

WHITE PAPER SERIES

Next-Generation Content Analytics



Introduction

Pioneering computer scientist Grace Hopper once remarked that the most dangerous phrase is, “We’ve always done it this way.”

Tech doc analytics is one area wherein doing things the way they’ve always been done is not only common but costly. Next-generation content analytics platforms offer the sorts of advanced analytics that tech doc teams need, but also have the capacity to power transformation right across the business, from content delivery to support to product design.

Let’s embark on a journey through the Analytics timeline, explore the boundaries of the methods and technologies currently in use, and uncover fresh insights of Next-Gen analytics and their game-changing potential.

NEXT-GENERATION CONTENT ANALYTICS

Table of contents

How, What and Why We Measure	3
Reading the Numbers	7
- Numbers in context.....	7
- Data visualization.....	7
- Taking action.....	8
From Tracking Stats to Transforming Business	9
- Better documentation.....	9
- Better products.....	9
- Personalization.....	10
- Informed customer support.....	11
- Predictive customer support.....	11
Conclusion	12

How, What, and Why We Measure

As long as we've had technology and technologically sophisticated products, we've had technical documentation to support users. Whether delivered as a spiral-bound manual, a QR code linked to a PDF, an app on a smartphone, or HTML pages on a website, the core focus of technical documentation remains the same.

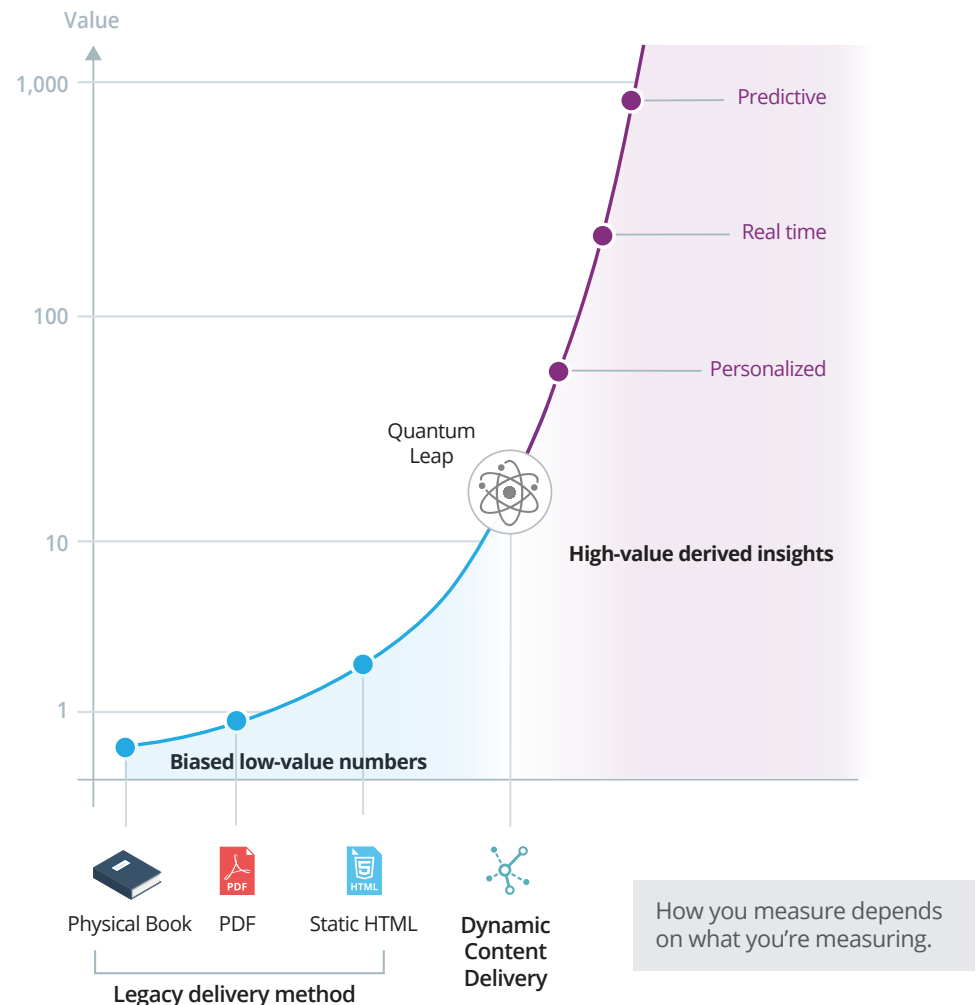
The available analytics, though, change significantly with the way content is delivered. What is measured and what is possible to measure is intrinsically bound to the format and delivery method of the technical document.

Physical guidebooks

Supplied with the product it supports, a hard-copy guidebook is consumed entirely offline and the manufacturer and retailer have no knowledge about how much of the book is read, or if it even is consulted at all. The sole metric is the 1:1 ratio of guidebooks to products, a statistic with little utility.

PDFs

With PDFs there is a dematerialization or digitization of the content in the physical guidebook. Manuals are converted into digital files, and posted to a website. You might imagine that the same guidebook delivered as a PDF might offer more analytical insights because, after all, counting total downloads of a PDF is trivial and, with some basic web analytics, you might also find out where those downloads are coming from. Yet download counts don't tell you much more than that someone clicked a link. There is no way to know if the PDF was read, how much of it was read, what part of the document was useful, or if it was shared with another person, a team, or an entire company. Was the answer found immediately and the file discarded? Did the document become a critical resource, a daily tool in someone's workflow? On questions like these, the analytics fall short.



How, What, and Why We Measure

HTML web pages

In terms of analytics, HTML pages are a stark improvement on PDFs.

First, you know how many times a page has been loaded, as well as when and how long it took before it was clicked away from. If that time is long enough, it has probably been read. And, since the manual is probably not on a single web page, you have a somewhat finer grain of knowledge. Instead of just being aware that a PDF has been downloaded, you find out that the HTML page containing Chapter 5 of the manual has been loaded into someone's browser. This is already an improvement as you could then report to management that Chapter 5 has received some attention.

Now, if you could tell where on the page the customer was looking, that would be even more informative. Sadly, you (probably) don't own the browser, so you can't get that kind of detailed feedback. If you broke the chapter into many small web pages, you would learn with more precision what parts of the chapter have been read, but it would be a pain for your users: they would spend their time clicking Next and Back buttons to navigate your content. **So you are forced to choose: either provide a superior user experience or try and gain more fine-grained metrics.**

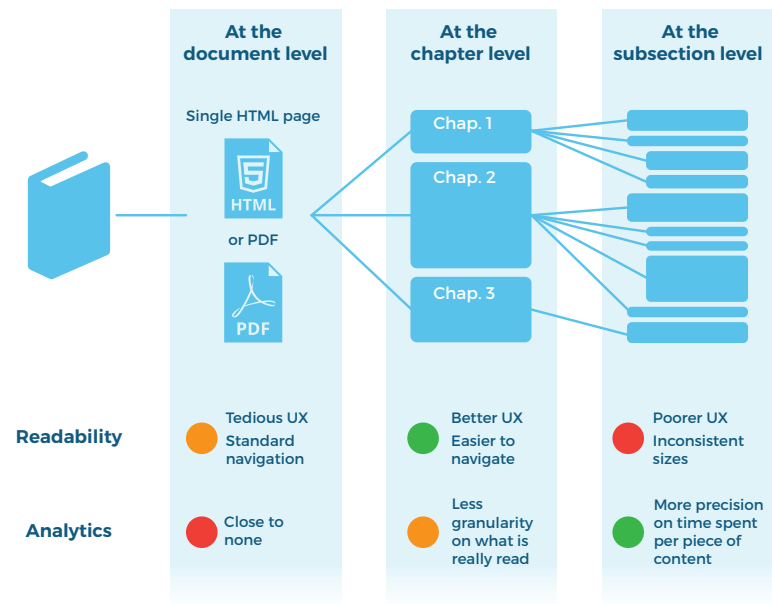
Even if more fine-grained metrics are available, it is important to be clear on what those metrics are. Web pages have been around for decades, and many people conclude they have access to all of the most useful analytics on them via **web analytics tools**, whether the market-leading Google Analytics or a competitor. As an analytics platform user has access to a lot of data, they may think that they know something useful. Yet often this is not the case and, instead, the user is suffering from what is known as **the 'data delusion'**.

At their core, web analytics products are designed for marketing. As the Google Analytics site itself explains,

"Google Analytics lets you measure your advertising ROI as well as track your Flash, video, and social networking sites and applications".

For a company website of some 20 to 200 pages, this is very useful. Users can generate a report with all of their URLs, the number of times each URL has been loaded, how users arrived on a page, where they arrived from, and where they exited. If you have a small number of pages, with a lot of viewers for each, and you want to finely focus your marketing spend, this is very useful information on which to base decisions.

Unfortunately, however, the typical tech doc portal has hundreds or even thousands of individual pages. With the various languages and versions of a product range, the number of HTML pages quickly skyrockets. That's without mentioning that many of these pages may only support a small number of readers.



Long web pages provide lots of info on a 'paper like' experience, but fewer details. Smaller pages provide more fine-grained metrics, but a weaker UX.

How, What, and Why We Measure

In such cases, the Google Analytics report generated will be a spreadsheet with thousands of items with impossible-to-decipher names and a wide range of numbers next to each name. It is all but impossible to make sense of these numbers without viewing the content at the same time. How, for example, can you make sense of a flow diagram showing the path users follow when thousands of small elements must be displayed? In addition, if two pages are related to the same task for the same product, but in two different versions, we need to see the metrics either separately or aggregated; but since web analytics tools don't understand the structure and semantics of tech doc, it's impossible to automatically create and navigate clustered numbers (for example, by book, by version, or by product).

Not to mention that changing the granularity of an HTML page (such as from topic to section) would render any comparison impossible.

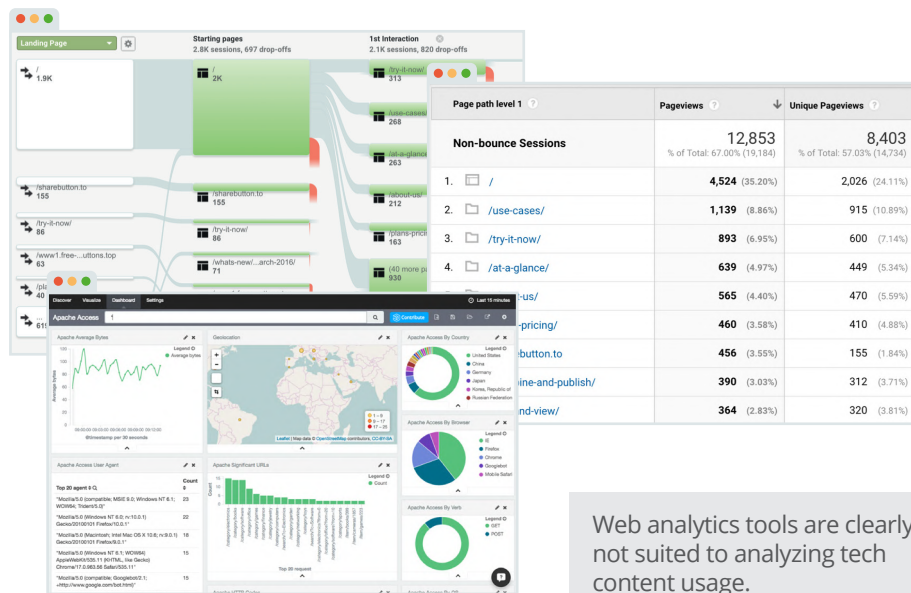
The tech doc decisions that can be made based on web analytics alone are limited. Metrics need context to be interpreted; it is important to review the numbers and the content simultaneously in order to gain insight. In addition, web analytics may reveal what has been read, but they cannot reveal what has not been read. Knowing what's happening at the web server level is not very useful and, for tech doc teams, renders marketing-focused web analytics platforms only slightly better than useless.

Counting the number of guidebooks sold is not useful, counting downloads reveals little to help make better tech doc decisions, and counting page loads is only slightly more helpful. The important information exists at the user level, something for which a different approach to tracking and logging is required.

What we need

In a perfect world, we would be able to stand over our users' shoulders and see what they read and for how long to get insight on our technical documentation. Of course, this might be a little difficult in reality... But it is exactly the level of detail modern analytics strive to offer.

In this data-driven era, enterprises can no longer ignore the opportunity of the wealth of information provided by users. Social media, online retail, mass surveillance, or census forms, there are now torrents of data available. Algorithms have thus been developed to process this data and produce insights. By having **technical documentation generate data** on its utility to users and applying advanced big data mining technologies, we can begin to truly understand how users read tech doc. What's more, you can **learn** not only more about the usefulness of our tech docs, but about **how users interface with products**.



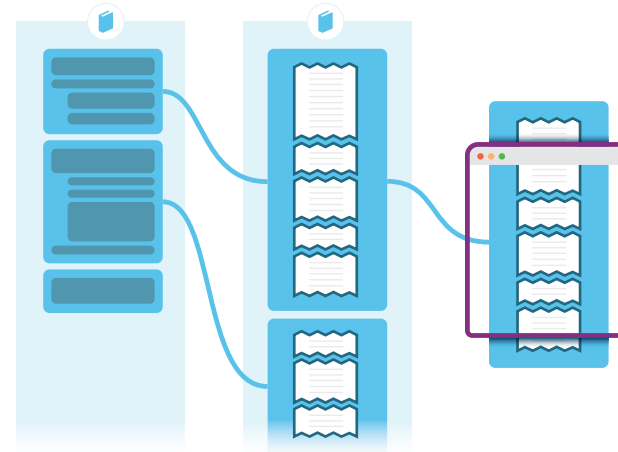
How, What, and Why We Measure

With modern tech doc analytics, it is almost as if a single topic in a manual can send a message back to the business saying, **“Hey, I’m being read now!”** However, for this to be possible, two things are required.

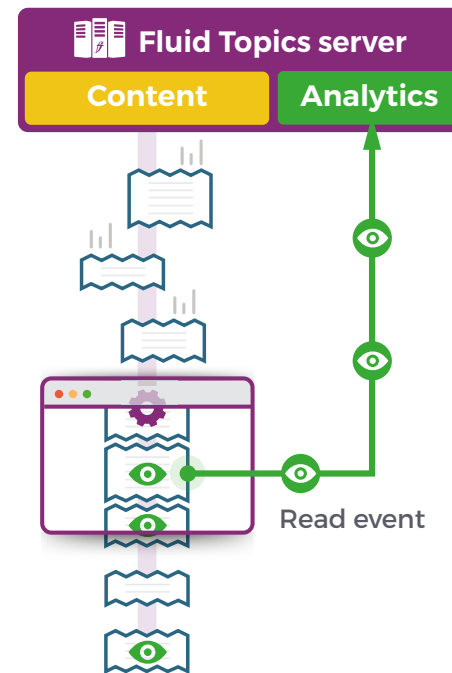
First, the content needs to be granular. Granular content allows you to track the consumption of each fragment individually. Structured documentation (such as DITA) helps with this, but otherwise it is the job of the dynamic delivery platform to break the content into chunks, no matter the source format. The platform must take differently-formatted inputs and present them in a consistent and uniform manner. Indeed, everything that could be considered documentation – wikis, knowledge bases, user forums, trouble tickets, catalogs, as well as product manuals and guides – should be searchable through the same portal. All of this product information can contribute to customer success, thus it should be tracked and analyzed with the same tool and at the same level of detail.

Second, a custom-designed reading and tracking technology needs to be deployed. While you can’t control the browser (you don’t own it), you can have customers view the documentation through a web-based reader that you do control. This will send these messages back when it displays a fragment for long enough to be read. Tracking the display at the device level instead of the download at the server level offers infinitely more relevant data.

As a result, the benefits of a dynamic delivery portal go in both directions. Customers accessing your tech doc benefit from contextual and tailored documentation, while you gather metrics about their access that offers detailed, unbiased data on everything they see, click on, and read — all generated on the fly.



Displaying in a way that hides its fragmented nature creates a good UX while still allowing fine-grained control.



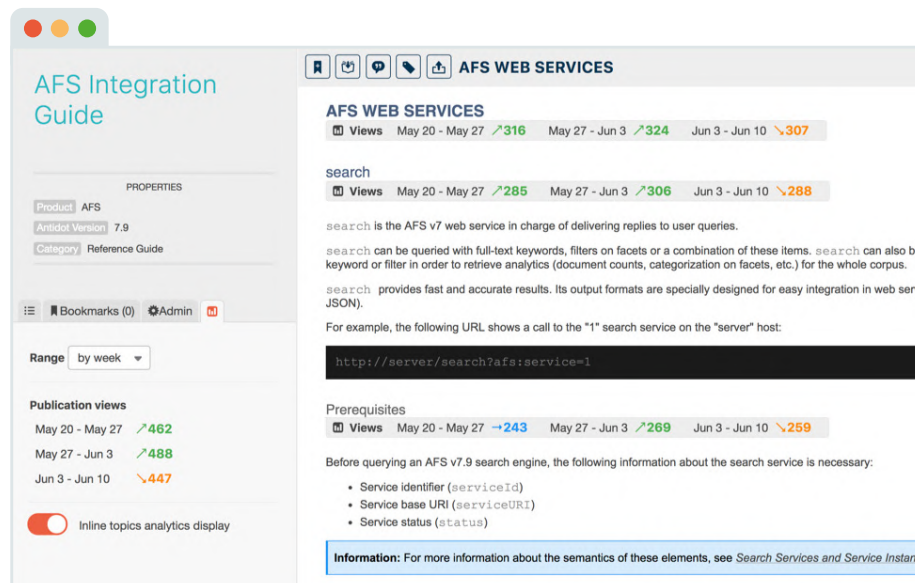
Displaying topics individually allows to track what is really being read and when.

Reading the Numbers

Numbers in context

With granular content served via a custom reading and tracking technology, a tech doc team has a much better grip on what parts of their documentation are getting attention, and which are not. You can then provide greater context for such data by displaying the metrics within the content itself. This makes it easy to understand the numbers by identifying what information they refer to.

Embedding the analytics inside the content is one of the benefits of dynamic content delivery as, when the rendering is built on the fly, it becomes easy to enrich the pages.

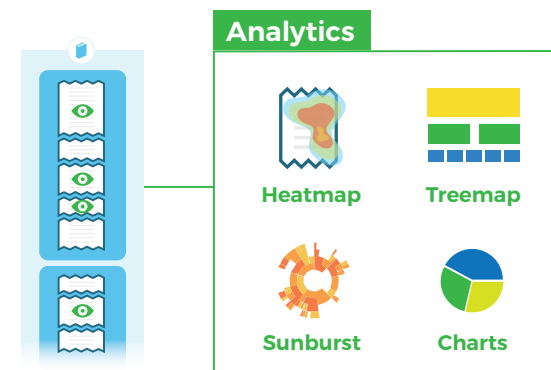


Data visualization

With a more sophisticated visual representation, you'll be able to see the entirety of the documentation via graphical tools rather than as a list of numbers and indecipherable URLs. These representations allow tech doc teams to learn where readers are focused, where their attention is drawn, what interests them, and what they need. What's more, this can be done while viewing the content, providing essential context.

These numbers can be further aggregated to provide greater meaning. For example, you may have an installation guide for each version of your software, but you could gather insights about the process of installation by aggregating all of these numbers, regardless of version.

Metadata, too, is not only useful for users conducting searches but can also be used to create clusters and axes for exploring numbers. You can analyze content consumption per product line or by the type of content. Indeed, the possibilities are almost endless and creating aggregates dynamically helps teams understand and reveal otherwise hidden data.



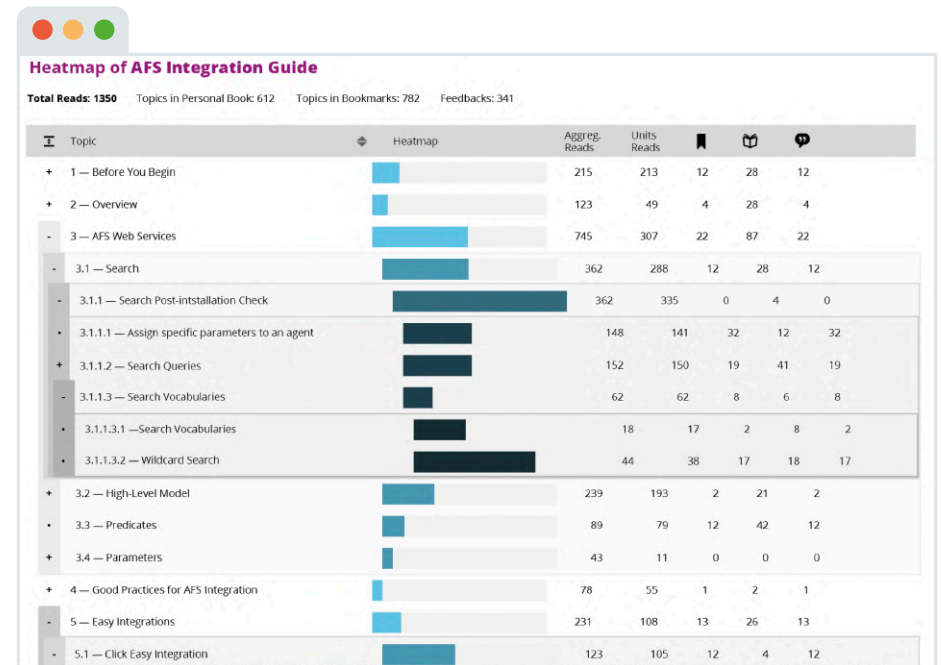
Using sophisticated data visualization such as heat maps, sunbursts, tree maps, etc., you can gain immediate insight into what is driving your users.

Reading the Numbers

Taking action

Maybe your users need to figure out how to set the clock on the stove. If your technical writers look at the data and see that this is a popular topic, maybe they could shoot an email over to Customer Support, and give them a heads-up that this is bothering users. Or they could forward reports to Product Design and let them know that their interface is confusing - maybe they can do something about that in future models. Or they could contact Sales and suggest that when they sell an item, they chat about how to set the clock.

The possibilities are suddenly blossoming because you have discovered the data hidden in your documentation consumption. Your data-driven organization now has a data-driven techdocs group



From Tracking Stats to Transforming Business

With the data flowing from your dynamic delivery portal, there are many more opportunities to unlock and generate new value for a business within and beyond the tech doc team.

Once you can collect data on every topic your users read, and know when they read that content and for how long, you can create a data portrait for each of them. Software can begin to group them in novel ways that people, full of biases and preconceived notions, could not.

While establishing a clear view of how much the technical documentation is utilized and sharing this insight with your company is a first and obvious benefit of Dedicated Content Analytics, there are dozens of ways in which next-generation content analytics can add value to a business. Four, though, stand out as particularly representative for their transformational business impact.

Better documentation

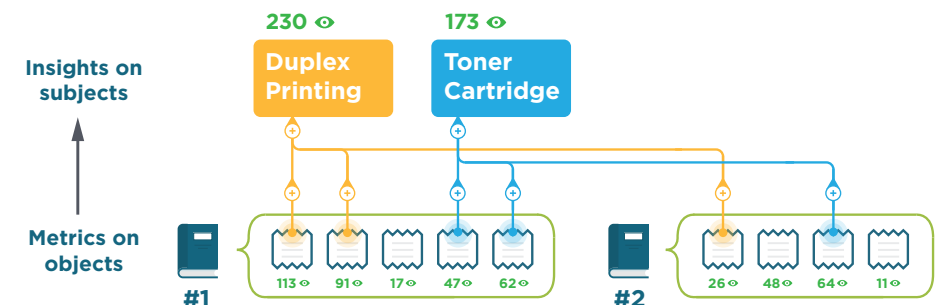
At their core, dedicated content analytics are here to help improve content. Next-gen analytics help identify content gaps by providing valuable information on, for example, “Searches with no results”. Often there is nothing more frustrating for a user than not being able to find the information they seek. This metric displays all the searches left unanswered and helps tech doc teams find out what is not yet in the documentation, but which is much needed by users.

Another key application of dedicated content analytics is for prioritizing content work. It’s no secret that documentation teams are overwhelmed with requests for new content, updates to existing content, and fixes. One of the main difficulties is to set priorities and, in this context, metrics such as “Most searched topics” and “Most read topics” help prioritize this task.

Better products

Product managers can benefit from the insights of next-generation analytics, too. Traditional marketing or web analytics will show a product manager that a particular user manual was downloaded 225 times and a troubleshooting guide 85 times. However, these document-level metrics tell the product manager nothing about the concerns of the readers: for example, are they trying to learn how to use a feature or do they want to know how to fix a problem?

Next-gen analytics, however, can add meaning to these metrics and report what it was in the user manual or the troubleshooting guide that caught the reader’s attention. What really drove engagement with both pieces of content might be something the product manager had not considered at all and, by uncovering the interactions that users had with the technical content at a fine-grained level, identifying the search queries, and extracting the subjects talked about in these sections, the product manager is far better equipped to plan the development of the product in line with customer expectations.



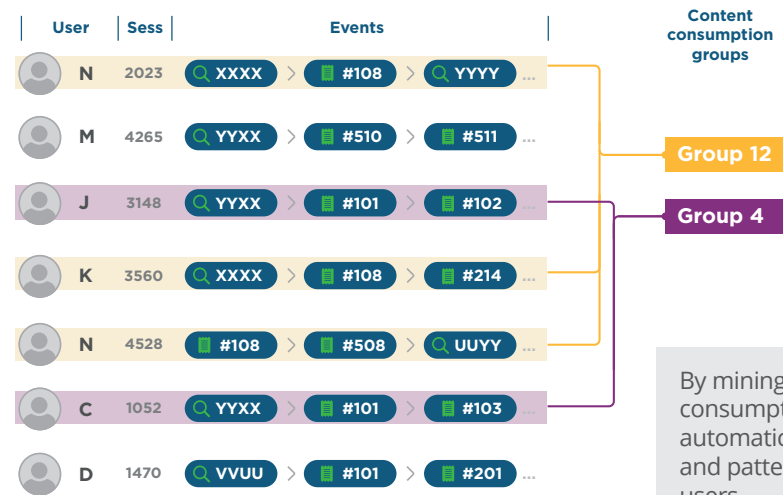
Numbers on distinct topics can be merged at a higher level to generate more insightful analytics.

From Tracking Stats to Transforming Business

Personalization

No other trend in the creation and consumption of content has revolutionized the field more than personalization. The personalization of search results, of store suggestions, of news content on a media site, of a video on YouTube – all of these aim to provoke a “stickier” content environment, a better user experience, and the capacity to serve users the content and documentation that they need, sometimes even before they even know they need it.

Next-gen content analytics offer the ability to detect patterns in the consumption of technical documentation and automatically create categories of users. These user categories can then be served content, search results, and reading suggestions that are more relevant. This, in turn, improves their engagement with the company as they are offered a truly personal interaction and service experience.



By mining content consumption, we can automatically detect and pattern groups of users.

User Group	Content				
	Book 1	Book 2	Book 3	Book 4	Art. 508
1			✓		
4	✓	✓		✓	✓
12					✓
274			✓		

Based on user C search history, promote content preferred by group 4

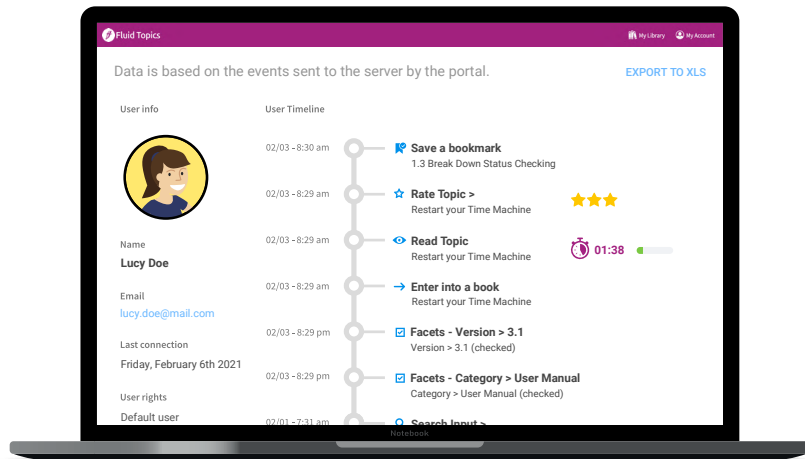
Search engine results and suggestions can be based on user characteristics.

From Tracking Stats to Transforming Business

Informed customer support

When it comes to customer support, next-gen analytics can empower support teams as never before. No longer will support agents need to ask if a user has read the manual. Instead, they will have access at a glance to everything that a user has searched and read in fine, granular detail. What's more, the support agent will have this knowledge in real time, so if a user is directed to read part of a manual and does, the support agent will immediately know. This means faster and smoother interactions for the support agent, and more effective interactions, too.

On the other side of the support interaction, the benefits that flow from the analytics improve the user experience, too. A user that has already taken pains to solve an issue themselves won't be directed to read content or review documentation that they have already seen. Instead, the support experience will be fluid, personalized, and superior to anything that is informed by page views alone. Empowered support agents deliver better customer support, and customers recognize the difference.



Predictive customer support

Customer support does not need to end with empowerment. Fully featured next-gen analytics can drive predictive customer support, with positive flow-on effects for support agents and users alike. Using cutting-edge machine learning trained on and consuming real-time data, you can identify trends and correlations in support experiences.



Popup windows could be displayed at the right moment by tracking user behavior and content consumption.

For example, an analysis of content consumption might suggest that – after a 20-minute period of reading technical documents – 85 percent of queries around a particular search term result in a call to a support line. Knowing this and coupling this insight with sales volume data and product specifications, businesses can drive proactive interactions and help predict appropriate staffing levels for support teams. Additionally, it allows support teams to receive real-time warnings when search term volume for a particularly problematic phrase spikes. Teams are thus better prepared for the volume and topics of the calls that will shortly arrive.

ABOUT US

Next-Generation Content Analytics

Conclusion

So, we can see that Grace Hopper's warning rings true in terms of analytics. That's why it's time to embrace innovation and leverage the opportunities for business transformation that they offer.

By focusing on user behavior to capture interactions with content at the finest level and derive truly relevant metrics, it's clear that next-generation analytics offer a real opportunity for businesses to transform their tech documentation into a true business asset.

But they can also help transform other parts of their companies as well. In fact, as we've seen, they can drive business across the board.

Dynamic delivery of content sets a new standard for publication and delivery. As a result, you can now expect tech content and analytics to offer deeper insight and choose solutions that support these changes.

What is Fluid Topics?

Our AI-powered Content Delivery Platform captures all technical documentation and transforms it into a smart knowledge hub to deliver actionable information that is tailored to the user, adapted to the situation, and suited to the channel.

Looking for a custom demo?

If your tech doc team and organization is ready to level up with a CDP, reach out to the Fluid Topics team for a custom demo and a hands-on experience with our market-leading CDP solution.

Contact us:

For technical questions:

fluid.support@fluidtopics.com

Website:

www.fluidtopics.com

Blog:

www.fluidtopics.com/blog