

# Virtual Index Architecture (VIA)

An architectural framework for tackling data sprawl

# Table of Contents



***Two of the greatest information trends today are the huge value of digital information within an organization and the near impossibility of accessing and using this information because of data sprawl.***

## Introducing VIA

Executive Summary .....	2
What is VIA? .....	2

## Architectural Principles

Strategic Principles .....	3
Non-Disruptive .....	3
VM-Optimized .....	3
Location Independent .....	3
Security Integrated .....	4
Data Extensible .....	4
Device Agnostic .....	4

## Architectural Overview

The Building Blocks .....	5
Universal Index Framework .....	5
Cross-Indexers .....	6
Security Integrator .....	6
Index-Enabled Applications .....	7

## How VIA Works

VIA in Action .....	8
An Open API Ecosystem .....	9

## About VirtualWorks

The Content Virtualization Pioneer .....	10
--	----



## Executive Summary

VirtualWorks™ has developed an innovative enterprise information framework, the Virtual Index Architecture™ (VIA), which promises to contain and reverse the overload of information suffered by organizations of all sizes.

VIA defines a distributed system that collects, normalizes and catalogs corporate data into a common company index. The information in the index is then accessible by applications specifically written to the VIA Application Programming Interface (“API”), which enforces the underlying security associated with the data elements.

With this inherent flexibility, VirtualWorks will develop or partner to develop plug-ins to cover the vast majority of commercial applications and security subsystems currently deployed by mid-market organizations.

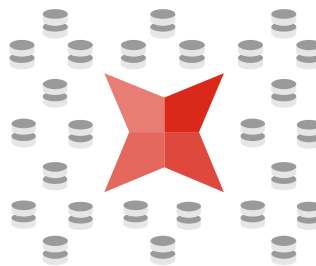
The VIA-enabled applications will provide a real-time image of the most relevant, business-critical content in a security context that mirrors the rights of the end user. These applications will lead to more productive, nimble and competitive organizations.

## What is VIA?

The Virtual Index Architecture is a roadmap that defines a highly extensible framework that supports standardized data collection plug-ins, security plug-ins and Index-Enabled Applications. This architectural framework defines how VirtualWorks’ content virtualization solutions understand how to track, interpret, index and display corporate content that may span disparate storage architectures, applications and data structures.

VIA enables Index-enabled applications that leverage a deep understanding of application data and stored metadata to render images of enterprise content stored in a variety of application systems, network topologies, cloud architectures, data storage systems and formats.

*VIA is designed from the ground up to make it fast and easy for developers to create new index-enabled applications and plug-ins that optimize the value of information for organizations.*



# Architectural Principles

*VIA works across all applications and platforms seamlessly without manual tagging, ripping or replacing; just plug and play.*

VIA defines how VirtualWorks' content virtualization system works and how the system will be enhanced with new capabilities over time. The architecture has been specifically conceived to allow VirtualWorks and its partners to easily create new Index-Enabled Applications and plug-ins that optimize the value of information for organizations.

## Strategic Principles

The Virtual Index Architecture is built on a set of core principles that reflect the tenets of content virtualization, making it a uniquely powerful solution for taming data sprawl. The premise behind VIA is simple: to deliver a pre-integrated content virtualization platform that makes it easy to find data and put it to work across the enterprise.

## Non-Disruptive

Creating a Universal Index for the enterprise is not the same as building yet another massive data warehouse. In fact, it's quite the opposite. VIA leaves content where it naturally lives, working across all applications and platforms seamlessly without manual tagging, ripping or replacing. The system does not interfere with the structure or integrity of data in performing its role. It just sits on top of other platforms and existing security systems in a lightweight model while building the Universal Index dynamically and continuously.

## VM Optimized

VIA is based on a distributed architecture that's well suited for virtual environments. It's a modularized and scalable approach that fully parallelizes the Universal Index processes. The individual indexing processes are designed as a modular stack and can be spread across multiple physical or virtual machines, offices and times of day, making it quick and easy to create, maintain and query a Universal Index. The entire system can dynamically expand or contract to multiple virtual hosts as VIA workloads fluctuate.

For instance, when the demands are greatest, VIA can 'fan out' to run workloads on any number of virtual machines to maximize throughput, then collapse back to one machine to minimize footprint when things slow down. And the system can be managed using standard virtual machine tools. This virtual design means VIA is high-performance, flexible and manageable.

# Architectural Principles

## Location Independent

VIA is designed to easily integrate any cloud-based application or data store. Its expandable architecture indexes all data, whether on-premise or in the cloud, regardless of format and structure. And VIA is dynamic: if data moves, the Universal Index tracks it without the heavy-lifting of data replication or physical integration. For the first time, content can be stored anywhere: behind the firewall, out in the cloud or anywhere in between.

## Security Integrated

VIA doesn't replace existing security schema; it's designed to fully integrate directly into it using all existing rights and privileges across the enterprise. This provides two major benefits. First, it automatically adapts to whatever security authorization model is in place. The index is created within the context of the existing security authorization model. Second, the system provides real-time security. Permissions are checked whenever the Universal Index is queried; and you never have to modify the index, even if you update your security on the data. All queries go through additional security verification prior to delivery of results to the end user, so they will only be able to view documents they are authorized to see.

VIA's single sign-on support makes it possible for end-users to sign-on using Microsoft Active Directory or other authentication systems while

performing secured index operations against sources using different security models (such as LDAP). Different access control mechanisms can also be combined in the same index operation. Because VIA is service-oriented, it is designed to adapt to any security authorization model using simple, fully integrated plug-ins.

## Data Extensible

Data never stops evolving. VIA has been architected from the ground-up to embrace new applications, databases and file formats. It provides an extensible indexing framework using non-disruptive plug-ins for any application, system or database - without having to extract a copy of the data before it's tagged and indexed. VIA uses Cross-Indexers to capture the data schema of applications and plug them into the Universal Index. Cross-Indexers are standard, pre-integrated programs that tell the Universal Index how to act by showing them where the data targets are by creating searchable, de-normalized data records. They understand the context of the data inside each application, so the knowledge inside is usable by the Universal Index. That's what makes the Universal Index so extensible - by embracing all data types and file formats, structured and unstructured.



# Architectural Overview

*Cross-Indexers understand the structure and metadata of specific applications and are used to plug any application into the Universal Index without custom integration.*

## Device Agnostic

VIA's exportable API means any new device or platform can be supported - physical, virtual, distributed and mobile - without changing the core engine. VIA can be expanded to leverage the native features of any platform including Windows, Mac, iOS, Android and more, so any user and any device can harness the power of content virtualization.

## The building blocks

VIA is based on an open, service-oriented architecture that embraces all of a company's applications and data stores, while keeping the indexing overhead light. The scalable system is engineered to index both static and dynamic data and work seamlessly with any SQL-based application. The VIA system can be extended to include different data types, data organization schema or security subsystems with architected data interpretation or security mapping plug-ins.

Additionally, VIA has been conceived as a fully parallelized and distributed architecture, allowing data collection components to run locally or remotely within the enterprise firewall or in a public or private cloud. It is designed specifically to address the challenges posed by data sprawl in the enterprise through a layered architecture that makes extensive use of data schema and security plug-ins. The architecture has four key building blocks:

## Universal Index Framework

The heart of the VIA framework is a single, unified view of enterprise data assets called the Universal Index. The Universal Index Framework employs a highly distributable/parallelizable model to manage the data collection, conversion and indexing processes, which run continuously and record changes in data in real-time. The Universal Index Framework components may work in a single machine or they may be distributed among multiple machines (physical or virtual). The Universal Index Framework also manages the security authentication, the data plug-ins and implements the Universal Index Application Programming interface ("API"). VIA's indexing framework is platform agnostic and provides support for all major virtualization systems and applications.

## Cross-Indexers

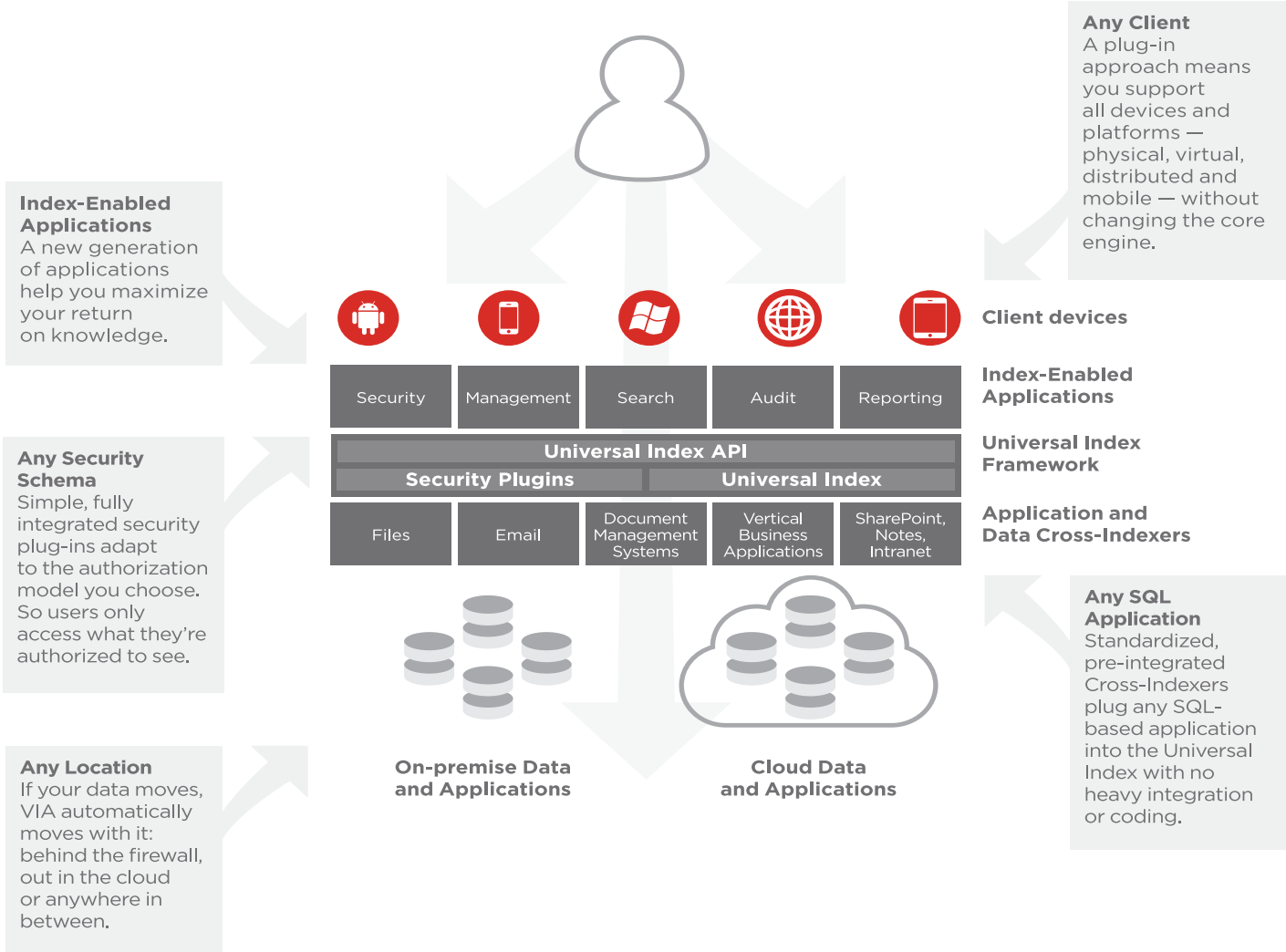
In VIA, the data interpretation plug-ins are called application Cross-Indexers. All Cross Indexers are standardized, modular and pre-integrated to their target applications. A primary task of an application Cross-Indexer is to de-normalize an SQL application's data schema into a small number of record types that can be added to the Universal Index.

For example, a CRM application can be reduced to "people," "company" and "transaction" records. Subsystem Cross-Indexers support non-relational data stores.

# Architectural Overview

All application or subsystem Cross-Indexers have the ability to insert data or application-specific metadata, which could later be used by an application to filter the Universal

Index query results. Cross-Indexers can easily be installed, fine-tuned and maintained by value-added resellers - with no system integration requirements.



# Architectural Overview

*Once the data is indexed, VIA opens up a whole new set of options for securely accessing content across a variety of locations and devices to increase employee productivity and information insight.*

## Security Integrator

The Universal Index Framework collects, converts and adds data to the Universal Index with 'super user' privileges. Once data is safely sequestered in the Universal Index, it may only be retrieved via the Universal Index API. To insure no user retrieves data for which they do not have access rights, the security integrator performs a run-time security check for each data element. Thus, an Index-Enabled Application will only see data records that its user has the authority to see.

The Security Integrator supports standard security schema plug-ins, which define how and where the Security Integrator determines the access rights for each record type. This extensible approach embraces new security schema as they emerge, without expensive re-integration. 'Run-time' security evaluation ensures changes made by an administrator to access rights are immediately reflected in the Universal Index API results.

## Index-Enabled Applications

A trusted, all-inclusive index that's maintained dynamically opens up a new world of applications that couldn't exist without a Universal Index.

Once the data is indexed, VIA opens up a whole new set of options for securely accessing content across a variety of locations and devices to increase employee productivity and information insight. These Index-Enabled applications can do powerful things with enterprise content, allowing users and administrators to:

- **Find it** - using next generation search to look across all applications and file stores in one step.
- **Secure it** - based on total information visibility by monitoring, modelling and managing security access.
- **Inventory it** - to make better resourcing decisions.
- **De-dupe it** - to improve data quality enterprise-wide.
- **Manage it** - to increase the value and availability of your content while managing its impact on your infrastructure.
- **Learn about it** - to gain insight into how people use knowledge as an indicator of performance, providing early insight into important trends and issues.

# How VIA Works

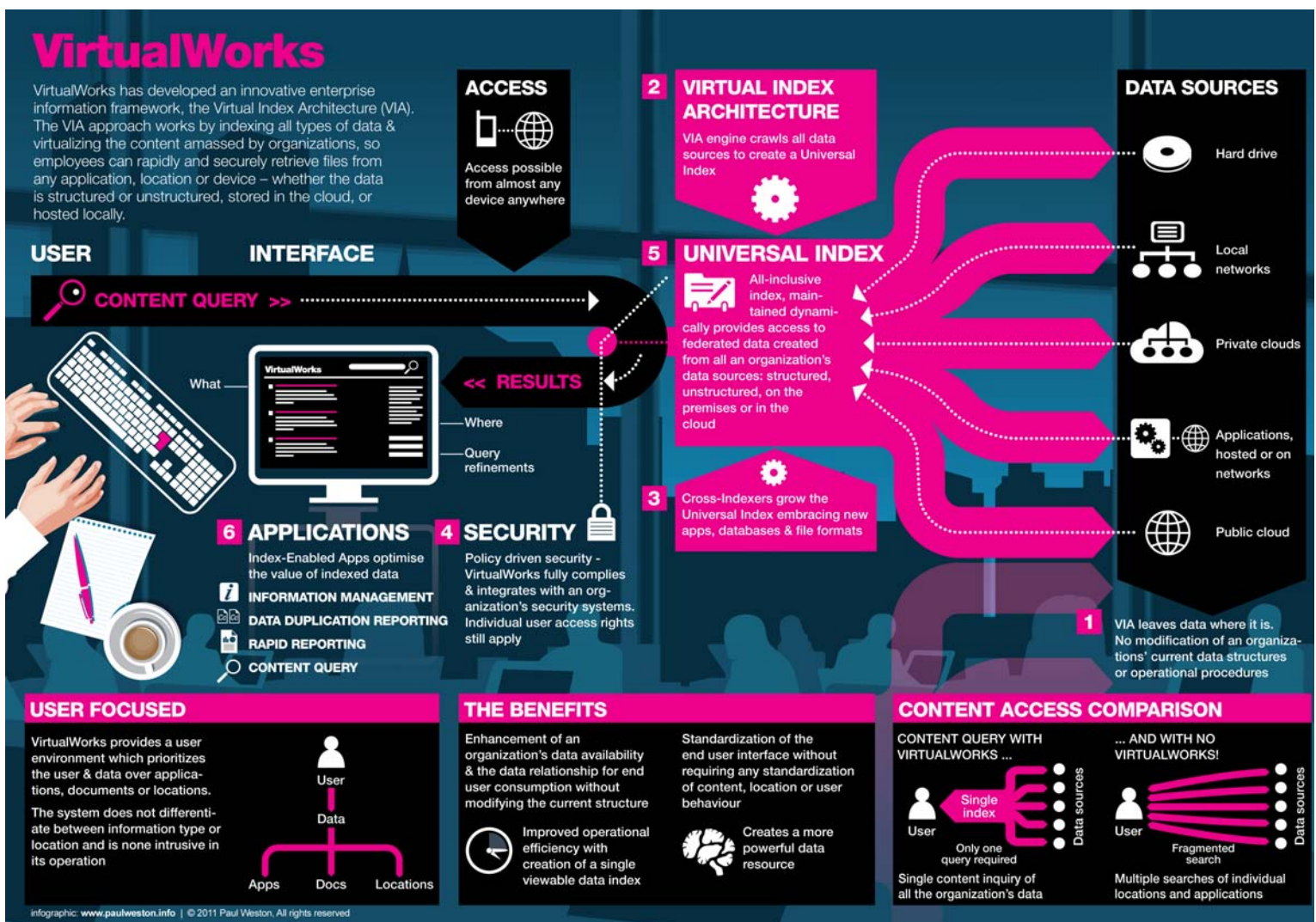
## VIA in Action

VIA inserts a virtual content layer above an enterprise's data silos to index every application, map every content store and record all metadata into a single Universal Index. The index is federated across the enterprise and is available to users and administrators, so they can securely find and access everything from anywhere without leaving their applications or devices.

Here's a quick snapshot of the Virtual Index Architecture in action:

## 1. Leave content where it lives.

Unlike complex systems that try to impose a structure on data or new processes on users, VIA lets data stay where it naturally lives but indexes it for easy access. All data is left intact on its original source without movement of data, disruption of source data or manual tagging of data types.



# How VIA Works

*VirtualWorks will publish the APIs, so third-parties can develop value-added applications and plug-ins for the platform that improve any company's Return on Knowledge.*

## 2. Content virtualization layer attaches to all data silos.

VIA crawls data sources in a non-disruptive manner, using a limited amount of available bandwidth. The engine converts this indexed data into a usable format for any application powered by the Universal Index. This converted data element is placed into the Universal Index and becomes usable by any authorized application.

## 3. Cross-Indexers grow the Universal Index.

VIA embraces new applications, databases and file formats through the simple addition of Cross-Indexers, a standardized component that quickly plugs any application or data element into the Universal Index.

## 4. Policy-driven security infrastructure performs real-time authentication.

VIA integrates with Active Directory and other authentication systems for secure access control. Permissions are checked whenever the Universal Index is queried, so users only view documents they are authorized to see.

## 5. All content is federated into a single user interface.

VIA provides single-point access to all enterprise data stores, structured and unstructured, on premise or in the cloud, without leaving your application or device.

## 6. Index-Enabled Applications optimize the value of indexed data - whether for reporting, auditing, search or even creating new applications based upon a common index.

## An Open API Ecosystem

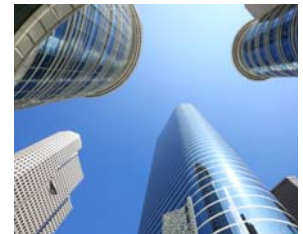
VIA embodies a rich set of APIs available to applications, which VirtualWorks is using to develop a range of Index-Enabled Applications. In addition, we will publish the APIs, so third-parties can develop value-added applications and plug-ins for the platform that improve any company's Return on Knowledge.

# About VirtualWorks

## The Content Virtualization Pioneer

VirtualWorks is pioneering a new software category called content virtualization that breaks down the walls between application silos to ease data sprawl and deliver a more productive end-user computing experience. Headquartered in Boca Raton, FL, with offices in Norway, the company is led by Ed Iacobucci, a prominent high-tech entrepreneur and co-founder of Citrix Systems.

The company's technology is currently used by more 300 private and public sector organizations in Norway, Sweden and the United States to unleash the power of information no matter where it resides within the organization. For more information, please visit [www.virtualworks.com](http://www.virtualworks.com).



***“This is a major  
paradigm shift  
without the major  
disruption.”***

Ed Iacobucci  
CEO, VirtualWorks





**USA**

5301 N. Federal Highway  
Suite 190  
Boca Raton, FL 33487  
561.327.4900

**Norway**

Lysaker torg 45  
1366 Lysaker  
(+47) 81 536 863

**Sweden**

Kanalgatan 51  
931 32 SKELLEFTEA  
(+46) 70 543 7700

[www.virtualworks.com](http://www.virtualworks.com)