How to Get Started With Content Analytics

Your organization is sitting on a treasure trove of information! How can you get to those treasures? Content Analytics is the answer. AIIM defines content analytics as using defined rules as well as mathematical, statistical, and semantic models to automate how content is captured, analyzed, and governed over its lifecycle. Seth Grimes shared that 'text analytics uses linguistic and statistical techniques to extract concepts and patterns that can be applied to categorize and classify documents, audio, video and images.' (Source: 2011 Text Analytics Summit)

Through the use of content analytics, you will be able to help your organization make better decisions, improve products and services to customers, and improve internal processes. Content analytics is also driving auto-classification, content remediation, Information Governance, security correction, adaptive case management, risk reduction, and operations monitoring.

This toolkit will help you to get started analyzing your data to be able to harvest the valuable nuggets of information from the vast amount of information in your organization's systems.

Step 1 - Identify the Value of Information

Every organization has information. Some of the information is captured and maintained to fulfill regulatory requirements and some of the information is captured and maintained to support the day-to-day operations of the business. In this step, we need to identify the valuable content in our organization, and to whom it is relevant.

AIIM Tip

The "who" of the content will shed light into the relevance of your content on a different scale: for example, a reference document may not be relevant to you, but may be critical to the task of a colleague. Knowing the "who" will also help ensure you don't mis-classify information as junk when it really isn't.

Begin by analyzing the information to determine the type of information it is and how long the organization needs to maintain it for either legal or business reasons. If you have not already created a retention schedule this exercise forms the basis of your retention schedule. Now would also be a good time to establish classification rules and train your auto-classification system, if you have one, with examples of key pieces of information to assist with future classification activities.

Look across the range of your business activities to see where analytics could provide insights to better meet customer needs, improve competitive advantage, or improve internal business processes.

For more information:

- [Six Ways to Measure the Value of Your Information Assets](#)
- [#infochat - The Value of Information](#)
- [I Can't Can You? Valuing Information](#)
- [The More Things Change](#)
- [The Information Boom](#)
- [I Can't Can You? Valuing Information](#)
- [I Think I Can - Valuing Information](#)
- [Why and How to Value Information](#)
- [6 Questions To Ask When Preparing Data for Analysis](#)
- [Fester, Fester, Fester. Rot, Rot, Rot](#)

Step 2 - Define Clear Corporate Objectives for Content Analytics

Clear corporate objectives should be laid out prior to the purchase or development of any content analytic systems. Your objectives will determine what data needs to be captured and at what level of detail. Your objectives will also drive the reporting needs of the project. As you are setting the objectives for your content analytics project, make sure you take time to align your retention policies to the new business needs of content analytics. The business needs for information may be longer than the legal retention requirements for the same information. This will help you to improve your compliance and risk exposure as well as to control increasing storage requirements and help you to identify potential uses for content analytics.

AIIM Tip

Make sure you keep the BIG PICTURE in mind as you are setting your corporate objectives for content analytics. You don't want to only think of one potential use. This means you should keep in mind the audience of your analytics. If you have a very broad audience, you may look at capturing different data than if you are targeting a very specific team.

You may wish to render your objectives and goals in the machine-readable, XML based StratML format as specified in ISO 17469-1. Using this standard will enable efficient discovery, sharing, and analysis of the content your organization maintains.

For more information

- [Data Sets You Free](#)
- [More Companies Need Strategy to Manage Content](#)
- [Better Questions to Ask Your Data Scientists](#)
- [Question the Data: How to Ask The Right Questions To Get Actionable Insights](#)
- [5 Ways Data Scientists Add Business Value](#)
- [Comparing Analytics and Data Science](#)

- The Importance of Clear Objectives, Directions & Project Plans
- How Setting Goals and Objectives for Projects Leads to Successful Outcomes
- Project Management: Project Goals and Objectives
- ISO 17469-1: 2015, Document management – Strategy markup language (StratML) – Part 1: StratML core elements

Step 3 - Determine Technology to Use

All organizations have content in numerous formats including documents, spreadsheets, images, emails, etc. The sheer volume of content organizations manage is staggering. What is even more staggering or perhaps daunting is the task of analyzing that content. The analysis of the content can be made easier through the use of automation. The industry offers tools that organizations can use to analyze their content. It is important to select the correct software to meet the business requirements.

In addition to analyzing the content, organizations should look at how they can manage information as it enters the organization. Establishing a digital mailroom will help to automate the capture of information from the very start, setting you off on the most correct path towards success.

> *AIIM Tip*
>
> *Content analytics should not be thought of as simply an ECM solution. It is much more. It extends elements of ECM to broader scale applications of information. If your organization is not using digital mailroom technology, now is a good time to implement it.*

For more information

- Best practices for evaluating and selecting content analytic tools (Webinar)
- Best Practices for Evaluating and Selecting Content Analytic Tools (webinar transcript)
- Content Analytics: The Next Big Thing in ECM?
- Selecting Software: How to Choose the Best Apps for Your Business
- A Software-Assisted Qualitative Content Analysis of News Articles: Example and Reflection

Step 4 - Access Data Structure, Tagging and Permissions at Creation

As information comes into an organization, it is important to make sure it is properly tagged with the appropriate metadata so that it can be found later, as well as reported on correctly. Since this process can be time and labor intensive, organizations will often delay tagging efforts, hoping to "get back around to it someday". Unfortunately, that day doesn't come. Or when it does, the volume of content that must be tagged is simply too overwhelming and the tagging efforts fall short.

The best practice to implement is to tag and determine the permissions for all content as it enters the organization, which is another reason to implement a robust digital mailroom in your organization.

AIIM Tip

It is important to use the people who ultimately use the content to establish the schemas for tagging, permissions and data structure. They understand the information and how it is used.

Review your metadata to ensure it is being used accurately to separate the ROT from the valuable information your organization receives. Align your content and metadata.

For more information

- Data Structures & Algorithm Basic Concepts
- Introduction to Data Structures
- Data Structures and Algorithms Complete Tutorial
- What is Tagging?
- What Are Tags and What Is Tagging?
- File and Folder Permissions

Step 5 - Improve Search with Auto-Classification

Through the use of auto-classification, you can ensure that your contextual search is properly tuned to provide the correct results to help your organization to be successful. It is also important to make sure that your staff knows how to use the auto-classification tool to further ensure that the content they create is tagged so that it will be able to be easily found in the future.

Your auto-classification tool is not only good for your contextual search and for ingesting your content. This tool can help you to determine the content to delete that no longer has value to the organization. The auto-classification tool can also improve all searches in terms of accuracy and completeness. Be aware that the auto-

classification tool has a built-in level of error. You will need to have procedures and processes in place to validate the classification performed by the tool.

> *AIIM Tip*
>
> *Remember that humans hate sifting through ROT to find the files they need. This is where the use of auto-classification can really be a benefit to the organization. As you are setting up the auto-classification system, make sure to identify suitable criteria to allow for proper cleansing of the files. Be aware that your information governance policies need to be updated and consistent.*

For more information

- [How to Improve Search Results with Auto-Classification](#)
- [Just the Facts - Auto-classification and Taxonomies Webinar](#)
- [How "Auto" is Your Document Auto-Classification? Five Questions to Ask](#)
- [Information Governance Training](#)
- [The Future of Search Engines is Context](#)
- [Contextual Search is the Future of Mobile](#)
- [Using Auto-Classification to Improve Your Information Governance Practices](#)

Step 6 - Improve Processes Based on Content Analytics

We are making progress towards our goals of increased operation efficiency and the extraction of more value from the information we already have. Eliminating manual processes and taking paper out of business processes is a great first step, but we still have a long way to go. By analyzing inbound content, we can improve the processes. Analytics can identify the common metadata for the content entering our organizations, which then helps to feed the content into automated routings to specific processes.

> *AIIM Tip*
>
> *Take control of your emails. If you have no emails archived or your policy is to archive the email and forget it, you are losing potential corporate knowledge and potentially exposing the business to risk. This is a good time to review your processes to identify bottlenecks that can be freed up and remove manual steps to improve process productivity*

For more information

- [How to Use Analytics to Drive Process Improvement](#)
- [The Whole Picture Using Process Intelligence to Extend Business Intelligence](#)
- [Business Process Management Training](#)

Step 7 - Monitor and Tweak

Once you have completed analyzing content in your organization, you are not finished. Times change, business objectives change, and information will still flow towards you at an astonishing rate. Now, is time to monitor the improvements you have implemented and be on the lookout for things that need to be adjusted to further improve the analytics.

Establishing some success metrics will help keep monitoring efforts on track, while simultaneously telling you how successful content analytics has been deployed in your organization. In addition to success metrics, you will want to measure your progress in processing your content through the analytics software or process.

> *AIIM Tip*
>
> *Always try to improve what you have established. After monitoring for a while, begin to tweak your metadata, tagging, data structures and access to continue to meet your business objectives.*

For more information

- [4 Tips for Monitoring (and Measuring) Projects](#)
- [10 Ways to Improve Your Project Today](#)
- [Big Data and Content Analytics: measuring the ROI](#)
- [AIIM Trendscape: Content Analytics](#)
- [Content Analytics: automating processes and extracting knowledge](#)