



通过企业捕获 获得业务授权

捕获作为一个标准的，可共享的企业服务提供了引人注目的商业效益



目录

执行摘要.....	3
什么是企业捕获.....	1
企业捕获的价值.....	2
企业捕获的演变.....	2
企业捕获的关键因素.....	4
企业捕获的今天.....	6
IBM进入市场.....	6
企业捕获平台.....	7
结束语.....	8

执行摘要

在如今的经济形势下，每一个企业面临的挑战是提高业务流程周转率，同时尽可能地使用最有效率的操作方式。对IT部门来说这意味着有助于增加业务收益，增加现金流，降低成本和提升客户服务.IT也呼吁通过整合基础设施和标准化服务等计划来简化它的操作。

企业捕获使得输入可以理解业务流程的非结构化信息，它是一个新兴的解决方案，可以同时解决这两项挑战。据估计，每天有15千兆字节的新信息产生，很大一部分是起源于书面形式，大约80%的此类信息增长是非结构化数据。企业捕获可以帮助组织利用此类信息实现更好的业务成果和进行更有效的运营。

企业捕获加快了抵达纸张、邮箱、传真，PDF和HML等媒介上的非结构化信息的处理过程。通过对保险索赔、发票、贷款申请、纳税申报表、货运清单、销售订单和其他事务性文件中的业务关键型信息进行自动分组、理解和提取，企业捕获能以电子的速度将数据数据到后端系统。这远比劳动密集型、高成本和易出错的纸质文件管理要更优秀。并且通过将多个捕获方案整合为一个企业标准，企业捕获可以降低管理基础设施和供应商的成本。另外，捕获可以作为企业内容管理（ECM）和业务流程管理（BPM）中不可或缺的战略。

近几个月来，企业捕获供应商的局面因为IBM的Datacap收购而发生了改变，Datacap 是一个独立软件供应商，它的Taskmaster Capture 解决方案已经瞄准了企业捕获领域的领导地位。通过将本产品添加到其内容管理解决方案范围中，IBM把自己定位为一个企业捕获和ECM的单一供应商提供者。

本白皮书探索了企业捕获的潜能，并解释了一个典型系统的组成部分。还介绍了IBM、Datacap、Taskmaster Capture背后高管的观点，他们就一个成功的企业捕获解决方案所需的架构、功能性和集成功能提供了自己的愿景。

什么是企业捕获?

企业捕获使用通用软件在企业范围内从文档中捕获和提取信息，要么是面向位于多个位置的单一企业功能，要么是面向整个企业内的多个企业功能。该功能的示例包括在多个分支机构/国家实现面向应付帐款/发票处理的单一捕获解决方案，或是一个获取和处理文件、电子邮件附件和传真服务索赔处理、人力资源、应付帐款和订单/应收帐款的通用解决方案。

企业捕获需要一个在其进入企业时可以捕获文档的智能流程。即.文档分类、数据识别和提取，以及路由实现自动化。通常，企业捕获通过集成案例管理、ERP、CRM、BI或其他应用程序，将所提取的信息提供给业务流程，同时为ECM存储建立索引和准备文档图像。

企业捕获的价值

企业捕获的设想是使捕获技术作为一个标准、共享额服务，替换通常分布在企业中的多个捕获和文档管理孤岛，从而可供整个企业使用。这就可以提供战略和战术两方面的商业效益，包括以下内容：

- 用更快的自动路由和捕获信息的处理过程来提高业务速度。效益可以在任何业务流程中产生，具体依赖于非结构化文档中的数据。
- 当他们进入企业时，通过捕获文档简化管理纸质文件的“截断”流程操作。
- 在即时分类、索引、添加元数据和跟踪文档方面实现更好的合规性。
- 降低在消除将纸质文档传输到集中中心或捕获服务部门过程中的安全风险。
- 在减少数据录入需求、文件传送和文件保存上节省成本。
- 更有效的IT管理，降低在巩固基础设施和供应商管理上的成本。

通过使纸质的和其他非结构化文档能够以数字化速度进行处理，企业捕获可以履行无纸化办公的承诺。

企业捕获的演变

企业捕获已经逐渐变成通过多种方式从企业文档中获取信息。从输入纸上数据到部门中的专门按需扫描，再到现场的基于Web的扫描、高速的批量扫描和表格处理，企业捕获可以利用扫描仪、图像分类和数据提取的识别软件，以及企业级解决方案架构和处理能力中的改进部分。

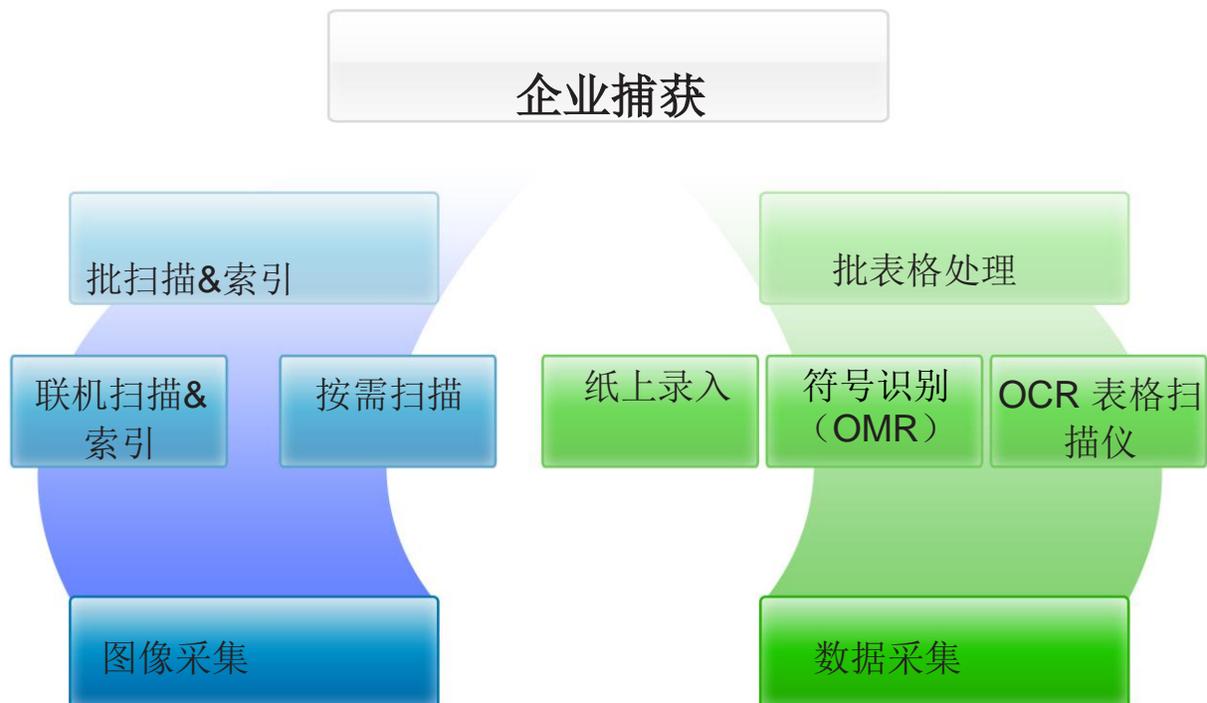


图1-企业捕获汇集了2个企业仍然在他们的文档战略中使用的长期过程。图像采集源自于微缩胶片行业，文档在该行业中可以成像和编成索引以供检索，而数据采集是表格处理行业的产物，数据在该行业中为了交易的目的从文档中提取出来。

从历史的角度来说，企业捕获开始于欧洲的邮件收发室。开发邮件收发室软件的目的是准备和扫描收到的邮件然后进行电子分发。这就为企业捕获搭建了舞台，因为当他们进入企业时会需要扫描和分配多个类型的文档。在美国，美国邮政服务的提供电子送货服务已经被法律阻止，但是关键事务型邮件如纳税申报表、表格和零售付款已经通过邮政信箱进行了预先分拣。然后才进行扫描和处理。

企业也可以在邮件进入组织时使用邮件收发室软件进行扫描。捕获软件使用自动分类来决定路由途径。在欧洲和其他地方，这种方式已经扩展到可以帮助执行客户服务和其他业务流程，已经成为企业捕获的一个关键要求。

企业软件在捕获技术上的进步与市场趋势十分吻合。今天的企业从整个企业的标准化服务上实现商业价值。通过整合IT平台、供应商和流程，他们可以降低成本并简化管理。通过实现企业范围可用的捕获服务，他们可以培养创新的解决方案。一个或者多个部门实现无纸化办公的商业价值时，其他部门就会效仿。而且一个管理非结构化内容的无纸化解决方案可以有助于合规性方面的工作。

企业捕获的关键元素

在进入点捕获文档

企业逐渐认识到了将文档转换为图像的价值，而不用考虑它们位于地方以及它们进入公司的方式。因此，企业捕获必须在它们的进入点支持通过多种方式获取文档。

邮件收发室

部门

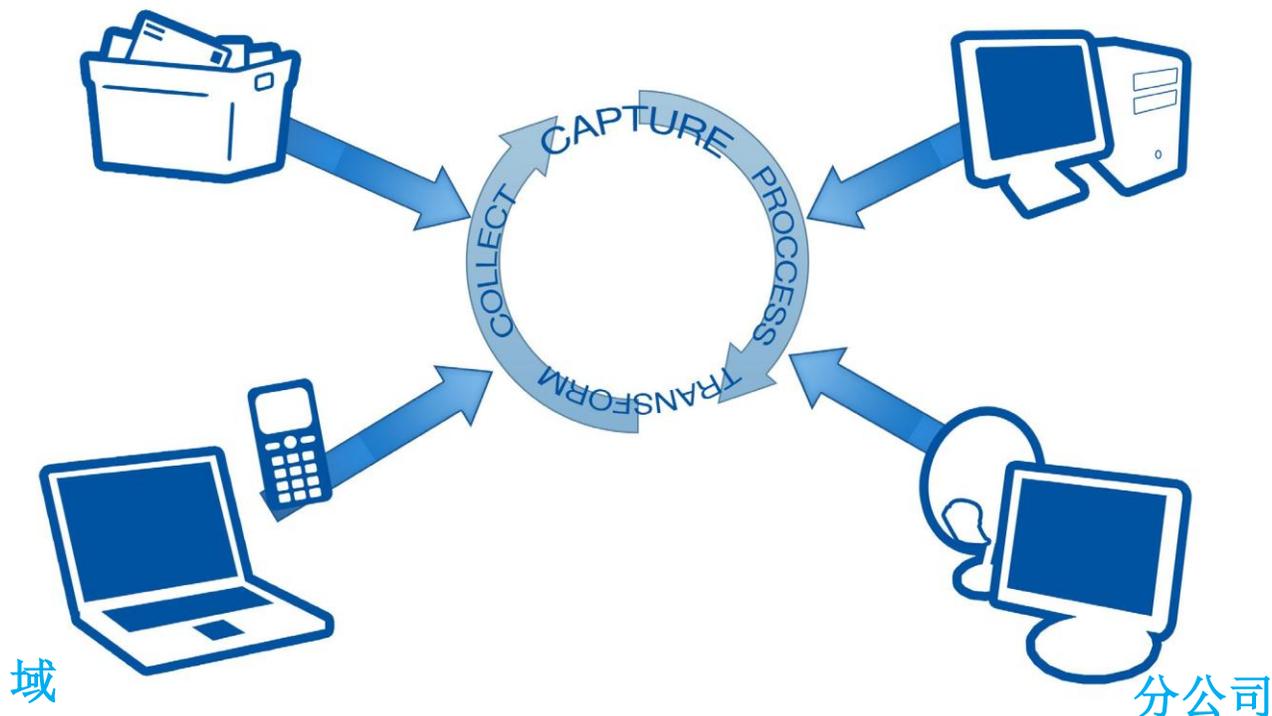


图2—当文件进入到组织，企业捕获软件必须可以从任意位置进行图像和数据的采集，作为一个捕获门户将索引图像和精确数据汇集到组织使用的所有目标系统，从ECM和ERP到BI和分析。

当使用高速集中式扫描仪或小型邮件收发扫描仪打开文档并进行扫描后，中央捕获就需要将文档交付到一个邮件分支、总部或邮政信箱。分支捕获能够使用传统扫描仪批量扫描接收到的文档，或根据需要使用网络扫描仪或MFP进行扫描。有代理商或附属公司参与的地方，就会出现扫描操作。

部门捕获通过内部邮件处理接收到的文件，或在具有共享网络扫描仪或MFP的部门里作为按需扫描结果而打印的文件。部门捕获包括传真服务器，依赖基于图像的处理过程，使用与处理扫描的纸质文件同样的方式来理解和提取数据。另外，Word文档、PDF和XML文件，以及电子邮件本身等附件文件必须集成到企业捕获中。

域捕获需要代理商使用智能手机、平板电脑、移动扫描仪或数码相机，或者通过网页界面立刻扫描文件。具体示例包括订单、支票和收据。

智能分类、询问和提取

文件被扫描后，一些会保留为索引图像（例如报告），然而文件上的其他数据比图像更重要（例如发票）。企业捕获通过早期在捕获时集成业务流程集成开始增加价值。

文档类型的容量和多样性对最小化高昂的文件准备费、预分类和数据录入费用至关重要。基于图像和文本的分类软件，通过对图像的识别和分组降低这类费用。紧接着是“询问”，使用基于模式的识别技术（包括OCR）理解基础数据， 然后进行验证和“提取”，将需要提供给业务流程的准确数据删除。最后，将图像和/或数据有选择地路由到适当的业务领域，和用于存储和管理的ECM系统。

为了准确地对图像进行分类、正确地理解他们并且确保能提取到有效和准确的数据， 捕获解决方案必须与适当的业务流程和企业系统紧密集成。元数据在任何阶段都可以添加到图像中以便实现更好的记录控制和管理。

作为企业捕获快速创造价值的一个示例，考虑一个已申请抵押贷款的银行客户。要获得批准，他需要和放贷员坐在一起并提供抵押品材料，如提供收入证明和财务报表。一个企业捕获解决方案可以使放贷员在为客户服务的同时扫描贷款抵押品，并且获取软件自动分类和检查文件（包括直接收入和资产核查）。进而，捕获启动一个流程，由银行员工分析申请人的投资并且建议附加服务（如资产管理）。通过这种方式，企业捕获帮助银行提高了客户服务和增加了服务收入。

一个灵活的, 可扩展的和可伸缩的解决方案

企业捕获软件必须足够灵活以供整个组织内的不同部门和流程轻松采用。一个真正的企业解决方案易于编制捕获流程所需的用于文档分类、数据定位和提取的各种规则和惯例。一旦程序化，这些规则在概念上就应该可拓展，作为网页服务可供任何应用程序调用。例如，客户账户或社会安全查找过程应该由正式捕获进程以外的多个应用程序使用。

在企业捕获中的大容量文档和它的分布特性表面需要一个可伸缩的分布式架构。解决方案应该能够利用今天的多处理器系统的能力，而且捕获应该能通过网络扩展到胖客户端和瘦客户端。云计算也对企业捕获很重要。为了更好地了解传入文档并证实和验证提取数据，云服务允许企业捕获访问远程特性、功能和数据库。另外，云可提供企业捕获部署所需的可伸缩性、计算周期和按需存储系统。

今天的企业捕获

企业捕获处于采用的早期阶段，这个时期的企业以两种方式之一安装了企业捕获。有些是采用现有的部门系统并拓展到多个部门。这并不一定会将运营方式改变为真正企业捕获解决方案，但是我们相信它将演变成为真正的解决方案。其他用户购买了面向业务流程的解决方案，如集成采购至付款的发票处理，并且在多个站点进行了安装，有时带有远程扫描。这更具战略性，因为它可以即刻提供提高的商业价值。在将来，我们相信这两种环境会共同推动集成的捕获解决方案。

注意，企业捕获解决方案可能需要业务流程重新设计为一体化捕获和后台企业功能。虽然这对那些考虑企业捕获的客户来说提高了赌注，但它也表明一系列的业务流程和应用程序中需要灵活且易于自定义的捕获软件。

IBM进入市场

在这种局面下IBM出现了。2010年8月10日，IBM宣布收购 Datacap有限公司，这是一个向世界范围内的机构提供文档捕获软件的死人供应商。

Datacap成立于1988年，提供自动捕获、数据提取和内容分类（包括存储在ECM系统之前的结构化和非结构化文档）的解决方案。Datacap的客户群超过了200家企业，为捕获自动化提供了行业特定的应用程序，如发票处理、医疗保健和政府。其他功能包括部门和中间市场实现协作和基本内容服务；为客户和合作伙伴实现行业快速启动包以便构建垂直解决方案；以及针对许多ECM存储库进行批量捕获。

作为IBM软件解决方案部内的IBM ECM产品组合的一部分，Datacap的产品增强了IBM的FileNet ECM的发行，同时增加了支持非IBM存储库和业务流程。Taskmaster Capture, Datacap的主打产品专为企业范围内的捕获部署，并将IBM定位为综合捕获和ECM解决方案的一站式供应商。

Datacap Taskmaster Capture 的最新版本 v. 8.01 已经集成了IBM的所有三个文档图像和管理存储库：IBM FileNet 内容管理器、IBM内容管理器和 IBM FileNet 影像服务。它也为IBM FileNet 捕获专业版提供了智能文档识别。在收购时，当时的IBM企业内容管理副总裁Ron Ercanbrack说过，“我们选择让Datacap的方法成为 IBM的文档捕获战略的基础。Datacap 的图像捕获方法使用复杂的业务规则管理功能，在业内处于领先地位，是 IBM 最完美的补充”。

一个企业捕获平台

一个IBM收购的关键原因是 **Datacap Taskmaster Capture**推动企业捕获部署的潜力。高管们参与开发了该产品和将产品集成到IBM的ECM产品组合，他们相信，该产品可以满足企业捕获的关键需求，包括以下：

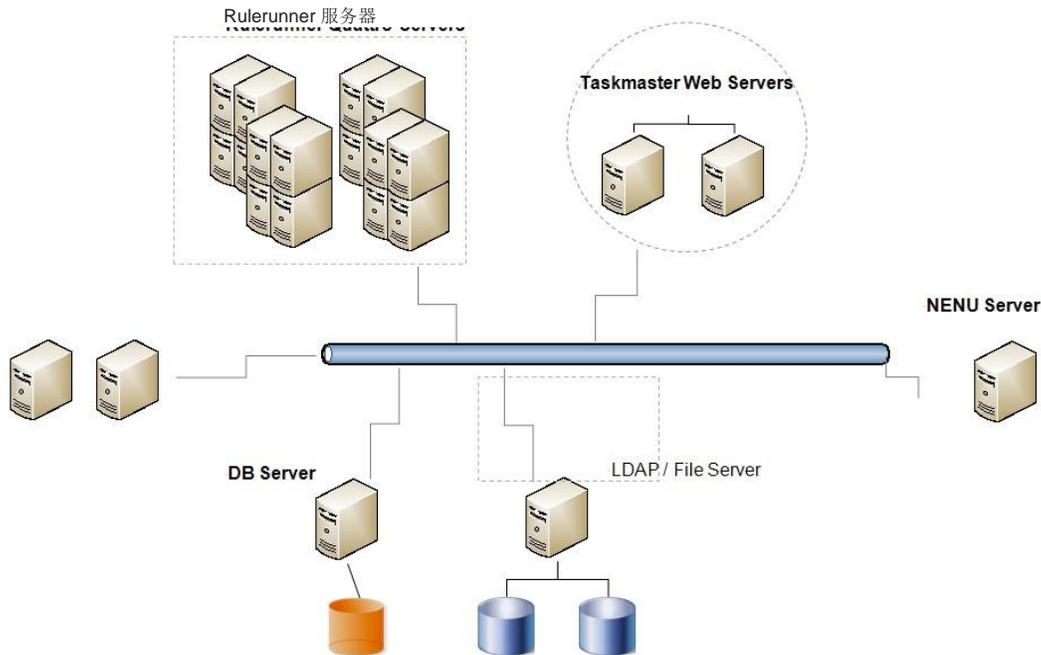


图3-此图显示 **Datacap Taskmaster Capture**的可拓展性。因为**Rulerunner** 服务器处理所有的捕获处理过程，如图像增强和识别，随着容量的增加，组织只需添加更多**Rulerunner** 服务器就可处理额外的工作。因此，**Taskmaster** 可以扩展为每天捕获数百万个文档。

捕获和处理的灵活选择

Datacap Taskmaster Capture 产品具有三层式、主从式架构的特点。捕获可以通过一个瘦客户端的网络界面分布到域、分支和其他文档进入点，同时处理过程可以委托给胖客户端和服务层。随着组织的发展或随着越来越多的文档和部门的加入，处理能力可增加到维持性能水平。

适应性强的, 可扩展的流程管理

为了让捕获成为一个企业标准，自动捕获流程必须在整个组织内易于编制和扩展。**Datacap Taskmaster** 的引擎是 **Rulerunner Service**，这是一个可配置的程序规则引擎，几乎驱动所有的文档捕获处理功能，包括扫描、文档编号、图像增强、文档识别和格式化数据输出。数以百计的预构建规则和操作作为存储在库中的独立执行功能包括在内。使用 **Datacap Studio**，一个应用程序程序构建工具，无需编程经验就可创建或修改规则。针对一个应用程序程序开发的规则可以很容易地应用到其他应用程序和业务流程中。

整个企业可调用的扩展性捕获

企业级捕获解决方案需要可扩展的处理能力，以及捕获流程和惯例可作为服务调用。Datacap Rulerunner Enterprise通过启用Rulerunner 服务的虚拟化实例在多核处理器上运行实现扩展。此外，Datacap Taskmaster Capture 支持面向服务的架构(SOA)来使得捕获流程（比如文档识别和数据定位特性）可以在正式捕获解决方案以外被调用。

结束语

企业捕获可以通过简化流程提高一个组织的经营成果，包括图像、纸张文件和其他非结构化文档。实现这种好处的关键因素包括选择在整个企业内灵活、适应性强且可拓展的捕获软件，然后用企业标准替换多个捕获孤岛。这将支持多个部门将捕获集成为各种各样的业务功能。尽管企业捕获距离它的成熟期还早，但无穷的潜力值得每个企业去探索。

授权组织企业捕获

Scott Blau的采访、IBM Datacap的联合创始人和企业内容管理捕获解决方案主管；以及 Reggie Twigg，全球产品经理，IBM ECM内容和捕获集成。

Q. ECM和捕获之间的关系是什么？

Reggie Twigg:IBM 向企业出售ECM已经有一段时间了。ECM系统 可以视作非结构化内容的记录系统，很像 ERP系统是金融和供应数据链的记录系统。我想我们会开始看到组织在记录系统捕获方面的标准化。

Q. 客户需要什么样的功能可以使企业捕获取得成功？

Scott Blau:在企业层面, 客户需要被授权这项技术; 能够控制它, 随时随地可以使用它并且能根据自身需求进行定制。它远不是一种“万能的”解决方案, 因为每一个部署都有其独特的细微差别。

在过去的五年中, 我们的工作主题是通过减少设置的数量获得授权, 并揭示控制级别, 这样企业用户就可以很容易编排各自的捕获步骤来满足自身的确切需求。

Q. 企业捕获和传统捕获解决方案有什么不同？

Reggie Twigg:它不是企业捕获, 除非有一个企业标准- 所有部门应该使用一种技术进行捕获并遵循一定的最佳实践。我们通常看到整个企业的捕获岛屿, 不同的技术来自于不完整的不同人群。没有一致性, 并且他们在应对4-5-6 或者更多的供应商。这应该引起 CIO发问, 为什么我们需要提供相同技术的所有这些供应商？

Q. 企业捕获怎样授权给一个组织？

Scott Blau:传统上, 捕获被视为一个孤岛, 您可以在文档管理的前端加上这项功能, 这也是使用它的惟一地方。您在进行 OCR、提取信息、集成电子邮件和传真系统, 所有这些都有一点严格的前端进行。但是后续文件通过邮件发给一个管理欺诈案件的组织。您可以打印这些文件并发送给正在进行正式捕获的部门, 但是这种做法太笨了。企业捕获调用您在产品捕获案例管理系统中相同的功能, 并且得到相同结果。能够利用现有技术大大加速投放市场的时间和降低开发和拥有成本。

关于 Harvey Spencer 和联合公司

自1989年起, Harvey Spencer 联合公司总部设在纽约 (US 电话: 1-631-368-8393) 已经专门从事电子信息 (基于图像和电子交易) 捕获技术。我们的服务包括市场分析、技术规划援助、产品定位、产品管理、客户发起的调查和战略计划服务, 这些产品包括高速文档扫描硬件、图像采集软件、字符识别软件 (OCR, ICR)、光标识别技术、条码识别和其他模式的识别和分类工具。

当我们进入电子商务环境, 很明显纸质信息必须进行更有效的管理。做这件事的最好方法是将纸张尽可能快地转变为电子表格。我们专门研究硬件、用来从纸和微型胶片上捕获图像的软件, 以及使用OCR和其他相关识别技术进行电子文档成像和表格处理的软件。我们相信, 先进的模式识别是创建这些文档的理解力的关键, 而且XML和相关技术是这些文档交换的基础。

我们是惟一遵循这个市场规律, 并帮助在这些领域增加价值的专业公司。我们建议开发标准和其他技术, 并且在这方面扮演着领导者的角色。作为一家分析师公司, 我们已经测量并分析了文档捕获软件的市场。

有关更多信息, 请联系: jim.everett@hsassoc.com

关于IBM企业内容管理的联席担保人

IBM 企业内容管理软件能让世界顶尖的公司更快地制定更好的决策。通过控制非结构化信息, 公司可以用新的方式访问、协作和影响业务决策, 使得内容成为一流的洞察力来源。使用特定于行业的IBM ECM 解决方案, 公司可以在整个生命周期内捕获、管理和共享内容, 从而帮助确保合规性、降低成本和生产力最大化。IBM ECM 产品组合包括大量的功能, 可与现有系统集成并帮助组织实现信息价值最大化, 具体包括: 文档捕获和成像、社交内容管理、高级案例管理、信息生命周期管理和内容分析。超过13,000个环球公司、组织和政府组织依赖于 IBM ECM 改善性能和通过创新保持竞争力。

www.ibm.com/software/data/db2