 12 月 11 日。

⁸ Kenneth Chin and Carolyn DiCenzo, “Key Issues for Information Archiving and Retention,” 2007 年, Gartner Research, ID 编号: G00146211, 2007 年 2 月 9 日, 第 2 页。

⁹ Ibid. 第 3 页。

¹⁰ Alan Brill 和 Michele C.S. Lange, “The New E-discovery Rules: What You Don’t Know Can Definitely Hurt You and Your Company,” Computerworld – Storage, 2006 年 12 月 11 日, 第 2 页。

¹¹ Barry Murphy, “Abysmal: The State Of Retention Management,” Forrester Research, 2007 年 7 月 17 日, 第 10 页。

IBM、IBM 徽标和 Optim 是国际商业机器公司在美国和/或其他国家/地区的商标或注册商标。所有其他公司或产品的名称都是其各自所有者的商标或注册商标。

本出版物中对 IBM 产品、计划或服务的引用, 不代表它们可用于所有 IBM 运营或从事业务的国家/地区。

客户有责任确保遵守法律要求。客户自己全权负责向合格律师请教哪些法律与他们的业务相关, 并请求律师对这些法律条款进行解释, 客户还应全权负责采取适当行动来满足此类法律的要求。IBM 不提供法律意见, 也不对 IBM 服务和产品能够确保客户遵从此类法律提供任何陈述或保证。

***TAKE BACK CONTROL WITH* Information Management**