Governance and Compliance for Unstructured Data: A White Paper

What's the Problem? Why Should I Care?

Starting from the days when organizations provided their staff with computers connected to network shares, runaway data has been an issue. Back then, no one was considering compliance or regulation and so the principal concerns were cost saving and efficiency. As backup times increased and storage costs rose, forward-looking businesses took steps to curb this growth through a range of approaches.

Electronic discovery for litigation raised the profile of runaway unstructured data – that is, data in file shares, mailboxes and elsewhere. In the last decade, the rise in prominence of privacy regulation and the inevitability of data breach has placed a far clearer cost on poorly managed and governed data. Organizations must now consider the risk of sanction, the cost of discovery, the cost of breach notification and liability and, after the event, the cost of recovery.

Whether directly through terms or agreements, or indirectly through buying patterns, enterprise customers are increasingly showing that they care about how their data is handled. Businesses are recognizing that being seen to be a good steward of data has impact on brand too.

ActiveNav and Unstructured Data

Since our very beginning, when organizations were concerned with enterprise search and 'bringing google-like' experiences to its staff, ActiveNav was developing solutions focused on empowering its customers to deal with the implications of the exponential growth of user generated data. Our aim was to enable them to understand data at scale and act on what they found. This led to approaches to handle a range of problems, from Redundant, Obsolete and Trivial (ROT), through duplication, to sensitive data and records classification.

The knowledge that ever-growing volumes of unstructured user-generated data places all organizations at risk has led us to our North Star. Zero Dark Data (ZDD) describes a state where an organization understands the data it holds and, as a result, take policy-based decisions to defensibly, and proactively, mitigate the hazards within. In a way, we think that modern regulations are only just catching up with what we've known all along – that all organizations must undertake a journey toward ZDD if they are to become good stewards of the data they hold, regardless of its origins.

What's Special About Unstructured Data

We have invested our time and energy in addressing unstructured data because it is uniquely difficult for organizations to govern systematically. Why difficult? Well, unlike application data (or structured data) - which exists in a service or database with a clear business purpose - unstructured data has no such context. In discovering a server full of files, an Exchange mailbox or a chat channel in Slack, you can expect to find content from any business function, or even a personal file cache.

Driven by our ZDD North Star, we've chosen to take this problem head on. Our solutions are dedicated to digging deep and providing the insights necessary for our customers to take decisions to mitigate the risks in their data. As a result, we believe that any organization that is serious about the compliance and stewardship of its data will recognize that our solutions provide a foundational part of their information governance program.

Americas
Reston, VA, USA
+1 571 375 2780

EMEA
Winchester, UK
+44 01962 454401

APAC
Melbourne, Australia
+61 3 9982 4543

Contact sales@activenav.com
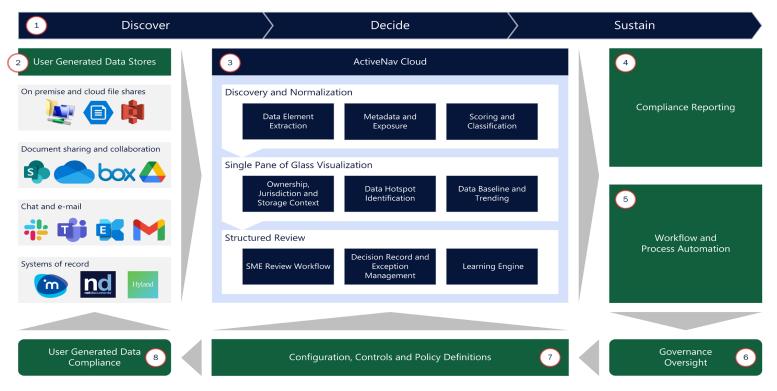**Tweet** @ActiveNav
Visit ACTIVENAV.COM

Our Solutions

We provide data discovery solutions for unstructured data. Our software is designed to enable our customers to discover that data, wherever it is stored, and enact policy-based decisions to control that data based upon its risk and compliance profile. We focus only on unstructured data because it's difficult to handle and because we know that our competitors don't adequately address its inherent challenges, leading to project failures and, as a result, programs falling by the wayside as they cannot be sustained.

The following shows how our Cloud service fits into any enterprise's existing environment to discover and monitor connected unstructured data repositories as part of an overall governance program.

1	Discover	Decide	Sustain

2 User Generated Data Stores

On premise and cloud file shares

Document sharing and collaboration

Chat and e-mail

Systems of record

3 ActiveNav Cloud

Discovery and Normalization
- Data Element Extraction
- Metadata and Exposure
- Scoring and Classification

Single Pane of Glass Visualization
- Ownership, Jurisdiction and Storage Context
- Data Hotspot Identification
- Data Baseline and Trending

Structured Review
- SME Review Workflow
- Decision Record and Exception Management
- Learning Engine

4 Compliance Reporting

5 Workflow and Process Automation

8 User Generated Data Compliance

7 Configuration, Controls and Policy Definitions

6 Governance Oversight

Here's what each piece of the model provides:

1. **Discover, Decide, Sustain.** A consistent approach to dealing with unstructured data initially, and in sustainable business-as-usual.
2. **User Generated Data Stores.** Any repository where user generated data is habitually stored, be it on-premise or in the cloud.
3. **ActiveNav Cloud.** Our Cloud service provides single pane of glass insights into all connected data and includes the workflow and decision support required for customers to make mitigation decisions.
4. **Compliance Reporting.** Third-party management reporting tools (such as privacy data maps or senior leadership dashboards) that draw insights from ActiveNav Cloud drive program reporting, specifically, how much data, of what type, is where.
5. **Workflow and Automation.** Robust change management requires that data disposal be handled through an organization's existing IT workflow. Decisions made in ActiveNav Cloud are passed into these services for action.
6. **Governance Oversight.** A functional governance program needs an oversight group comprised of data owners, compliance, IT and leadership to set policies and controls.
7. **Configuration and Controls.** From records schedules through acceptable use policies to privacy controls, ActiveNav Cloud identifies controlled or regulated data wherever it lies.
8. **Data Compliance.** As controlled data is identified and disposed of, monitored data sources become and remain defensibly compliant.

ActiveNav Solutions Overview

We provide two software solutions; first, our new Cloud service sets new standards for performance and usability in Data Discovery, and, second, our on-premise DC Project Suite which excels in focused and highly-detailed unstructured data transformation projects.

ActiveNav Cloud

A new data discovery service, hosted in Microsoft's secure Azure infrastructure, providing secure, high-performance, low-cost discovery and compliance monitoring for unstructured data wherever it lies

DC Project Suite

An expert information toolkit, matured and hardened over many thousands of hours of real-world projects from data clean up, through records and sensitive data identification to migration preparation

Key Capabilities

Discover aging, unwanted, duplicated and sensitive data in any connected repository

Classify and score controlled data against any regulation or policy

View data distribution, age, duplication and compliance in a single pane of glass across geographies, repositories or business units

Identify and prioritize controlled data hotspots according to risk score

Drill into hotspots to review responsive data in detail

Triage authorized data, mark unauthorized data for disposal

Target search across all discovered data for Data Subject Access Request Fulfilment

Integrate with third party reporting, privacy and workflow tools as part of a compliance program

Key Capabilities

Discover aging, unwanted, duplicated and sensitive data in any connected repository

Configurable classification for metadata creation, sensitive data, labelling or record schedule mapping

Report on age, size or other metadata, duplication or any configured policy

Model data against hierarchical classification or file plans

Delete, quarantine, migrate discovered data directly from reports

Transform and re-organize data in place or on migration

Write metadata or MIP labels to data in place or on migration

Map data against new file plans as part of a file plan migration or re-structure

Supported Repositories

Windows and network shares on premise
Azure, Amazon S3 or GCP file shares
SharePoint Online and OneDrive
Teams documents and chat, Exchange online
Google Workspace Drive, iManage
Soon: Box, Netdocs, GoogleMail, Slack

Supported Repositories

Windows and network shares on premise
SharePoint Online
SharePoint on premise
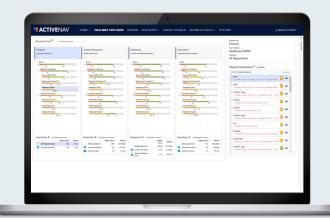OpenText Content Server

Why Use ActiveNav?

Any organization addressing its governance and compliance obligations needs to set its risk appetite against its budget and resources. When the time comes to address the risks inherent in its unstructured data it must consider some harsh realities:

- Unstructured data governance projects treated as one-time efforts frequently fall short of and yield diminishing value over time.
- Seeking 'easy' fixes by asking end users to govern data is fraught with issues and, at minimum, needs to be complemented by comprehensive data discovery capabilities.
- Vendors offering 'do everything' solutions have genuine execution blind spots for unstructured data where their feature-function lists don't translate into practical solutions.

This means that there is a very real likelihood that any unstructured data governance implementation will fail because the chosen vendor lacks the experience and specialized features required for success. Simply put, if your organization considers governance of unstructured data to be important, ActiveNav has the experience and the focus that will get you started fast and maximize the chances of a sustainable governance program which becomes embedded in your business-as-usual rathe than a one-time effort.

About Us

We're data experts, and our North Star is Zero Dark Data. We believe that all organizations should be aiming for a state of Zero Dark Data so that they can act as good stewards of that data to minimize their cyber risk surface area, protect the interests of their customers, their staff and their other stakeholders. We've been working continuously with unstructured data in the wild for well over a decade and we think the market deserves data discovery products that just work. As a result we are trusted by leading companies and government agencies to help them understand and control their data assets to drive regulatory compliance, reduce the cost of data ownership and improve data quality.

Aside from our fascination with Dark Data, we're engineers, designers, runners, gardeners, chefs, photographers, bikers, parents, skiers, cosplayers, hikers, gamers, travelers, mountain climbers, friends, race car drivers, readers, volunteers, and car enthusiasts. We value loyalty, accountability, and communication and care deeply about creating a sustainable future, supporting charitable causes as proud members of Pledge 1%.

Request Demo

→ 15 Years of Experience

→ 6 Continents and 28 Countries

→ 15+ Billion Files Discovered

→ 30K+ hrs Customer Deployments

→ 300+ Customers and Counting