Measuring the value of technical content

A comprehensive and holistic architecture for content analytics

Michael Iantosca
Senior Director of Knowledge Platforms and Engineering
Avalara

At a recent industry conference on content, moderators conducted a live poll asking attendees to identify the biggest challenge facing technical content organizations today. When the results were revealed, no one was shocked: the top challenge was proving the value of content to the business.

What shocked the audience, however, was that the results being read weren't from the poll just taken; they were from the same survey conducted 30 years ago. The results of both polls were identical. Despite decades of technological advancements and evolving business strategies, content teams, including those at leading global brands, still struggle to demonstrate their impact objectively and quantitatively.

Why does this challenge persist? The reasons are complex, but one thing is clear: solving them remains as critical - and elusive as ever.

Measuring the value of technical content is difficult and costly

The impact of technical content is often indirect and difficult to isolate. Unlike direct revenue-generating activities like sales or marketing campaigns, technical content primarily enhances customer experience, reduces support costs, and improves product adoption—factors that are not always straightforward to measure but affect revenue and expense.

Another reason is the complexity of attribution. When a customer successfully resolves an issue using documentation, it's challenging to quantify how much that contributed to customer retention, satisfaction, or reduced support interactions. Organizations struggle to connect content effectiveness to business outcomes without a clear attribution model.

Access to technical documentation is often free and doesn't require users to authenticate, so teams can't immediately track precisely who is reading the content, how they use it, and how they measure its impact.

Additionally, many content organizations lack standardized metrics and KPIs. While marketing teams have well-established KPIs like conversion rates and customer acquisition cost, technical content teams rely on qualitative feedback and page views, which rarely translate into monetary value and impact.

Organizational content silos further complicate the issue. Content teams, support teams, training teams, and product teams often operate separately, making it difficult to align on shared goals and integrate data. Without cross-functional collaboration, tracking the business impact of content remains fragmented.

Bean counting

Many business leaders see content as little more than a necessary expense—like office supplies or the company coffee machine. This view is not only misguided but also overlooks a critical fact: after the sale, technical documentation is the second most interacted-with element of a product, right behind the product interface itself.

Technical documentation is the second most interacted-with element of a product, right behind the product interface itself.

Despite its undeniable role in customer experience and profitability, content teams are constantly asked to justify their existence with data. "*Prove your value*," said the executives, as if measuring the nuanced impact of content is as straightforward as tracking sales figures. The reality? It's an expensive, complex, and often impractical endeavor that requires sophisticated tools, expert analysis, and significant investment—ironically, content teams struggle to secure the resources without first proving their worth.

And so, the cycle continues: **no investment without proof, no proof without investment**—a classic corporate Catch-22 that leaves content teams scrambling to measure the immeasurable while business leaders wonder why their customer satisfaction scores are tanking.

Yes, everything *is* measurable – at a cost

Some folks will inevitably push back and insist, "Everything is measurable!" And do you know what? They're absolutely right. But here's the catch: just because something *can* be measured doesn't mean it's practical—or even remotely feasible, especially regarding technical content.

The challenge isn't just collecting data; it's collecting *useful* data. Most commercial content platforms do offer metrics, but they exist in separate silos, blissfully unaware of one another. You might get analytics from a content management system, but what about the content delivery system? Those two often live in entirely different worlds, refusing to share secrets like feuding relatives at a family reunion.

A well-functioning content supply chain spans everything from authoring tools and multiple CMSs to delivery channels, localization, taxonomy, and quality management, not to mention whatever hot new trend emerges next week. While individual systems churn out raw numbers, making sense requires stitching together multiple data sources to identify cause-and-effect relationships. And spoiler alert: that's neither quick nor easy.

Take something as seemingly straightforward as bounce rates. Knowing people leave a page quickly is helpful, but *why* are they bouncing? To figure that out, you'd need to analyze content quality. Unless you have an automated quality management system, that means manually scoring content—a task as enjoyable and scalable as counting grains of sand at the beach.

The usual fallback is a manual correlation. But since doing that continuously is expensive, time-consuming, and borderline soul-crushing, teams usually attempt it sparingly or in small samples. Gathering customer feedback is another option, but it's costly and subjective and provides only fleeting insights rather than a comprehensive, ongoing picture.

So, the next time someone confidently declares, "Anything can be measured!" smile, nod, and hand them an invoice for the army of analysts, developers, and data scientists required to make it happen.

Or, better yet, consider a more sustainable approach—one that too many teams still overlook.

Define your audience

Before discussing the "*how"* of content analytics, let's first focus on the "*why"*—and, more importantly, the "*who."*

As professional technical communicators, we learn early on that understanding our audience is the foundation of high-quality technical content. The same principle applies to content analytics: to build a meaningful measurement strategy, we must first identify *who* we are measuring for and *why*.

A one-size-fits-all approach won't cut it. Different stakeholders have different needs, expectations, and definitions of success. A content strategist might be focused on engagement trends, while an executive wants a clear ROI. Without aligning analytics to these distinct perspectives, we risk gathering interesting but not actionable data.

Let's break down key audiences for content measurement and what they truly need.

Executives

- Needs: Business impact metrics
- Key Metrics: Cost savings from ticket deflection, customer retention rates, content ROI, revenue impact, operational efficiency

Writers

- Needs: Content performance and user engagement
- Key Metrics: Page views, time on page, content quality scores, readability scores, bounce rates, user feedback

Editors

- Needs: Content accuracy, consistency, and quality
- Key Metrics: Content error rates, readability scores, editorial cycle time, content revision history

Content Planners

- Needs: Content coverage, gaps, and alignment with business goals
- Key Metrics: Content reuse rates, content lifecycle tracking, coverage analysis, user search trends

Content Strategists

- Needs: Alignment of content with business and user needs
- Key Metrics: Content engagement, content effectiveness, topic performance, knowledge gaps, customer journey mapping

Information Architects

- Needs: Content structure, navigation, and usability
- Key Metrics: Findability scores, search success rates, content hierarchy efficiency, metadata effectiveness

Writing Team Managers

- Needs: Team productivity and content effectiveness
- Key Metrics: Writer output rates, content quality scores, content cycle time, feedback resolution rates

User Experience (UX) Professionals

- Needs: How content affects user experience
- Key Metrics: User satisfaction scores, heatmaps, usability testing results, content navigation efficiency

Localization Managers

- Needs: Content translation and localization effectiveness
- Key Metrics: Translation turnaround times, localization quality scores, multilingual content engagement, cost per translation

Knowledge Management Professional (such as a Taxonomist and Ontologist)

- Needs: Taxonomy effectiveness and content categorization
- Key Metrics: Search success rates, content tagging accuracy, metadata usage trends, taxonomy adoption

Product Owners

- Needs: Impact of content on product adoption and support reduction
- Key Metrics: Product usage analytics, support case deflection, customer feedback, documentation-driven adoption rates

Support Teams

- Needs: Content effectiveness in reducing support requests
- Key Metrics: Ticket deflection rates, self-service success rates, common support queries

Sales and Marketing Teams

- Needs: How content supports lead generation and customer conversion
- Key Metrics: Content-driven conversions, engagement rates, case study effectiveness, lead nurturing impact

Developers and Engineers

- Needs: Access to accurate and effective API/technical documentation

- Key Metrics: Documentation adoption rates, API call success rates, developer feedback, issue resolution efficiency

Compliance and Legal Teams

- Needs: Content compliance with regulations and policies
- Key Metrics: Compliance audit success rates, legal issue tracking, policy adherence scores

This isn't the only dimension. Do we need to measure value and impact in content production or consumption, and do we need to measure different channels and more? Content shops need more than metrics; they need holistic ***content analytics architecture***.

This *business* of content

But let's not forget what we said earlier—raw metrics alone are as helpful as a GPS without directions. We need analytics that tell us what's happening and why it's happening—making content production and impact measurable, trendable, and, most importantly, *actionable*. Think of it as giving folks a dashboard with all the right buttons, dials, and levers so they can determine what changes to make and measure the effectiveness of those changes afterward. Whether it's cost savings, cost avoidance, faster product adoption, operational efficiency, customer retention, upsells, cross-sells (yes, technical content is secretly a revenue ninja), or even reducing support costs—content plays a critical role.

A business will only adequately fund and value technical content if we can convincingly demonstrate its impact on profits, which is simple math: revenue minus cost.

Of course, we already know the intrinsic value of technical content, but businesses? Again, they're run by actuaries who demand cold, hard data. And since proving that value at scale is often a logistical nightmare, most content teams either don't do it at all or settle for the occasional vague snapshot, relying on anecdotal evidence that screams, "*Trust us, it's important!"* which inevitably is ignored.

I get it. You want to grab the next person who says, "*Everything is measurable*," shake them, and shout: "*Oh yeah? Try running your business WITHOUT technical content and see how well that works out for you!*" Go ahead—make our day.

Change the approach: Develop a content analytics architecture

The key to solving this challenge isn't to dive headfirst into collecting metrics—it's to focus on *why* those metrics matter in the first place. If you want the investment to do it properly and efficiently, you must demonstrate the value of gathering the data, not just the mechanics.

When collecting metrics, there are typically two methods: manually or automatically via an API or similar automation. The goal isn't to pick one over the other; it's to automate as much as possible while strategically augmenting the rest. But where do we start?

We've already established the importance of audience analysis and segmentation. We need a *vehicle*— **a Digital System of Record (SoR)**—to centralize and standardize content data.

This is the point when most content teams hit a wall. The classic mistake is thinking this is solely your problem to solve. **It's not**. Please stop trying to go it alone! You need engineering support from your IT department, content platform engineers, dedicated developers, contractors, or consultants. The biggest hurdle isn't collecting the data; the challenge is securing the resources to make it happen.

Out of the hundreds of content leaders I've informally surveyed, I can count on one hand the number of teams with a digital system of record for their content inventory—let alone one for their content metrics. Most are stuck managing their content data in spreadsheets (if they even have that). It's an industry-wide blind spot, and it's holding us back.

A digital system of record isn't just a nice-to-have—it's the *foundation* of a scalable, sustainable technical content analytics architecture. Without it, you're not measuring content; you're just chasing numbers in the dark.

Building a robust content analytics architecture

A well-designed analytics architecture isn't just a collection of metrics; it's a structured system with a solid foundation, multiple layers, and key pillars that work together, such as the wiring and plumbing of a well-constructed building. A mature technical documentation analytics framework provides deep, actionable insights into documentation usage, engagement, effectiveness, and overall business value. Let's break down what it takes to build one.

Foundation: Core data and infrastructure

The infrastructure that enables data collection, processing, and storage is at the base of any analytics system. This includes:

- **Tracking and Logging Mechanisms** – Capturing user interactions, search queries, page views, feedback submissions, and other behavioral data.
- **Data Storage** – Centralized repositories such as cloud-based data lakes or relational databases to store and manage historical and real-time data.
- **Integration with Documentation Platforms** – APIs, webhooks, and log ingestion pipelines that extract structured and unstructured data from knowledge bases, help centers, and CMSs.

Without a strong foundation, analytics efforts become fragmented, unreliable, and difficult to scale.

Layers: Data processing and analytics stack

A mature analytics system transforms raw data into meaningful insights through multiple processing layers:

1. **Ingestion Layer**—This layer collects and standardizes data from various sources, including user logs, search behavior, and support tickets.

2. **Processing Layer** – Cleans, aggregates, and enriches raw data to create structured datasets.
3. **Storage Layer** – Organizes and optimizes processed data for efficient querying and visualization.
4. **Analytics & Insights Layer**—This layer applies statistical analysis, machine learning models, and visualization tools to extract patterns, trends, and correlations.
5. **Presentation Layer** – Provides user-friendly dashboards, reports, and API endpoints to ensure insights are easily accessible and actionable.

Without these structured layers, organizations risk drowning in data without extracting real value.

Pillars: Key architectural principles

For analytics architecture to be truly effective, it must be built on **three essential pillars**:

- **Scalability** – The system must manage increasing volumes of documentation data and user interactions without performance bottlenecks.
- **Security & Compliance** – Robust access controls, data anonymization, and adherence to privacy regulations (e.g., GDPR, CCPA) are non-negotiable.
- **Actionability** – The ultimate goal isn't just generating reports—it's driving real improvements in **content strategy, search optimization, and user engagement**.

A well-designed content analytics architecture isn't just a luxury; it's necessary for organizations that want to optimize their content, improve user experience, and confidently make data-driven decisions.

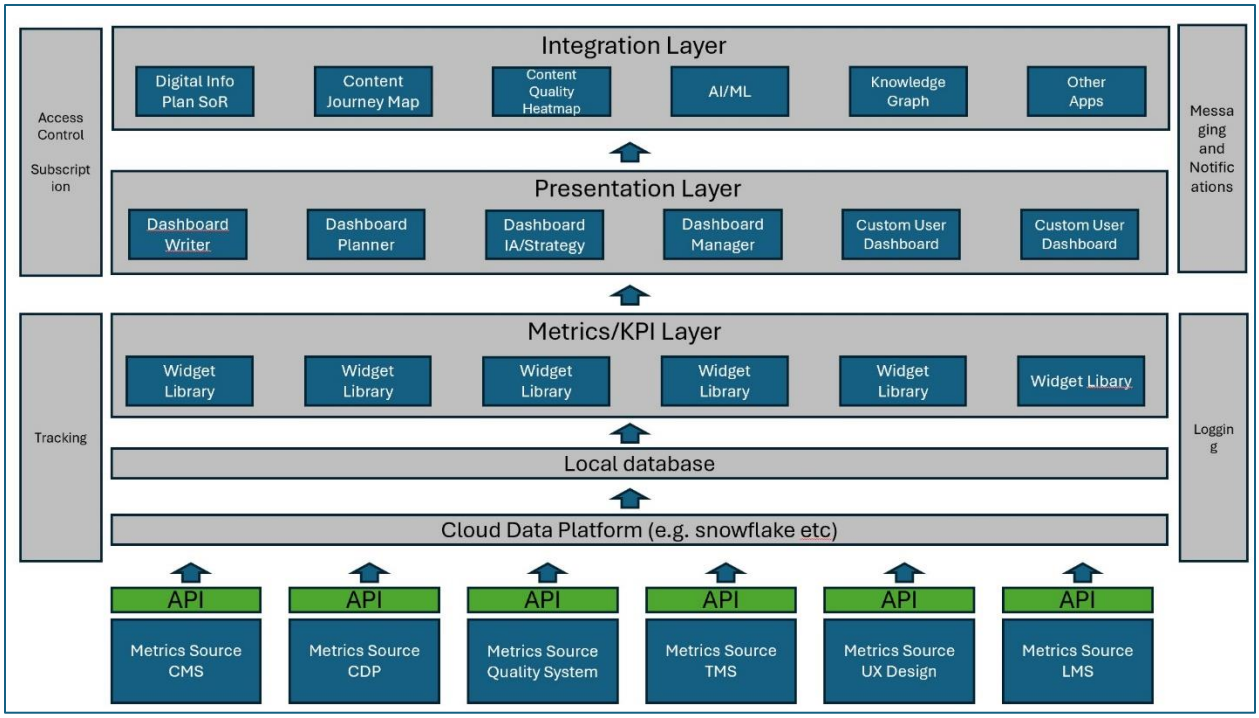

Figure 1 A conceptual architecture

Metrics, KPIs, and baking cakes

Let's organize our analytics strategy for technical content into a framework. Metrics are often confused with KPIs

Metrics are raw data points that measure specific aspects of user interactions, such as page views, bounce rates, and time on page. They provide foundational insights but do not indicate success on their own.

KPIs (Key Performance Indicators) are high-level, goal-oriented measures derived from metrics. They are tied to business or content strategy objectives, such as engagement rate, self-service success rate, and content ROI. While metrics track general activity, KPIs focus on performance relative to goals and drive decision-making. Put another way, all KPIs are comprised of metrics, but not all metrics are KPIs. KPIs reflect what matters, while metrics provide the underlying data and actionable insights.

Just as we need architecture for our analytics system, we need a **content analytics framework** to capture and manage the data and turn it into usable and actionable information,

A content analytics framework for technical documentation

The following provides a structured approach to measuring the effectiveness of technical documentation. It organizes analytics into key categories to help organizations track content performance, user engagement, and overall impact.

Core Data Categories

Metrics represent the raw, quantitative data collected from user interactions. These include fundamental data points such as page views, unique visitors, bounce rate, time spent on a page, and search queries within documentation. We'll not list all the potential metrics here, as many dozens or hundreds exist. Instead, we'll provide a sampling in the appendix.

Dimensions provide contextual attributes that categorize and segment data. Typical dimensions include user type (such as new vs. returning users), content type (articles, FAQs, release notes, API documentation), device type (desktop vs. mobile), and traffic sources (organic search, internal site navigation, direct visits). Geographic and language-based segmentation also fall within this category.

Actionable insights

Key Performance Indicators (KPIs) are high-level **business measures** derived from raw metrics. They provide a goal-oriented perspective on documentation effectiveness. Examples include content engagement rates, conversion rates, self-service success rates, customer satisfaction scores (CSAT), and content ROI.

User behavior insights focus on how users interact with documentation. These insights come from tools such as heatmaps that track scrolling and clicks, session recordings that reveal real user journeys, and path analysis to understand how users navigate between documentation pages. Exit pages provide additional context for where users disengage.

Sentiment and qualitative feedback offer a subjective perspective on documentation quality. This includes user ratings, feedback forms, support ticket sentiment analysis, and survey responses measuring Net Promoter Score (NPS) or usability feedback. These inputs help assess content clarity and effectiveness.

Advanced analysis

Trends and historical patterns reveal how documentation performance evolves. Seasonal trends, post-release traffic spikes, and the rate at which content engagement declines provide deeper insights into content relevance.

Benchmarks and comparative analysis allow organizations to assess their performance against past performance and industry standards. Internal benchmarking compares current engagement against previous periods, while industry benchmarks provide context for how documentation performs relative to similar organizations. A/B testing refines content by comparing formats, structures, and presentation methods.

Anomalies and alerts detect unexpected changes in documentation performance. A sudden drop in traffic could indicate indexing issues, while a surge in search queries for an unresolved topic might suggest missing or insufficient content. An increase in negative feedback on specific pages could highlight content gaps that require immediate attention.

Predictive analytics and AI-driven insights use machine learning models to forecast documentation performance and identify areas for improvement. Predictive models can anticipate which topics

will become outdated, highlight weak spots in content based on engagement trends, and generate AI-based recommendations for content restructuring.

Baking our analytics layer cake

When thinking about metrics and KPIs, it helps to think of metrics as the individual ingredients for a recipe, such as a list of ingredients for baking a cake. It is entirely possible that a single metric can be consumed individually and serve as a KPI (admit it, you've stolen quite a few Nestle chocolate morsels when baking; I've seen you!). However, valuable KPIs are built from a collection of individual metrics that are weighted, correlated, and triangulated. So, think of a KPI as the cake, the metrics as the ingredients, and the weighting, correlating, and triangulating as the baking instructions and techniques.

Materializing the architecture

Search for books on Amazon and other sources about measuring the value and impact of technical content, and you'll find very few, if any, dedicated ones. Such information is usually buried in a precious few books, such as Developing Quality Technical Information: A Handbook for Writers and Editors, but none of them tell you how to implement and operationalize doing so in a feasible manner. Search the web for the same, and you'll find a swath of articles about what to measure, again with little or no guidance on *operationalizing* it other than referencing a few sources for raw metrics.

To put it bluntly, we need a **vehicle**—not just any vehicle, but a virtual assault vehicle—a relentless battering ram to break through the siege of those who have long overlooked, dismissed, ignored, undervalued, and underinvested in technical content. It's time to drive change, shift perceptions, and ensure our work gets the recognition it deserves—not just for ourselves but for the customer and the business's success.

> To put it bluntly, we need a **vehicle**—not just any vehicle, but a virtual assault vehicle—a relentless battering ram to break through the siege of those who have long overlooked, dismissed, ignored, undervalued, and underinvested in technical content.

This is the harsh reality few are willing to face: content shops need to develop tools to make gathering metrics and generating analytics a continuous, feasible, sustainable, and, most importantly, impactful process. Some shops use spreadsheets, but that's not an automatable and scalable method. You must step up to business analytic (BI) tools and technology, which will require development work and resources. If you don't have development resources for data analysis and reporting, visualization tools such as Power BI, Tableau, and Google Data Studio help

transform raw data into actionable insights. However, as shops will eventually discover, using off-the-shelf tools won't get you to the holistic content analytics architecture in the long run – at least not quickly because it takes more to achieve an actionable **360-degree content analytics strategy** and architecture that includes elements such as integrating metric and KPIs **throughout the entire content supply chain** that includes content strategy, information architecture, content creation, content journey mapping, and content inventory management. We'll get to those later.

A digital system of record (SoR)

Just as looking at only metrics is akin to feeling only one part of an elephant in a dark room, so is looking at just the BI tool. A BI tool is part of a much larger animal.

An in-your-face interface

No amount of data will move the donkey if you can't kick it. Said in a politically correct way, data alone won't drive action—you need the proper force to move things forward,

The message must be clear, crisp, concise, and constant. Thus, the primary assault vehicle is a **personalized content analytics dashboard** for each role we listed earlier—especially for decision-makers. The analytics must be highly relevant and of interest to each individual in each role, and the data must be current and continuously updated. It must be a pull-and-push model—push because if Mohammad won't come to the mountain, we'll push the mountain to Mohammad using a subscription service.

That sounds like a daunting and impossible goal, right? It isn't. A handful of content professionals and a competent software engineer can develop such a strategy and architecture over the course of one or two quarters and gradually expand upon it over time.

So, imagine this: You develop an analytics website that your target audience can easily access. Let's call it Content Central. Content Central will eventually become more than an analytics dashboard – an enterprise content operations management center akin to NASA's Mission Control. We'll get to the other elements of Content Central that can be built out later. Let's focus on only analytics with the caveat that the same metrics will eventually permeate the depth and breadth of every function and service of our content operations ecosystem and supply chain. That's why we call this approach an analytics *architecture*.

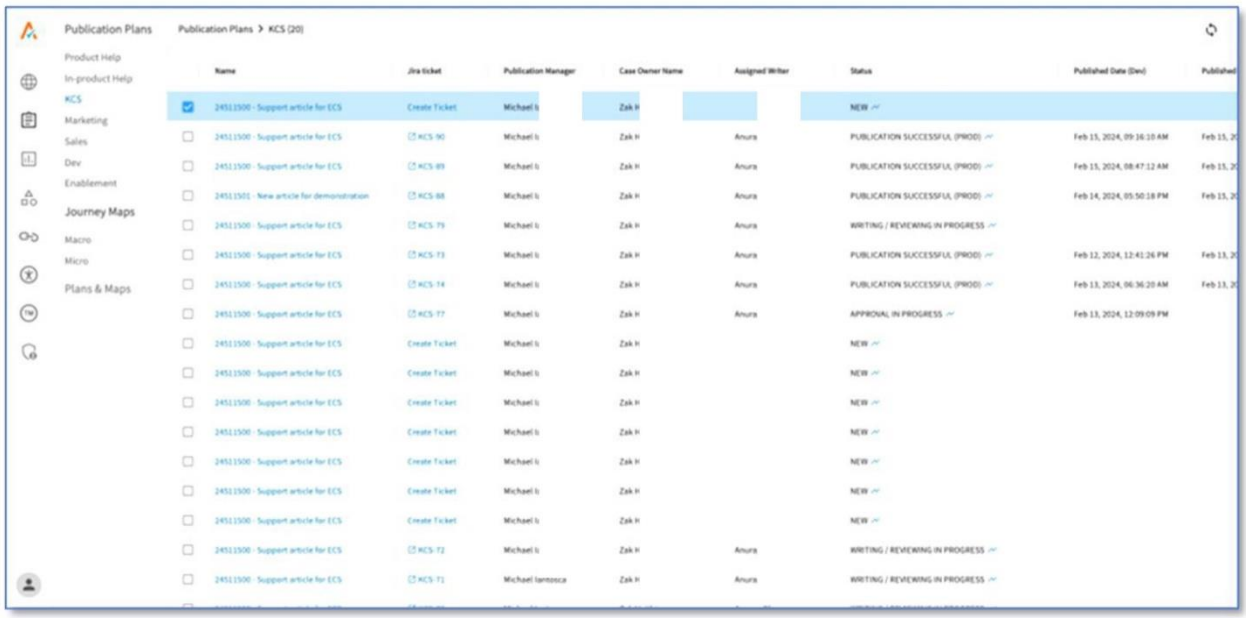

Figure 2 Digital content plans and inventory (SoR)

Plumbing and wiring: The essential connective tissue of an analytics system

In our discussion of analytics architecture, we outlined its core components: foundation, layers, pillars, wiring, and plumbing.

The foundation consists of the data and data model, forming the bedrock upon which everything else is built. The layers include storage and data management, the source systems and APIs that supply raw data, and the widget and KPI libraries that transform that data into usable insights.

But what about wiring and plumbing? These behind-the-scenes systems ensure everything runs smoothly, securely, and efficiently. Without them, even the most powerful analytics engine can become disorganized, ineffective, or underutilized.

Access Control: Managing who sees what

A strong analytics system must ensure that the right users have access while restricting sensitive data to authorized personnel. This often requires integration with an organization's identity management system, such as Okta, Azure AD, or another authentication platform.

By aligning analytics access with existing identity systems, we can:

- Simplify user management by grouping permissions based on role or department.
- Ensure that confidential or high-level data is visible only to those who need it.
- Reduce administrative overhead by eliminating the need for separate login credentials.

Without proper access controls, analytics data can either be too restricted to be useful or too exposed to be secure, both of which undermine the system's value.

Notification service: Keeping users informed without overwhelming them

A well-designed analytics system shouldn't require users to hunt for critical insights constantly. Instead, it should proactively alert them when key thresholds, anomalies, or performance changes occur.

Users should be able to receive notifications via email, dashboards, or in-app alerts when:

- A metric crosses a predefined threshold (e.g., content engagement drops below a target percentage).
- New data that impacts decision-making is available.
- Issues require immediate attention, such as compliance failures or broken links in published content.

Without a proper notification system, users risk getting lost in a sea of data, only realizing too late that a critical issue went unnoticed.

Subscription service: Ensuring analytics are used

One of the biggest frustrations for analytics teams is investing time and effort into creating insightful dashboards and reports—only for no one to use them.

An opt-out subscription service can help push analytics to the right people at the right time with the right message. Instead of expecting users to log in and manually review dashboards, we can:

- Send automated reports to key stakeholders at regular intervals.
- Allow users to subscribe to specific metrics or KPIs relevant to their role.
- Provide customized insights that highlight trends, anomalies, and recommendations.

Most decision-makers love hard data, especially when it helps validate their investments, track performance, and justify future decisions. By making analytics impossible to ignore, a subscription service ensures that data-driven decision-making becomes the norm rather than an afterthought.

The right message, at the right time, for the right user

Plumbing and wiring might not be the most visible parts of an analytics system, but they are critical to its success. Access control ensures security, notifications keep users informed, and subscriptions drive engagement. Together, they transform raw data into a living, breathing analytics experience that delivers real value to the organization.

It’s about the data, stupid

Hijacking a popular phrase about the economy and persuading constituents, we need the data. As we said earlier, we need to automate metrics collection and store them before we can work with them and turn them into valuable and usable information and actionable knowledge.

One option is to acquire a business intelligence (BI) tool or system. Another option is to have your engineering person or team build and maintain a custom data model and database for your metrics and analytic data. The latter approach is much preferred for flexibility, customization, and integration with the many other elements of the content supply chain and in Content Central to achieve the holistic solution we aim for. If you can’t, that’s OK – we’re shooting for the sun, but reaching the moon or even the bar down the block with analytics will be something big to celebrate.

It’s also about the stupid APIs and the clueless content tool vendors

As mentioned earlier, automating metric collection as much as possible is essential. What is the best way to do that? APIs. Lots and lots of APIs. But here’s the kicker—not every system has an API, and even when they do, they rarely offer the analytics you *need*. This is where things start to get… frustrating.

For the gaps where APIs don’t exist (or refuse to be helpful), you’ll need a way to input critical metrics manually. And let’s be honest—if the manual input interface isn’t dead simple, no one will use it. A few fancy automations won't help if your team has to navigate an Excel labyrinth to track basic performance data.

Beyond just collecting data, you’ll also need a data management layer to manage the flood of incoming information. Depending on your needs, which might mean Amazon S3 for raw storage or Snowflake for scalable data warehousing and analytics. These solutions help capture large, unwieldy data sets, manage them efficiently, and integrate them with other tools.

For example, if your CMS dumps a truckload of logs, Snowflake can help organize and query those data mountains instead of forcing you to shovel through them manually. Or, if you’re pulling case deflection metrics from a support system, Snowflake can be a bridge to bring those insights into your broader content performance analysis.

But what about when an API doesn’t surface the data you need? Simple: keep a list of must-have metrics by the system and relentlessly pester your CMS, CDP, TMS, or whatever acronym-laden provider is hoarding your data. If enough of us do this, vendors will eventually cave. Remember, *the squeaky wheel gets the oil*—or, in this case, a collective army of annoyed content professionals can become a giant, city-wrecking centipede of change.

A bit more about those “*clueless content tool vendors*” Harsh? Maybe. Earned? **Absolutely.**

Let me explain. Many CMS and CDP vendors believe their built-in analytics dashboards are *the* solution for analytics. They assume you’ll be perfectly happy looking at just their data, in their way, with zero external correlation. But real-world content analytics requires data from multiple sources, and these walled-off dashboards are about as useful as a submarine with screen doors.

Even worse, when vendors offer API access, they often limit what data you can retrieve—kneecapping your ability to make those metrics useful. And if you think they'll publish a data model to help you interpret what's available? Think again. That kind of transparency is too much to ask, even though hiding a data model in an analytics tool is tantamount to data management malpractice.

Let's talk about real-world consequences:

- Got a CCMS with content reuse? Great! But good luck quantifying cost savings if they won't give you the reuse metrics.
- Using a content quality platform? Fantastic! But what's the impact if they won't provide the number of issues found and fixed—by type, no less?

You'd think these vendors would be *eager* to provide these metrics so customers could prove massive ROI and easily justify renewals or upgrades. Instead, they seem determined to make it as difficult as possible. It's maddening.

So, yes, this process isn't going to be easy. But trust me, it will be worth it. Keep pushing for the data you need, demand better from vendors, and, if necessary, be the squeaky wheel… or the giant centipede of accountability.

Widgets, widget libraries, and personalized analytic dashboards – Oh my!

Now we're cooking! We've got our vehicle (the analytics website) and fuel (raw data), but we're still missing one crucial piece: the engine. That engine is powered by a robust collection of metric widgets organized into meaningful categories to populate analytics pages for different user roles.

Some pages will feature default, pre-built dashboards, giving users an at-a-glance view of key metrics. But we're not stopping there—oh no, my friends! We're going all-in on customization.

Let's empower our users to create personalized analytics dashboards by dragging and dropping the metrics and KPIs they care about most. No more hunting through irrelevant reports or waiting for someone else to build a custom view—just pure, user-driven flexibility.

Because real insights should come to you, let's also roll out a subscription feature while we're at it—the "push" element we mentioned earlier. Users can subscribe to key analytic dashboards and receive automatic updates, ensuring they always have the latest data without constantly logging in.

With this approach, we're not just building an analytics system, we're building an analytics experience that's dynamic, user-friendly, and tailored to individual needs. Buckle up!

Metric Widgets: The building blocks of insight

Before diving into the magic of correlating multiple metrics to create powerful KPIs, let's acknowledge an important truth: many individual metrics are already insightful. Read rates, content publishing frequency, bounce rates, content quality scores, utilization—there's a boatload of them (we'll even provide a sample in the appendix because we're just that generous).

The first step? Turning these raw metrics into interactive widgets.

Figure 3 A sample metrics widget library

For every metric that matters to at least one of our analytics user roles—writers, IAs, content strategists, managers, executives, and beyond—let's generate a corresponding widget and organize them into categorized widget libraries inside the analytics UI. From there, users can drag and drop their preferred widgets onto personalized dashboard pages, rearrange them, and resize each as needed for a custom analytics experience.

Of course, we're not stopping just making the numbers visible. We want these widgets to be powerful and flexible. That means:

- **Configurable Time Frames** – Each widget should allow users to display metrics by day, week, month, quarter, year, or any custom range to visualize trends over time.
- **Target-Based Insights** – Many metrics should allow for target definitions, enabling comparison against goals (e.g., an 80+ content quality score measured by Acrolinx, Congree, or similar tools).
- **Visual Indicators & Alerts** – Widgets should include color coding, trend arrows, or threshold-based alerts to highlight when something needs immediate attention (because no one has time to dig through every number manually).

For example, suppose a team is tracking content quality improvement across multiple collections. In that case, a visual progress tracker should highlight areas of improvement, alert users to

underperforming sections, and track progress over time—all without requiring users to calculate trends manually.

Because data isn't one-size-fits-all, we're ensuring these widgets can display and export data in multiple formats—graphical (charts, trend lines, heatmaps) or tabular (for those who love their spreadsheets).

With this approach, our analytics system won't just *display* metrics—it will bring them to life, making data more actionable, insightful, and, dare we say, actually helpful.

Metric widget libraries

Because there can be dozens or more metrics widgets, we'll classify them for easy access to drag-and-drop to personalized dashboards. Here are some examples of potential configurable classification widget library categories:

- **Content Creation** – Metrics related to content production rates, revision frequency, and workflow efficiency.
- **Content Quality**—This includes readability scores, compliance checks, grammar/style adherence, and an overall content rating.
- **Content Utilization** – Engagement metrics such as page views, time on page, bounce rates, and content reuse statistics.
- **Localization** – Metrics tracking translation accuracy, turnaround time, cost efficiency, and localization QA scores.
- **Content Performance** – A/B testing results, impact of content updates, and effectiveness of different content types.
- **SEO & Findability** – Search ranking, keyword performance, metadata optimization, and internal search success rates.
- **User Behavior & Engagement** – Click-through rates, navigation patterns, interaction heatmaps, and user feedback ratings.
- **Compliance & Governance** – Regulatory adherence tracking, accessibility compliance, and audit report statistics.
- **Cost & Efficiency** – Cost-per-content-piece, savings from content reuse, and workflow automation impact.
- **Knowledge Management** – Metrics related to taxonomy, ontology, knowledge graphs, and other semantic utilization and assets.
- **Personalization & Customization** – Metrics related to dynamic content effectiveness, personalized content usage, and adaptive content engagement.
- **Content Distribution & Syndication** – Content performance across platforms, content shares, and referral traffic.
- **Collaboration & Workflow** – Team productivity, content review turnaround times, and approval process efficiency.
- **Support & Knowledge Base Effectiveness** – Article effectiveness, case deflection rates, self-service success metrics.

- **AI & Automation Insights** – AI-assisted content improvements, auto-tagging accuracy, and AI-powered suggestions used.

Again, these are configurable and can differ from content shop to content shop. They can also be dynamically reorganized by user role. For example:

Writers & Content Creators

- **Content Creation** – Publishing rates, revision frequency, content velocity.
- **Content Quality** – Readability scores, style guide adherence, spelling/grammar errors.
- **SEO & Findability** – Keyword performance, search ranking, metadata optimization.
- **Collaboration & Workflow** – Review turnaround times and approval process efficiency.

Information Architects (IAs)

- **Content Structure & Reuse** – Component reuse rates, modular content effectiveness.
- **Content Performance** – A/B testing results, effectiveness of content templates.
- **User Behavior & Engagement** – Click-through rates, heatmaps, interaction patterns.
- **Localization** – Translation turnaround times, cost efficiency, quality scores.

Content Strategists

- **Content Utilization** – Page views, time on page, bounce rates, content consumption trends.
- **SEO & Findability** – Internal search success, ranking performance, metadata effectiveness.
- **Personalization & Customization** – Engagement with personalized content, dynamic content effectiveness.
- **Content Distribution & Syndication** – Content shares, referral traffic, multi-channel reach.

Content Managers & Leadership

- **Cost & Efficiency** – Content ROI, cost-per-content-piece, savings from content reuse.
- **Compliance & Governance** – Regulatory adherence, accessibility compliance, audit reports.
- **AI & Automation Insights** – AI-assisted content suggestions, tagging accuracy, automation impact.
- **Content Performance** – Content effectiveness by topic, audience reach, and engagement rates.

Executives & Decision-Makers

- **High-Level Content Performance** – Aggregated engagement metrics, business impact of content strategy.

- **ROI & Cost Efficiency** – Budget allocation efficiency, content investment vs. return.
- **Compliance & Governance** – Risk mitigation, legal compliance tracking.
- **AI & Automation** – Strategic value of AI-driven content initiatives.

Localization Managers

- **Localization Efficiency** – Turnaround time, cost, workflow delays.
- **Translation Quality** – QA scores, consistency across languages.
- **Global Content Utilization** – Regional content engagement and performance in different markets.
- **Cost & Efficiency** – Savings from translation memory, automation impact.

Customer Support & Knowledge Base Managers

- **Support Content Effectiveness** – Case deflection rate, helpfulness scores, self-service success.
- **User Behavior & Engagement** – Click rates, search success, time spent on help pages.
- **AI & Automation** – AI-suggested content improvements, chatbot-assisted resolutions.
- **Collaboration & Workflow** – Content updates based on support case trends.

Such a structure aligns widget libraries with user roles and key concerns, making it easier to drag and drop relevant analytics into dashboards.

More than widgets – composite metrics and KPIs

As we said, think of individual metrics as ingredients in a recipe—some, like chocolate chips, are great on their own, while others, like oregano, work best when combined. We're gathering an extensive collection of individual data points and creating independent metric widgets. The real power comes into play: **constructing composite metric widgets** that correlate, weigh, and interconnect individual metrics to form business KPIs that provide meaningful, actionable insights.

By going beyond isolated data points, we create dynamic KPI widgets that allow users to drill down into the underlying metrics, revealing cause-and-effect trends that were once hidden. This approach transforms the analytics experience from passive observation to active, customizable control—giving users the virtual buttons, levers, and switches they need to analyze trends, make data-driven decisions, and optimize performance with unprecedented precision.

We can organize composite business KPI widgets into structured categories as we've categorized individual metric widgets. These KPIs function as recipes, blending multiple data points to provide deeper insights into content performance, engagement, quality, efficiency, and more.

Content consumability

Web Experience Quality

- **Customer Ratings & Feedback**
    - Customer page ratings (exit surveys)
    - Support tickets that result in a doc correction
    - Experience scorecard (custom metric development)
    - Content helpfulness survey responses
- **Engagement & Interaction**
    - Time spent on page
    - Bounce rates
    - Scroll depth (% of page viewed)
    - Click-through rates on recommended content
    - Number of content interactions (e.g., downloads, video views)

App-Embedded Experience Quality

- **Usage & Behavior Analytics**
    - In-app help engagement (views per session)
    - Help widget usage frequency
    - Path analysis (user navigation trends)
    - Repeat visits to the same help topic
- **Retention & Effectiveness**
    - Time spent on in-app content
    - Percentage of users who resolve issues using in-app content
    - Interaction success rates (e.g., tooltips, guided walkthroughs)

Customer satisfaction & success

Customer Satisfaction

- **User Feedback & Sentiment**
    - NPS (Net Promoter Score)
    - Customer ratings on knowledgebase articles
    - Survey responses on content satisfaction
    - Number of problem pages reported by users

- **Issue Resolution & Content Effectiveness**
    - Help content effectiveness rating
    - Percentage of users who find the content useful
    - Time taken to resolve issues using content

Customer Success

- **Conversion & Impact on Business Goals**
    - Successful content-driven conversions (e.g., trial sign-ups, purchases)
    - Increase in self-service adoption rate
    - Reduction in support ticket volume due to improved content
    - Success rate of self-help content (percentage of users who find solutions without needing human support)

External effectiveness

Traffic & Engagement

- **User Interaction with Content**
    - Total traffic to content pages
    - Traffic sources (organic, direct, referral, social, paid)
    - Pages viewed per session
    - Most-used pages vs. least-used pages
- **Content Discovery & Consumption**
    - Zero-hit pages (unfound content gaps)
    - Search ranking performance of content
    - Click-through rate on search results

Social Impact & Reach

- **Social Sharing & Engagement**
    - Number of social shares per content piece
    - Engagement rate on social media (likes, comments, retweets)
    - Referral traffic from social media

Internal efficiency

Content Management Efficiency

- **Inventory & Governance**

    - Total number of pages (by product, update, channel, subject)
    - Number of outdated pages
    - Number of redundant or duplicate pages
    - Percentage of content reviewed within the governance cycle
- **Content Lifecycle Metrics**
    - Average content lifespan before updates
    - Frequency of content updates per quarter
    - Number of archived vs. active pages

Cost Efficiency

- **Financial Impact of Content**
    - Cost per delivered page (creation, design, maintenance)
    - ROI from content reuse
    - Cost savings from automated content processes
    - Localization costs per language

Content quality

Overall Quality Score

- **Content Scoring & Evaluation**
    - Content quality score (overall rating)
    - Editorial review scores
    - Grammar, spelling, and style guide adherence
    - Content clarity score
    - Readability index

SEO & Findability

- **Search & Discoverability Metrics**
    - Organic search impressions
    - Search ranking position for key content pages
    - Click-through rate on search results
    - Percentage of pages with optimized metadata
    - Number of broken links detected

Complexity & readability

Complexity Reduction

- **Simplification & Readability Metrics**
    - Sentence length reduction over revisions
    - Language simplification (average word difficulty score)
    - Percentage of passive vs. active voice usage
    - Average Flesch-Kincaid readability score

Content value & differentiation

Value Differentiation

- **High vs. Low-Value Content Analysis**
    - Ratio of high-value to low-value content pages
    - Percentage of pages classified as ‘critical’ or ‘core’
    - Number of high-performing vs. low-performing content pieces

Content Reuse & Efficiency

- **Content Duplication & Reuse**
    - Number of reused topics and maps
    - Number of reused text objects
    - Number of instances a word is published vs. authored

Integration & multi-channel performance

Integrated Experience Quality

- **Cross-Platform Performance**
    - Engagement consistency across web, mobile, and app
    - Content performance across different platforms (web, PDF, chatbots, etc.)
    - User preferences for content format

Usage Patterns & Analytics

- **Content Consumption Trends**
    - Percentage of users who consume content across multiple platforms
    - Impact of embedded content on user retention

This detailed KPI framework provides specific, measurable metrics that can be turned into widgets for content analytics dashboards.

Our goal here is to unlock business value through deeper analytics. This structured approach to business KPIs is just the beginning. Combining and analyzing content data at scale allows organizations to go beyond basic performance tracking and unlock new strategic insights. For instance, what if we could finally prove what many content teams have known all along - that post-sales technical content plays a massive role in closing deals? Imagine the look on a marketer's face when they realize buyers spend more time reading documentation than their carefully crafted sales materials. Or the surprise from leadership when they learn that over 50% of technical evaluations in B2B sales involve documentation, the same documentation that rarely gets a seat at the revenue discussion table.

We can continue refining and expanding these KPIs to finally give content teams the credit (and budget) they deserve. By providing accurate, data-backed insights into how content impacts the sales funnel, cross-sells, and customer retention, we shift the narrative from "nice-to-have" to critical business assets. The tools are here; it's time to ensure the right people start paying attention.

Calculating ROI

There are few commercial and public tools to calculate or project return on investment. One nifty online tool for estimating ROI for reuse and localization is here:

https://www.scriptorium.com/2024/02/estimate-your-roi-for-content-operations-with-our-calculator/

While such tools might not reveal their underlying algorithms and assumptions, they are often based on years and volumes of sample data. Still, it is best to extract raw reuse metrics from the source content repository, break the instances of reuse down by component type (collections, sub-collections, topics, chunks, and phrase-level element reuse), and apply an average cost savings per reuse instance based on the time it would take to create and maintain duplicate instances for each component type for actual savings.

The same applies to ROI for tools. For example, implementing a productivity tool for writers can save substantial costs by reducing time spent on writing, research, and editing while improving overall efficiency. The total savings can be calculated using a formula that factors in the burden rate (employee cost per hour), time saved per task, usage frequency, and the cost of implementing the tool. The burden rate includes salary, benefits, and overhead, while time savings are determined by comparing pre-and post-implementation task durations. Usage frequency measures how often the tool is used, and the tool cost accounts for subscription fees, training, and integration.

A general formula for cost savings might be:

$$\text{Savings} = (\text{Burden Rate} \times \text{Time Saved per Use} \times \text{Usage Frequency}) - \text{Tool Cost}$$

Where:

- **Burden Rate** = Employee's total cost per hour (salary + benefits + overhead).
- **Time Saved per Use** = Reduction in time-per-task after implementing the tool.

- **Usage Frequency** = How often the tool is used per writer per time period.
- **Tool Cost** = The total cost of acquiring, implementing, and maintaining the tool.

For example, if a tool reduces writing time by one hour per article and a writer produces five hundred articles per year, the savings can be significant when multiplied by the burden rate. Similarly, automating editing and proofreading tasks can save additional time, reduce manual effort, and improve turnaround. AI-assisted research tools can also enhance efficiency by streamlining information gathering and cutting research time by a considerable percentage.

To maximize cost savings, organizations should ensure high adoption rates, track time savings before and after implementation, scale usage across teams, and consider hidden costs such as training and transition periods. Periodic reviews should be conducted to assess the tool's effectiveness and optimize its impact. By structuring the cost analysis and optimizing the adoption strategy, companies can achieve significant returns on investment and enhanced productivity.

A few considerations to keep in mind when calculating ROI:

- **Adoption Rate Matters** – Savings are minimal if employees don't use the tool.
- **Measure Before and After** – Track old and new time spent on tasks.
- **Scale with Team Size** – The larger the writing team, the greater the potential savings.
- **Consider Hidden Costs** – Training, support, and transition time should be included in cost analysis.
- **Periodic Review** – Assess effectiveness every 6–12 months to ensure continued savings.

A 360-degree view: the holistic picture

We discussed early on that this technical content analytics strategy is a holistic architecture. Let's bring the architecture full circle.

You might recall that named application framework Content Central. The notion of Content Central is that of a content operations management center. It's not the same as the content management subsystem used for creating and managing source content, nor the content delivery system for delivering content to the various channels; it is instead a center for shared content business management and shared services of which our content analytics subsystem can be the only application or one of many that can include digital content inventory management, digital content plans, content use journey maps, localization request and tracking, and more.

We won't cover all the possible content operations management applications and services a framework such as Content Central can host. See the following article if you'd like a broader view: https://img1.wsimg.com/blobby/go/8d83b4d9-fe3b-41c3-ba74-a645effc4bf2/downloads/6d4f2344-8478-49f7-9377-c28e17893e27/Houston%20we%20have%20a%20problem.pdf?ver=1734976175557

Aggregating and reusing content analytics – sample use cases

We want to make content analytics more actionable by driving them back into our day-to-day processes. Let's examine a few examples – the possibilities are endless:

Actual savings from content reuse

Many content teams have adopted automated content reuse systems using component-based models like DITA or by extending Markdown with YAML, Jekyll, and similar techniques. These models allow content to be reused by reference on a large scale, leading to significant cost savings. Organizations can save millions of dollars, even at a mid-sized enterprise level, by reducing the need to scale the number of writers in proportion to product and content growth.

However, few teams can demonstrate these savings to prove their impact on the business and bottom line. Even fewer can track content growth and reuse trends to justify additional investments in people and technology when needed.

A well-designed content reuse savings metric can be one of the most compelling KPIs for decision-makers. By capturing the number of reused content components—such as collections, subcollections, topics, subsections, and individual elements—and assigning a cost value, it becomes possible to calculate actual savings. This can be done by multiplying the number of reused components by the estimated burden rate for creating and updating content.

Additionally, teams can forecast when to expand resources, including additional staff and technology by correlating the new and revised content rate with content production velocity. This data-driven approach proves the value of content reuse and helps organizations and executives make informed decisions about scaling their content operations efficiently.

Savings from assistive content quality tools

Many content teams have adopted advanced AI-powered content quality and governance tools like Acrolinx, Congree, and similar systems. These tools function as highly sophisticated technical content editors, incorporating hundreds of AI-driven style and quality governance rules, terminology enforcement, and consistency checks across large content libraries and multiple writers. Unlike general large language models, these specialized tools are explicitly designed for structured content quality management. Many have also begun integrating generative AI capabilities, combining rule-based AI with content generation to increase efficiency.

These systems analyze thousands of editorial elements, providing writers with automated suggestions that can be quickly reviewed, applied, and measured for impact. The time savings per writer are substantial. A leading technology company conducted a formal study and found that these tools saved thousands of dollars per writer annually in editorial review costs alone. The cost avoidance factor was even more remarkable; the expense of manually aligning, fixing, and optimizing content at scale is often prohibitive. The savings become significant when multiplied across dozens or even hundreds of writers.

Calculating these savings is straightforward. By sampling and quantifying the time required for a writer to identify and fix different editorial issues—such as terminology enforcement, clarity improvements, style guide adherence, translation efficiency, and grammar corrections, teams can estimate the total time saved. Applying the **employee burden rate** (the fully loaded cost of a writer, including salary, benefits, and overhead) to the number of issues flagged by these tools provides an estimate of **potential savings**. If teams can track the number of resolved problems versus those flagged, they can calculate the **savings**.

However, this requires access to the right metrics, which some tool vendors may not readily provide. If these data points aren't available, it may be necessary to push vendors to expose key analytics—without them, proving the full value of these tools remains an uphill battle.

Content quality scoreboards for writes and content scrums

One of content analytics' most influential and practical applications is using content quality heat maps to help writing teams improve their work. Suppose your organization uses a content quality governance system like Acrolinx, Congree, or similar tools. In that case, these platforms can batch-analyze content and generate a range of content quality metrics, including composite quality scores for entire document collections and individual topics or articles.

A highly actionable approach to content analytics involves running automated batch checks on content to generate initial quality scores, storing those metrics in a centralized database, and then visualizing the results in an interactive heat map. This gives writing teams clear, data-driven insights into content quality, making it easy to identify strengths, weaknesses, and areas for improvement.

For example, imagine a user guide containing one hundred topics. A dashboard could display a grid or checkerboard-style visualization, where each square represents a topic, showing its title and composite content quality score. Clicking on a square could reveal detailed sub-scores for factors like terminology, grammar, readability, and style, helping writers pinpoint areas that need improvement.

To make this system even more intuitive, the heat map could use color coding:

- Red for topics that fall significantly below the target quality threshold.
- Yellow for those that are close but still need improvement.
- Green for topics that meet or exceed the quality standard.

Of course, we'll also want to ensure the interface is designed for accessibility.

The UI design could also include navigation features like a sortable table of contents, allowing users to filter or scroll through extensive content collections based on quality scores. Additionally, writers could have direct links to quality scorecards within the content governance system for deeper insights.

This approach goes far beyond providing high-level KPI widgets. It delivers actionable intelligence that empowers writers to improve their work efficiently, resulting in higher-quality content, improved business outcomes, and, most importantly, better customer experience.

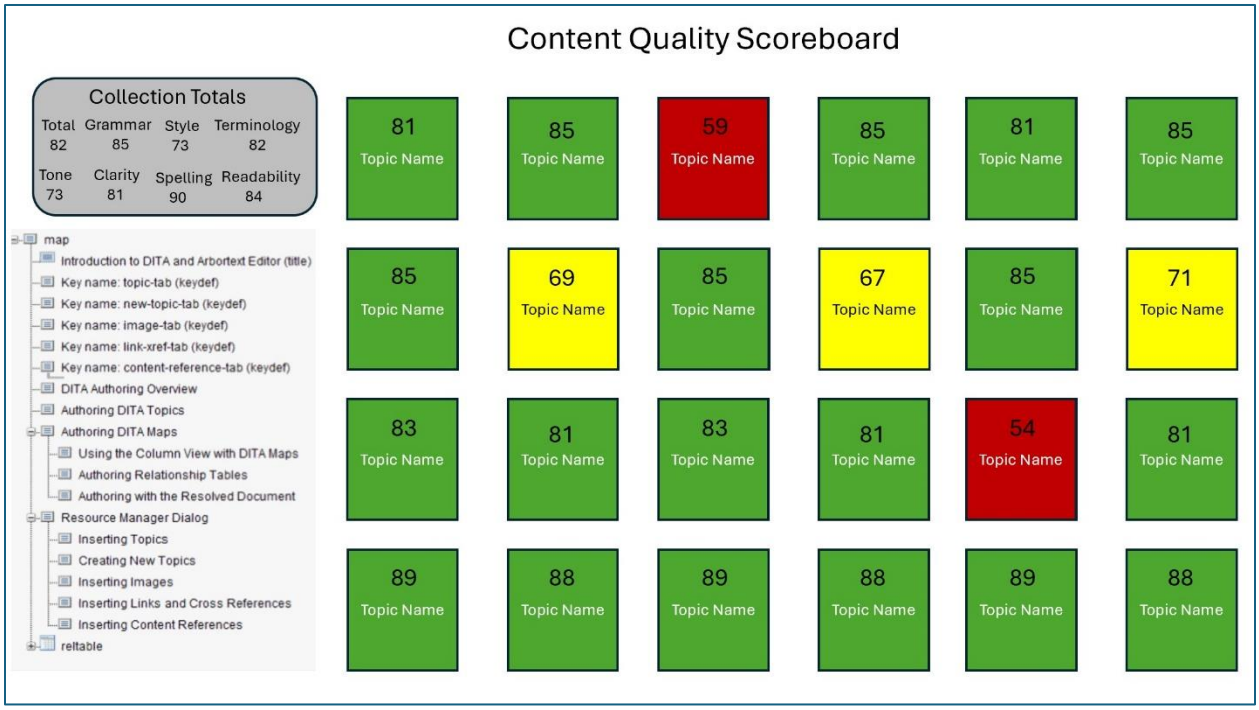

Figure 4 Content quality scoreboard for writers

Effectiveness of AI-assisted KCS technical support article generation

Many companies implement a support content program called Knowledge-Centered Support (KCS). Unlike primary product documentation, KCS articles are created in response to resolved technical support issues. Some originate from application defects, while others address highly repetitive inquiries, such as standard procedures that generate frequent support calls or tickets—costly for both the business and the customer. In some cases, these articles fill gaps left by deficiencies in the primary product documentation. Regardless of the cause, KCS articles are designed to help support agents quickly resolve common issues and, ideally, empower customers to find answers on their own. A well-run KCS program reduces support costs, improves efficiency, and enhances customer satisfaction, with potential savings reaching millions of dollars annually in support agent expenses alone.

Some organizations take this further by intercepting support tickets as they are being created and automatically suggesting relevant KCS articles to customers, preventing unnecessary ticket submissions. The same content can be used to train AI chatbots, further enhancing self-service capabilities. Every successfully deflected ticket can be quantified in terms of utilization and direct cost savings.

One particularly advanced organization leverages AI to generate draft KCS articles automatically from completed support cases. Subject matter experts review, refine, and validate these AI-generated drafts before publication. The company even tracks how much AI-generated content is modified, using the data to fine-tune and improve the AI's performance over time—slick! It's an incredibly efficient system—one that not only accelerates KCS article creation but also continuously improves its accuracy and effectiveness.

Content inventory digital system of record- operations automation

Managing your content inventory and plans in a digital system of record, such as the referenced Content Central model, allows you to seamlessly integrate key metrics into the metadata of each publication record.

For instance, if one of your metrics includes the results of an automated accessibility or trademark compliance check or an overall content quality score from an analysis tool, these results can be automatically stored in the metadata fields of each record. This creates a structured, trackable system where quality and compliance are built into the content lifecycle.

Beyond just tracking, these records can be used for gating publishing. You can define pass/fail criteria for accessibility and trademark compliance or set a minimum content quality score threshold that content must meet before it can be published. If a piece of content falls short, publication can be automatically blocked until it meets the required standards—unless a content planner explicitly overrides the restriction.

Integrating a digital system of record with automated checks and quality gates ensures that content meets high standards before it reaches your audience. This reduces risk, improves consistency, and makes quality control a proactive process rather than a last-minute fix.

The right content at the right time: Content lifecycle journey maps

Content lifecycle journey maps are powerful tools in some content organizations to visualize how content supports users across different journey stages. These maps track publications across key content lifecycle phases, such as Discover, Learn, Try, Buy, Use, Support, and Advocate. They can be presented in tabular formats or more visually engaging styles, such as a metro map, where each content type represents a stop along the user's journey.

The primary goal of these maps is to manage content strategy from the user's perspective. The more comprehensive they are—incorporating content from product, support, developer relations, partner networks, marketing, and beyond—the more accurately they reflect the actual customer content experience.

Content lifecycle maps help content strategists and planners identify critical gaps, eliminate redundancies, and make informed decisions about which publications should be:

- Updated to stay relevant for a product release
- Retired or unpublished to reduce clutter and outdated information
- Monitored for entitlement restrictions to ensure access aligns with user permissions

The value can go even further for teams that create digital versions of these maps. Making them interactive allows analytics to be embedded into each node, displaying real-time insights through pop-ups or overlays. This integration transforms static maps into actionable intelligence dashboards, providing content teams with immediate visibility into content performance, engagement, and effectiveness at every user journey stage.

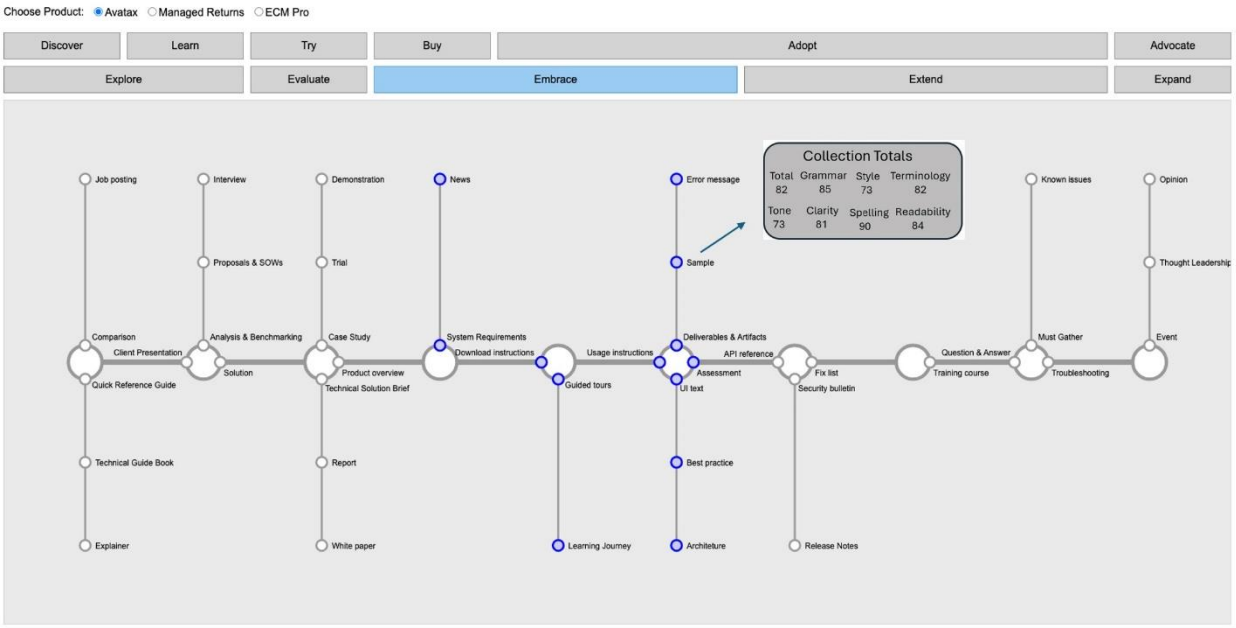

Figure 5 Interactive content usage journey map with pop-up metrics

The future: AI-assisted analytics and knowledge graphs

ChatGPT and other LLMs

Don't shortchange current LLMs to help with content analytics. One well-designed prompt can generate a wealth of information that would take hours or days to gather, if ever. Those who focus intensely on content analytics are often asked – *What metrics do you use, and what do you measure*? That's quite an open-ended question. We can't be experts on every facet of content! Instead, leverage the power of letting a generative engine spark ideas based on a wealth of existing information. For example,

Prompt: *How can I best demonstrate the business value and effectiveness of localizing our technical content to business leaders and highlight the efficiencies and value our content team brings to the process? What metrics can I consider, and what KPIs might interest executives most?*

ChatGPT's answer *(not a bad starting point!):*

It's best to tie localization efforts directly to business outcomes when making your case. Here are some categories and specific metrics to consider:

1. Traffic & Engagement Metrics

- ***Page Views & Unique Visitors:*** *Monitor the traffic of localized pages compared to non-localized ones, segmented by language or region.*
- ***Time on Page & Bounce Rate:*** *Longer time on site and lower bounce rates can indicate that localized content resonates with the audience.*

- ***Engagement Rate:*** *Track actions like downloads, video plays, or clicks that show active user engagement.*

2. Conversion & Revenue Metrics

- ***Conversion Rates:*** *Measure how often visitors to localized pages complete a desired action (sign-ups, inquiries, purchases).*
- ***Revenue Attribution:*** *Quantify revenue generated from leads or sales from localized content.*
- ***Return on Investment (ROI):*** *Compare the revenue uplift against localization costs to highlight cost-effectiveness.*

3. SEO & Organic Search Performance

- ***Keyword Rankings:*** *Track improvements in search rankings for region-specific keywords.*
- ***Organic Traffic:*** *Measure the growth in search-driven traffic from localized markets.*
- ***Click-Through Rates (CTRs):*** *Analyze how well your localized meta descriptions and titles attract clicks.*

4. Customer Experience & Support

- ***Customer Satisfaction (CSAT) & Net Promoter Score (NPS):*** *Use surveys to gauge satisfaction among users engaging with localized content.*
- ***Reduction in Support Inquiries:*** *Fewer queries or support tickets may indicate that localized content is clearer and more helpful.*

5. Operational & Process Efficiency

- ***Time-to-Market:*** *Track the turnaround time for producing and localizing content. Faster cycles can be a decisive operational win.*
- ***Cost Metrics:*** *Measure cost per localized piece or cost per word, highlighting efficiencies from translation memory reuse and streamlined workflows.*
- ***Error Rates:*** *Monitor for consistency and quality, which can reduce post-publication revisions and improve brand reputation.*

KPIs of Greatest Interest to Executives

- ***Revenue Impact:*** *Executives will look for direct links between localized content and revenue increases or lead generation.*
- ***Market Penetration:*** *Growth in new user segments or increased engagement from target regions.*
- ***Content ROI:*** *provides clear evidence that investments in localization yield measurable returns, whether through improved conversion rates or reduced support costs.*

- ***Efficiency Gains:*** *Metrics demonstrating improved turnaround times and cost savings in content production*.
- ***Brand Strength:*** *Improvements in customer satisfaction scores and social sentiment that drive long-term brand value*.

By presenting these metrics and KPIs in a clear, data-backed format, you can effectively demonstrate the quantitative and qualitative business value of your content team's localization efforts.

It's not cheating – you'd not think twice to do the same (painfully) using Google. Armed with this kind of detail, you can determine the available data, choose which metrics widgets are feasible, and create the recipe(s) for KPIs from correlated and weighted metrics.

Knowledge Graphs

We can't know the unknown—**or can we?**

Since we've already defined, categorized, and structured our metrics, we have the perfect foundation to transform them into a taxonomy and schema for an ontology. We can take the next step: automatically generating a knowledge graph by ingesting this ontology and metric data into a graph database. But why go through all this effort? An analytics knowledge graph can query relationships between data points, apply graph reasoning and inferencing, and uncover insights that would otherwise be too resource-intensive to extract manually.

This isn't just theoretical; it's entirely achievable today. By mapping analytics results against a graph of your content corpus, AI can identify patterns, detect correlations, and even predict optimization opportunities that human analysts might never catch. Instead of just analyzing what we already know, we can begin to discover what we didn't even think of looking for. For an example of how a knowledge graph can be automatically constructed from componentized content in a working model, see https://medium.com/@nc_mike/document-object-model-graph-rag-af8ae452b0b6. AI-assisted analytics and knowledge graphs aren't some far-off futuristic concept—they're a practical, next-generation approach to making content insights more intelligent, deeper, and actionable than ever before.

Summary

As we bring this discussion full circle, one thing is clear—content analytics isn't only about crunching numbers or tracking engagement. It's about transforming raw data into actionable intelligence that drives efficiency, improves quality, and delivers better content experiences. From cost-saving automation to AI-powered insights, we've seen how the right analytics strategy can do more than justify investment; it can shape the future of content operations. The possibilities are limitless, whether you're optimizing reuse, fine-tuning quality, streamlining support, or uncovering patterns through knowledge graphs. And while no single solution fits all, one truth remains—those who harness the power of content analytics today will lead the industry tomorrow. So, buckle up. The journey has only just begun.

Appendix: Content performance metrics

This appendix provides a sample list of content performance metrics for measuring content quality, engagement, efficiency, and business impact. The following metrics are categorized into various performance areas and structured in a multi-column format for ease of reference. It is not a comprehensive list by any means, but it is a start. The metrics available to your shop depend on the sources and their extraction methods.

User engagement

Category	Metric
User Engagement	Customer page ratings
	Exit survey responses
	Time spent on page
	Bounce rates
	Scroll depth
	Click-through rates
	Number of content interactions
	Content shares
	Social media engagement
	Referral traffic from social media

Search & discoverability

Category	Metric
Search & Discoverability	Search ranking
	Search results impressions
	Most used pages
	Least used pages
	Zero-hit pages
	Internal search success rate
	The click-through rate on search results
	SEO prominence scores
	Percentage of pages with optimized metadata
	Number of broken links

Content quality

Category	Metric
Content Quality	Overall content quality score
	Grammar scores
	Style adherence scores

	Clarity scores
	Terminology scores
	Readability index
	Language simplification metrics
	Percentage of passive vs. active voice usage
	Editorial review scores
	Content accuracy

Content effectiveness

Category	Metric
Content Effectiveness	Number of help sessions
	Support tickets that result in a doc correction
	NPS score
	Problem management tickets
	Conversions attributed to content
	Customer success rate via documentation
	Case deflection rate
	Self-service adoption rate
	Resolution time improvement
	Impact of content on sales funnel

Content utilization

Category	Metric
Content Utilization	Total content views
	Page views per session
	Pages per visit
	Content consumption trends
	Traffic sources (organic, direct, referral, social, paid)
	Referential links
	Traffic generated to other sites
	Multi-channel content access rate
	User preferences for content format
	Percentage of users consuming content across multiple platforms

Content governance & maintenance

Category	Metric
Content Governance & Maintenance	Total number of pages
	Number of outdated pages
	Number of redundant or duplicate pages
	Percentage of content reviewed within the governance cycle

	Content lifecycle duration
	The average time before content is updated
	Archival rate vs. active content
	Complexity reduction metrics
	Reduction in sentence length over revisions
	The number of compliance issues flagged

Content cost & efficiency

Category	Metric
Content Cost & Efficiency	Cost per delivered page
	Cost to design
	Cost to author
	Cost to maintain
	Localization costs per language
	ROI from content reuse
	Savings from automated content processes
	High-value vs. low-value content ratio
	Efficiency of content production workflows
	Cost per content interaction

Content reuse & modularity

Category	Metric
Content Reuse & Modularity	Number of reused topics and maps
	Number of reused text objects
	Number of modular content components
	Instances a word is published vs. authored
	Percentage of content available in reusable format
	Efficiency of structured authoring
	Reduction in duplicated content
	Improvement in translation efficiency
	Time savings from reuse strategies
	Impact of reuse on localization costs