

LavaCon 2024

The Content Strategy Conference lavacon.org

What does good look like? Measuring content performance

ROB HANNA

CEO & CO-FOUNDER, PRECISION CONTENT



Shaping the future of content





Precision Content

Precision Content is shaping the future of content. We transform your complex information into easy-to-understand content, making it accessible and intuitive for everyone. Leverage our expertise in structured authoring and DITA to enhance content quality, streamline workflows, and improve content operations. Visit our website for solutions that bring precision to your content challenges.



*AT PRECISION CONTENT, WE BUILD CAREERS IN
INFORMATION ARCHITECTURE*

Best
Workplaces™

Great
Place
To
Work®

CANADA
2024



Presentation agenda

Pressing need for quality

Measuring human performance
on content

- Creating quality benchmarks
- Usability labs
- Usability lab scenario
- Sentiment analysis

Measuring machine
performance on content

Role of microcontent

Conclusions



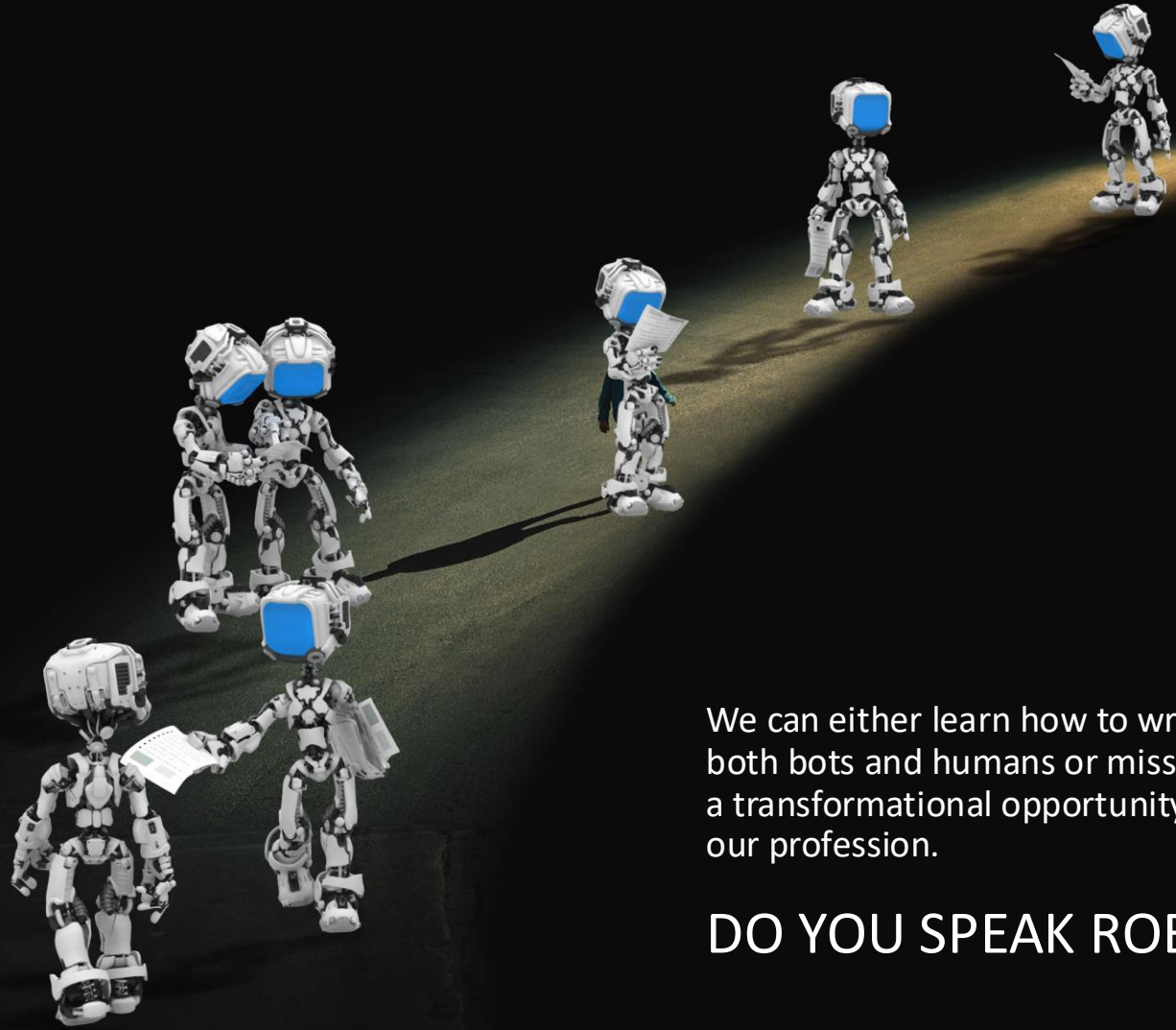


Why does it matter (now)?

Quality, consistency, and precision matter now more than ever with the increasing prominence of artificial intelligence and generative AI.

The maxim “**Garbage in: Garbage out**” was never truer than it is today as companies and professionals look to differentiate themselves and compete on a new playing field with machines.

The robots are
coming ...



We can either learn how to write for
both bots and humans or miss out on
a transformational opportunity for
our profession.

DO YOU SPEAK ROBOT?



Future of Information 4.0

Despite our focus on simplifying the technology around content – content itself is not getting simpler

Creating progressively smaller and smarter content components

Focus on answers that people are seeking

Information 4.0 for Industry 4.0 – Enabling Information Interoperability,

<https://www.slideshare.net/jgollner/information-40-for-industry-40-tcworld-2016>



November 30, 2022: ChatGPT First Public Release!





GPT

~~Father~~ Knows Best*



**CBS Television circa 1954*

GPT startled many who have worked with chatbots in the past

Instead of saying "I do not know" or querying the user for clarifying input, GPT just gives a pat answer regardless of the probability of being correct

This will fundamentally change how we interact with the connected cyber universe allowing for voice and other innovations in content delivery



GPT:> How can we help?

How can we help our clients to be successful with this new technology?

What role will tech pubs play?

Will structured authoring give us an advantage?





“Data is food for AI” – Andrew Ng

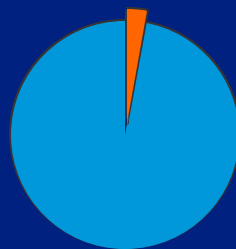


Andrew Ng, Co-founder, Head, Google Brain

Data quality is STILL a big problem:

OpenAI has openly stated that one of the biggest issues in training ML models like Dall-E, GPT-3, and ChatGPT were errors in the data and labels.

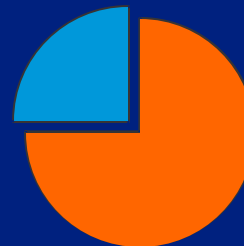
Academia



■ Datasets
■ Models and Algorithms

Algorithm-centric AI

Industry



■ Datasets
■ Models and Algorithms

Data-centric AI

<https://dcai.csail.mit.edu/lectures/data-centric-model-centric/>



Focus on enabling technology

Most sessions on GenAI focus on the tools and techniques needed to successfully leverage AI for authoring and delivery. Fewer sessions talk to the broader success of GenAI across companies. Far fewer deal with the content quality issue as it is a much tougher nut to crack.

How can you measure quality without a benchmark for quality?

This session will detail how we create benchmarks on your content and then test human performance on that content.

Improvements in machine performance is harder to measure. Instead we need to look at the architecture along with the quality of the content for machine performance.

What are you feeding your robot?





Measuring Quality

What is quality in technical communication?



What do we measure?

We look at many facets for measuring value and quality of content. Not all content is evaluated the same way.

For marketing type content, we consider

- activity
- engagement
- voice and tone
- relevancy, and
- user feedback.

These measures are similar yet different for technical communication.



Evaluating technical communication

We need to assess for quality through

- readability and plain language
- accessibility
- consistency
- accuracy
- efficiency, and
- trust.





Creating content quality benchmarks

Measuring human performance on your content



Writing Methodology

Precision Content® Writing Methodology™ is a research-based approach for breaking information into smaller components and organizing and structuring information.

Benefits:

- Easier to write and read
- More accurate
- Ideal for Component Content Management Systems (CCMS)
- More reusable

Process

Determine type of maintenance required:

Notes:

For a rush request instructions, see [Rush Request](#).

The CKO Maintenance Spreadsheet - Sample job aid can assist with the completion of the CKO Maintenance Spreadsheet.

If request is to:

Delete an account from CKO service

Then:

Record and track the request in Navigator and select the following:

Case Field or Tab	Selection
Sub Product	PRINT OUTSOURCING (JPMC)
Reason	Maintenance/Setup
Case Type	Change
Notes & Interactions	Correspondence related to the request

Complete the CKO Maintenance Spreadsheet.

Review the Instructions tab for guidance.

Navigate to the General and Accounts tab to enter the data for the client's request.

General tab: In the Work Summary column, enter detail of the request, such as "delete accounts 1, 2, 3," etc.

Accounts tab: Complete the Bank and Account (max 13) columns each account.

Email the completed spreadsheet to CKO Fulfillment. CKO Fulfillment sends a notification when the request is complete.

Notes:

These types of maintenance requests are in production within 48 hours of receipt of a properly completed spreadsheet.

Go to [Wrap-up](#).

Change billing account

Change email address

Review reports

Change address on check or remittance statement

Courier code changes:

Changes to the courier and/or the courier billing account

Change courier code

Change address

New special handling (courier) code for an existing client who already uses courier codes (listing is not required)

Record and track the request in Navigator and select the following:

Case Field or Tab	Selection
Sub Product	PRINT OUTSOURCING (JPMC)
Reason	Maintenance/Setup
Case Type	Change
Notes & Interactions	Correspondence related to the request

For requests to change the billing account where the Check Print activity account is an ERISA account, the billing account in the Check Print system must be the same as the activity account.

Complete the CKO Maintenance Spreadsheet.

Notes:

The CKO Maintenance Spreadsheet cannot be used for changes that require testing (e.g., adding a new form code, adding a new account, or adding a courier code for a client who has not used courier codes to date). The CKO Maintenance Spreadsheet can be used when an existing client, who uses courier codes, is adding a new courier code.

Email the completed spreadsheet to CKO Fulfillment. CKO Fulfillment sends a notification when the request is complete.

Notes:

These types of maintenance requests are in production within 48 hours of receipt of a properly completed spreadsheet; they do not require an install date.

Go to [Wrap-up](#).

Determine the request type

Determine the request type before facilitating a client request.

1. For client requests that need to be processed urgently, go to the [Process Rush Requests](#) section.
2. For regular maintenance requests, determine the type of maintenance before processing the request.

Notes: While processing maintenance requests, use the CKO Maintenance Spreadsheet - Sample job aid for help with completing the CKO Maintenance Spreadsheet.

Facilitating maintenance requests

The CSA facilitates maintenance requests in those cases where the LOB has set up an incremental business and in cases where LOB has not set up an incremental business.

These are the stages of facilitating a maintenance request.

1. The CSA
 - a. determines if the service that the client requested is to
 - i. Delete an account from CKO service
 - ii. Change the billing account, email, address, or details of the courier codes
 - iii. Replace a logo or signature in client's existing template, or
 - iv. Complete technical maintenance requests
 - b. creates the maintenance implementation request (R) in Treasury Services (TS) Deal Manager
 - c. completes the product line item (PLI) specification sheet
 - d. submits the request in TS Deal Manager, and
 - e. completes the steps in the [Close the case](#) section.

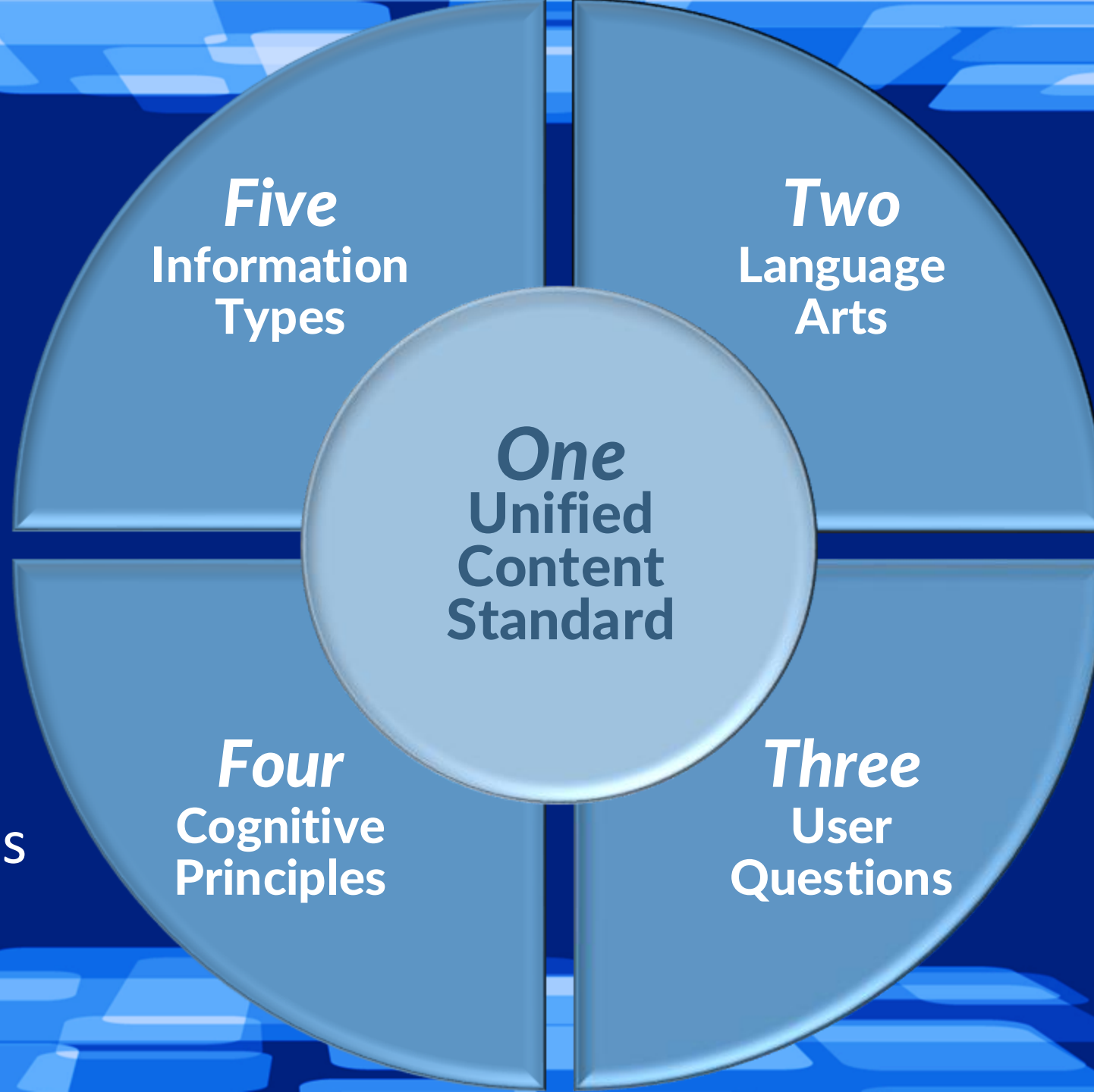
Delete an account from CKO service

Follow these steps to delete an account from the CKO service.

1. Record and track the request in Navigator.
2. Select the correct information for each case field or tab according to the values listed here.

Case Field or Tab	Value
Sub Product	PRINT OUTSOURCING (JPMC)
Reason	Maintenance/Setup
Case Type	Change
Notes and Interactions	Correspondence related to the request

3. Complete the CKO Maintenance Spreadsheet.
 - a. Review the Instructions tab for guidance.
 - b. Navigate to the General tab and enter the client request data, including details of the request such as "delete accounts 1, 2, 3, and so on," in the Work Summary column.
 - c. Navigate to the Accounts tab to enter the client request data and complete the Bank and Account (max 13) columns for each account.



Precision
Content®
writing
fundamentals



Editorial review of titles and structures

Before

After

The title is vague. A process for what?

The note text is not helpful

Steps are buried

Although titled a "process" this information is "procedural"

Process

Use Service Portal to determine if the client has Check activity on the account.

Note:

This information will be useful later in this procedure.

Use Service Portal and Access Administration Center (Admin Center) to determine whether the client has positive pay and/or PVE on the account(s) and the status of positive pay.

Use the following tools (in the order listed)	Then follow the steps to verify a client's information
Service Portal	Click the Client Profile tab and select Client Hierarchy from the drop-down list. Search for the client by typing the account number, ECID, client name, or TIN. Click Search . Select the client name and click View Client Hierarchy . Select the Product Setup Summary widget. Select Yes to view the client's Access details by ECID Level or Ultimate Parent Hierarchy. Select the Access ID to view the Access Profile Details . Select the Products tab and locate the Checks Module to verify if the client is has positive pay. Click the Launch Access button to enter Access Admin Center.
Access Admin Center	Search for the client GEID and click GO . Click View . Click Client Profile Products . Select the Products tab then select Checks . Select the Account tab. In the Show menu select Positive Pay . You will see the status of either Pending Positive Pay or Active Positive Pay and whether or not the client has PVE on the account(s).

Determine the status of Positive Pay or PVE on client accounts

Use Service Portal and Access Administration Center (Access Admin Center) to determine if the client accounts have Positive Pay or PVE and the status of Positive Pay.

Follow these steps to determine the status of Positive Pay or PVE.

1. Verify the client accounts have Positive Pay or PVE using the Service Portal.
 - a. In the Client Profile tab, select Client Hierarchy.
 - b. In the Search field, enter the account number, ECID, client name, or TIN.
 - c. Click Search.
 - d. Select the client name and click View Client Hierarchy.
 - e. Select the Product Setup Summary widget.
 - f. Select Yes to view the client's Access details by ECID Level or Ultimate Parent Hierarchy.
 - g. Select the Access ID to view the Access Profile Details.
 - h. In the Products tab, locate the Checks Module to verify if the client's account has Positive Pay.
2. Determine the status of Positive Pay or PVE on the client accounts using the Access Administration Center (Access Admin Center).
 - a. In the Checks Module, click the Launch Access button to enter the Access Admin Center.
 - b. In the Search field, enter the client GEID number.
 - c. Click Go.
 - d. Click View.
 - e. Click Client Profile Products.
 - f. In the Products tab, select Checks.
 - g. In the Account tab, under the Show menu, select Positive Pay.
You can see the status of the client accounts displayed as either Pending Positive Pay or Active Positive Pay and if the client has PVE on the accounts.

The title is clear. It informs and guides the reader

Clear instructions

Steps and sub-steps are clear

This is task information and follows specific writing rules



For more information about our methodology

To get started, follow this QR code or
visit <https://pcas.info/get-ai-ready> for
more information.





Readability statistics



Writing tools for quality

Technical communication groups will use a variety of tools to measure the quality of their writing. These may include

- Microsoft Word
 - for spelling and grammar
 - passive voice, and
 - readability scores.
- Acrolinx
 - for terminology management
 - voice and tone, and
 - scannability.
- HyperSTE, and
- Grammarly.



Scores and stats

Factors	Before	After	Improvement
Words per sentence	13.5	11.8	12.6%
Passive sentences	8.8%	2.3%	74%
Flesch Reading ease	63.0	59.0	-6.3%
Flesch-Kinkaid Grade level	7.7	7.9	2.6%



Usability labs



Facilitated user study

Both before and after samples are presented to a user in a controlled environment.

The facilitator will

- observe the behaviour of the user
- record the time-to-answer
- assess the confidence level, and
- evaluate the answer.





Test scenario

Your boss has just called you at your desk to ask a question. You've placed them on hold while you look up the answer.

Scan the procedure document and find the answer.

Once you have found the answer, pick up the phone and relay the answer to your boss.





Lab example results



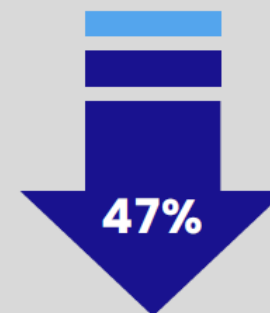
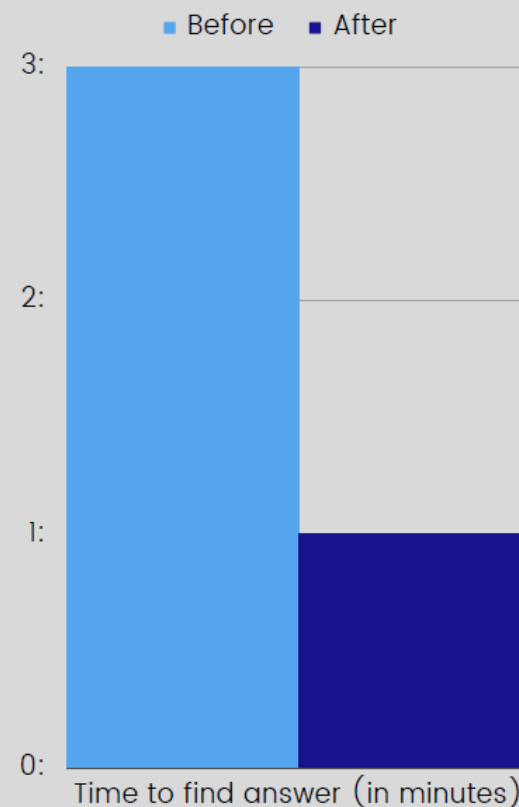
Usability lab test results

Participant	Procedure tested	Before Question				After Question				Difference
		Number	Time	Errors	Confidence	Number	Time	Errors	Confidence	
User 1	FISPAN	#1	04:08	N	80%	#2	01:49	N	100%	02:19
User 2	FISPAN	#1	05:15	N	80%	#2	00:55	N	80%	04:20
User 3	FISPAN	#1	03:12	N	60%	#2	01:41	N	100%	01:31
User 4	FISPAN	#13	02:08	Y	100%	#14	01:29	Y	100%	00:39
User 5	New Cashier's Check Request	#15	07:26	Y	60%	#16	01:25	N	100%	06:01
User 6	New Cashier's Check Request	#3	04:01	Y	60%	#4	00:55	N	100%	03:06
User 7	New Cashier's Check Request	#3	00:58	Y	100%	#4	02:45	N	100%	01:47
User 8	Initiation	#5	04:22	N	100%	#6	02:00	N	100%	02:22
User 9	Initiation	#5	03:20	N	60%	#6	03:01	N	100%	00:19
User 10	Initiation	#11	01:37	N	60%	#12	01:07	N	90%	00:30
User 11	PP/PVE	#7	00:59	N	80%	#8	01:06	N	100%	00:07
User 12	PP/PVE	#7	04:04	Y	100%	#8	01:16	N	100%	02:48
User 13	PP/PVE	#7	02:46	Y	80%	#8	02:29	N	100%	00:17
User 14	Maintenance/Incremental Business	#9	05:11	N	100%	#10	04:32	N	60%	00:39
User 15	Maintenance/Incremental Business	#9	02:47	Y	50%	#10	01:06	Y	80%	01:41
Averages			03:29	47%	78%		01:50	13%	94%	
		seconds	209			seconds	110			



User Performance Index

Time to find answer – results



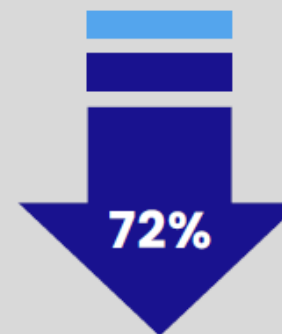
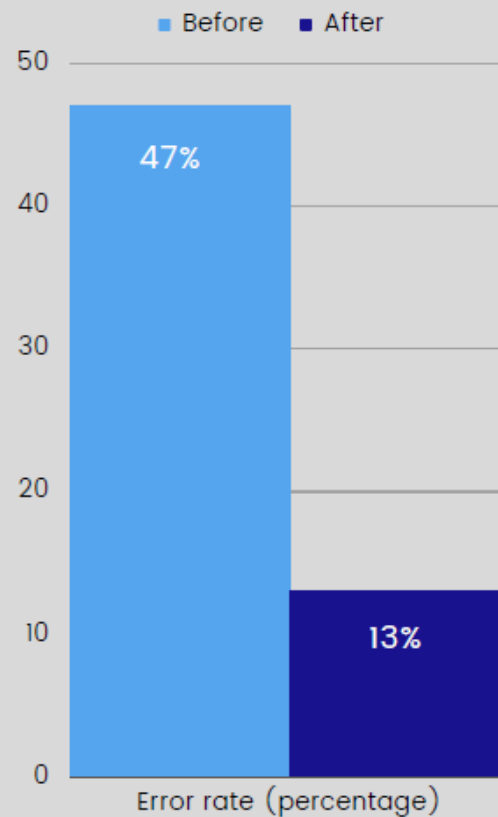
The time it took to find answers improved dramatically, resulting in an average reduction of 47%

TIME TO FIND ANSWERS	BEFORE	AFTER
Minimum time	0:58	0:55
Maximum time	7:28	4:32



User Accuracy Index

Error rate



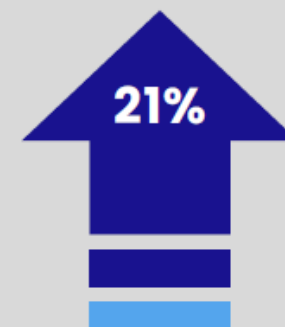
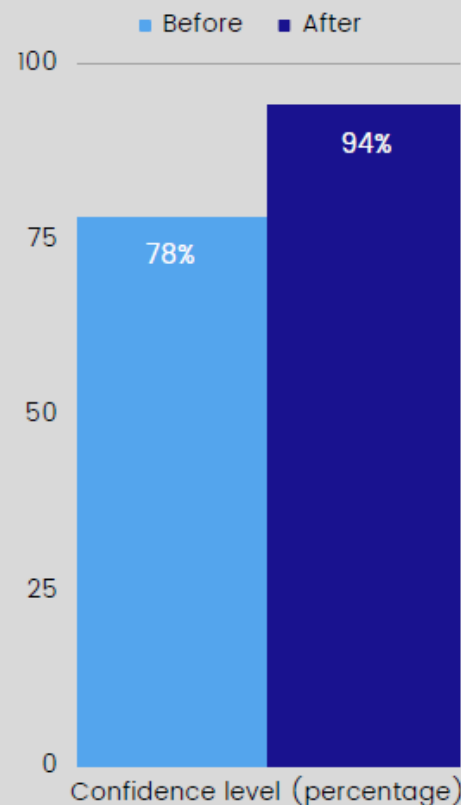
The participants made far fewer errors in finding the information they were looking for in the 'after' document, resulting in a average reduction of 72%

ERROR RATE (OUT OF 15)	# OF ERRORS
Before	7
After	2
Difference	5



User Confidence Index

Confidence level



Confidence in finding the right answer improved by an average of 21%.

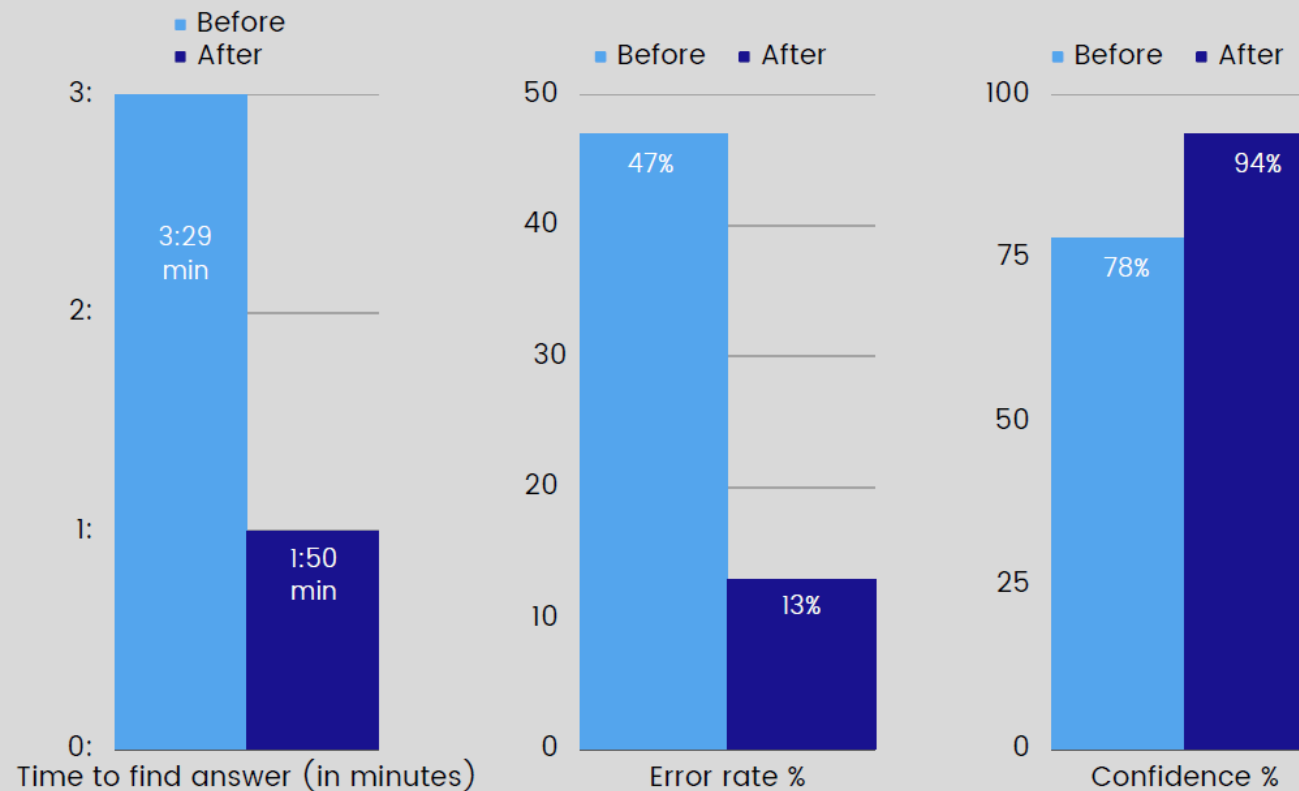
CONFIDENCE IN ANSWER	BEFORE	AFTER
Minimum	50%	60%
Maximum	100%	100%



Before and After Improvements

Usability Test Results

High level statistics from the Usability Lab data





Sentiment analysis

Using AI for objective measurement of user sentiment



Measuring sentiment

Observed behaviour

Transcript analysis

Facial analysis

Prosody analysis



Commentary

Excerpts from the participant transcripts. Some have been edited or paraphrased for clarity.

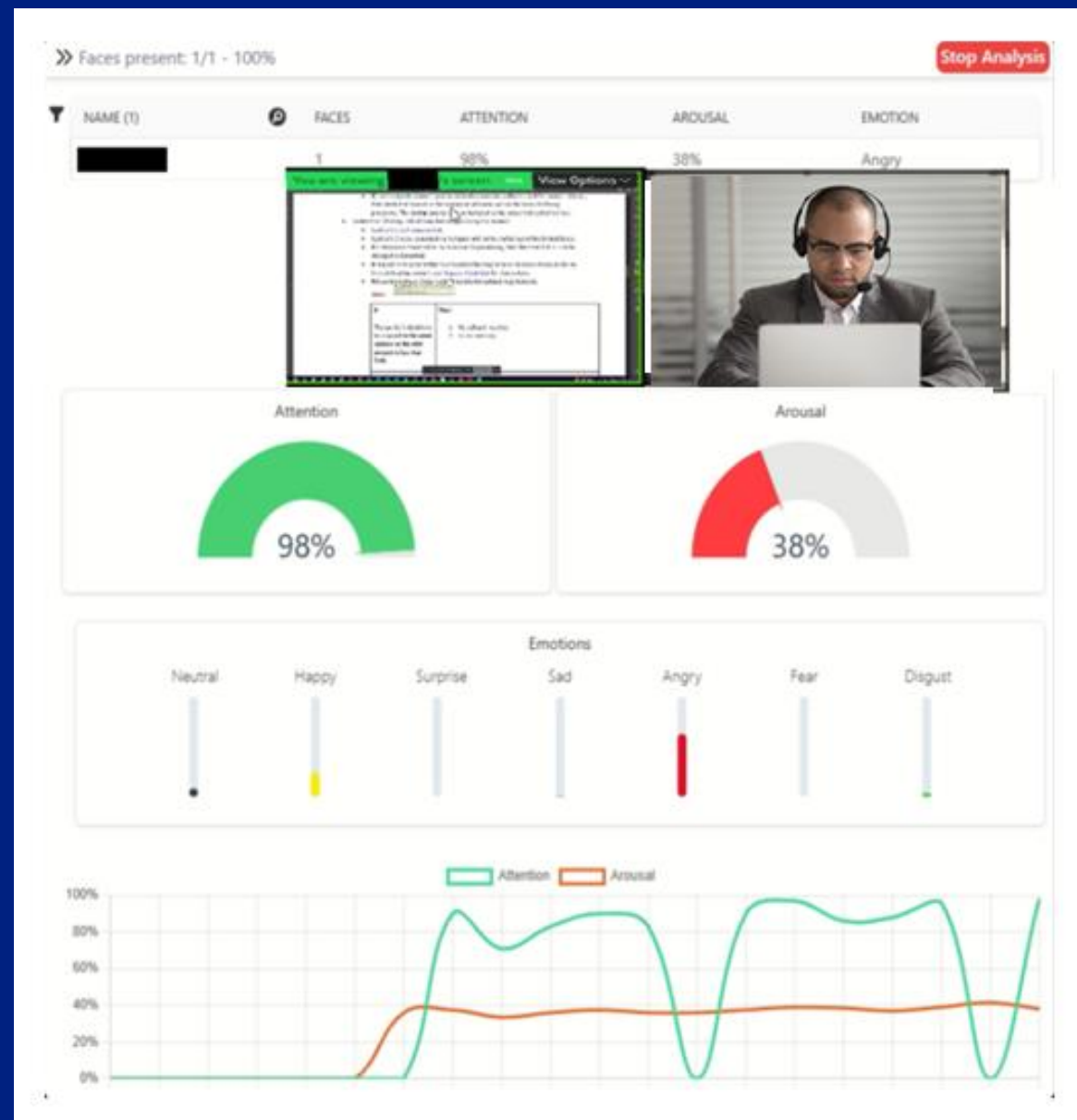
"It's the same information? You gotta be kidding me. It's not even comparable. The second one is much faster. I can't even believe it's the same material. It's laughable, it really is."

"The table of content helped a lot. Key words in the table of contents went right to the information to help find the answer. In the after, there were far fewer steps to read through, so I could jump right to the step I was looking for. Far less wording, more streamlined. And I like "If, then" tables a lot."

"The titles, the formatting; it's more styled. It's cleaner. Today, we have way too much info that we have to put in everything. There's way too much way, too many procedures and way too many things. This new way looks great."

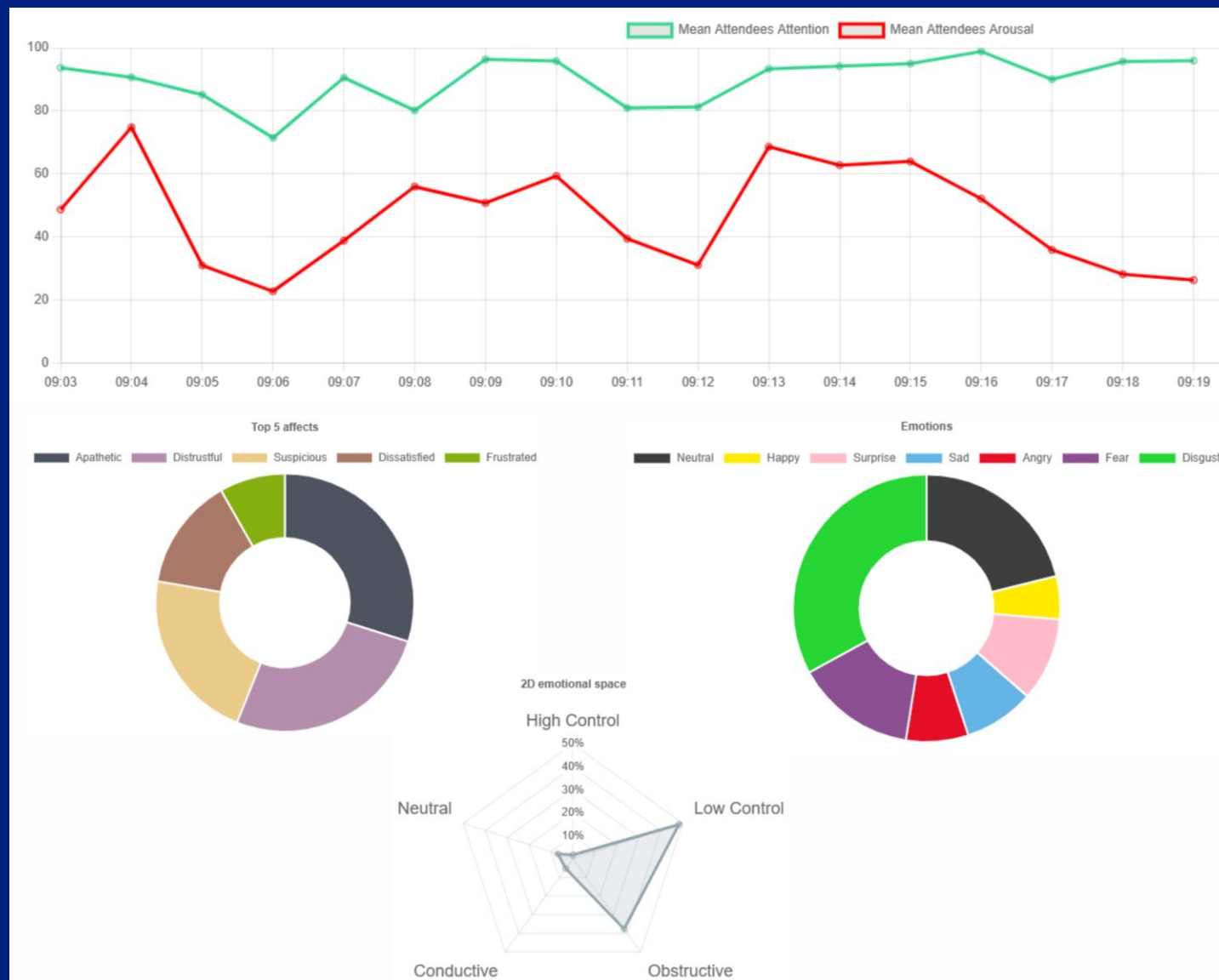


Real-time facial analysis



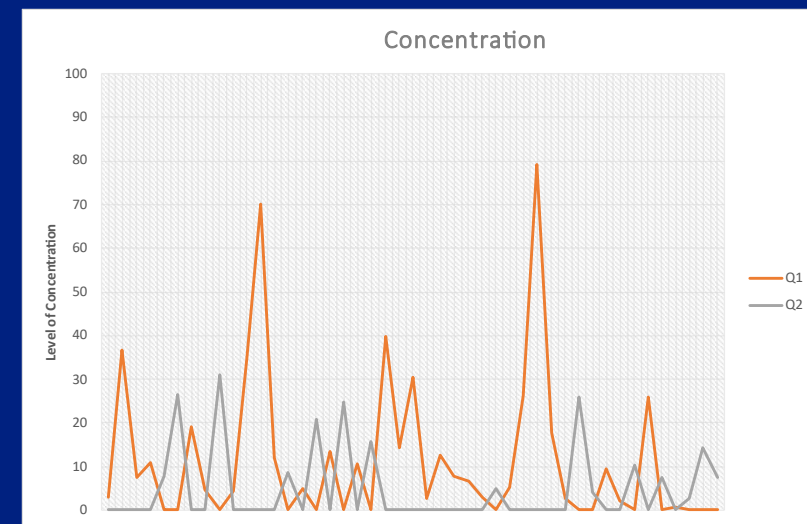
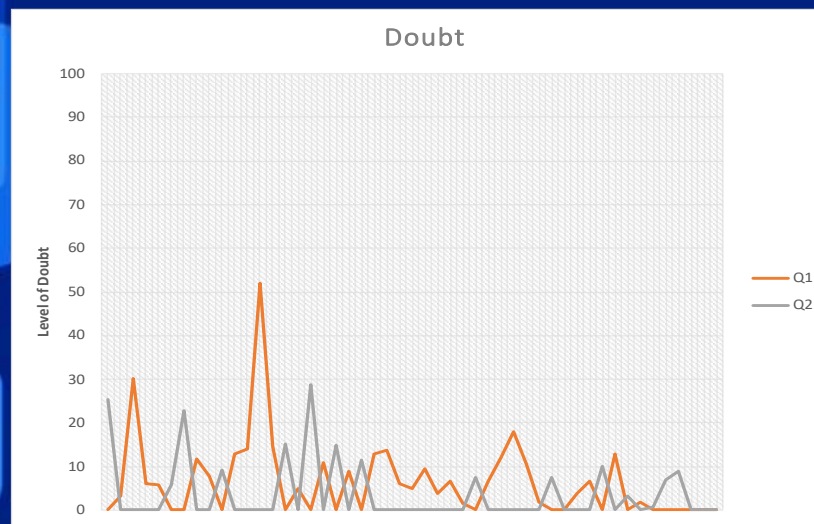
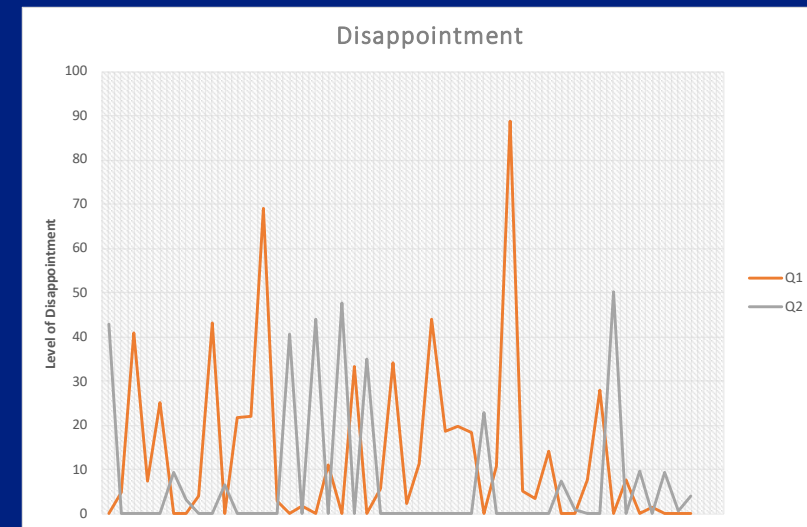
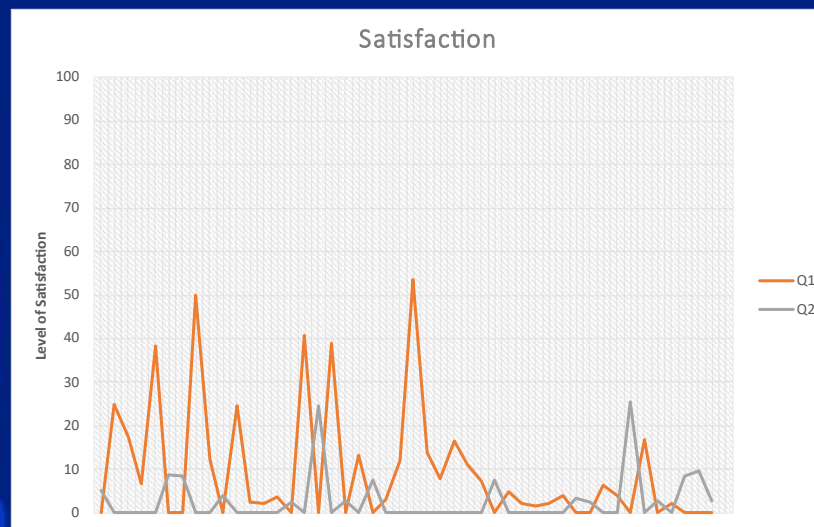


Session summary





Voice Prosody Results User 1





AI-interpreted sentiment analysis

An important gauge for usability of content will always come down to trust for the user. If the user does not trust the content, it will slow the user down in performance with that content and potentially cause fault with the content as the user doubts the veracity of the content if and when they find it.

Trust is of course a subjective measure.

AI can help to remove some of that subjectivity but the technology is still in its early development. While facial expressions and language used by the participants is heavily nuanced, prosody proved to be the most accurate when compared against observed results.



Measuring machine performance on content

Preparing your quality content for machine consumption



GPTs and RAG use in techcom

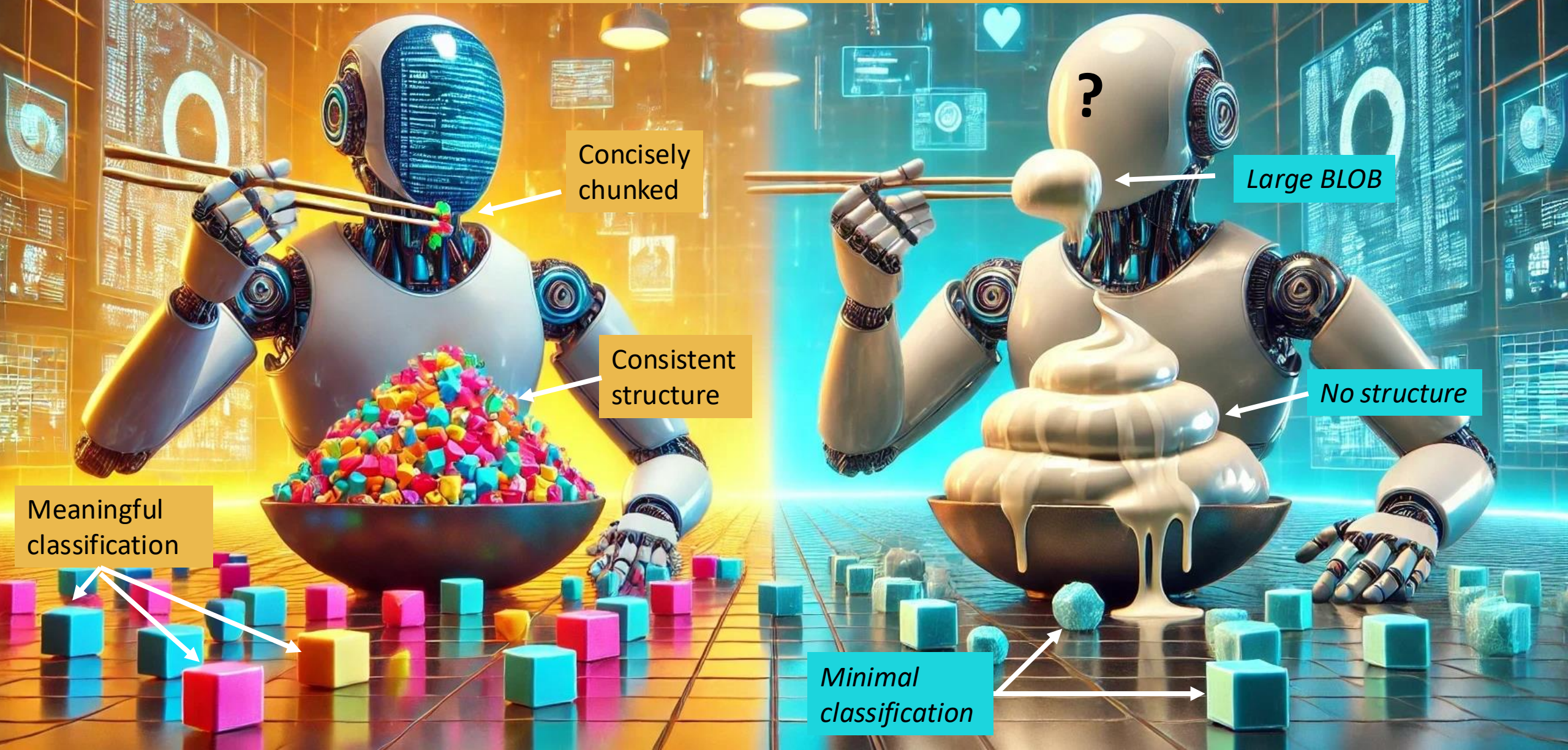
Generative Pre-trained Transformer (GPT) models are in widespread use as a writing tool to create text faster and with better consistency when used to re-write text or summarize text.

Retrieval Augmented Generation (RAG) is used primarily to make technical content more widely available to real-time chatbots and AI assistants.

It uses a retriever to locate answers to questions and an LLM to generate a response in context with the query. This solution relies on a high-quality corpus of content and many architectural features that make retrieval more reliable.

The retriever may rely on multiple technologies to improve accuracy including natural language processing, elastic search, knowledge graphs, and vector embeddings.

What are you feeding your robot?



Concisely
chunked

Consistent
structure

Meaningful
classification

?

Large BLOB

No structure

Minimal
classification



What you *should* be feeding your robot

A domain-relevant, accurate, and well-structured corpus is key to maximizing the precision, relevance, and efficiency of RAG solutions.

1. Relevance and Domain-Specificity

Ensure content is highly relevant to the target domain to improve response accuracy.

2. Data Freshness and Accuracy

Keep information up-to-date and prioritize correctness for reliable, accurate responses.

3. Balanced Coverage with Granularity

Cover topics with well-segmented chunks, aligned to likely user queries for targeted retrievals.

4. Diversity and Clarity of Language

Use varied phrasing to support query diversity, while maintaining plain language to reduce irrelevant content.

5. Consistent Structure and Formatting

Standardize formatting and structure for efficient and precise retrieval.

6. Noise Reduction with Minimal Redundancy

Eliminate irrelevant or duplicate information to improve retrieval efficiency.



Diversity of language

While diversity of terminology, phraseology, and language is important to meet the needs for diversity in the prompt queries, it works against best-practices for technical content where consistency is key.

To better cope with query diversity, RAG solutions should

- Use Paraphrasing and Query Rewriting Techniques
- Implement Synonym Mapping and Lexical Matching
- Enhance the Query Expansion Techniques
- Implement a Guided Query System with Suggested Queries



Redundancy and reuse

Content reuse is over-emphasized on many CCMS implementations without having a proper strategy and governance around reuse.

Reusing coarser, structured chunks of content rather than reusing anything that can be reused will reduce complexity to the source and redundancy to the output.

Redundancy will make it harder for Gen AI solutions to sort out the facts and bias responses based on the matching frequency of segments of text.

Feeding a model with discrete blocks of microcontent instead of larger topics or pages of content will reduce this redundancy while also supporting context

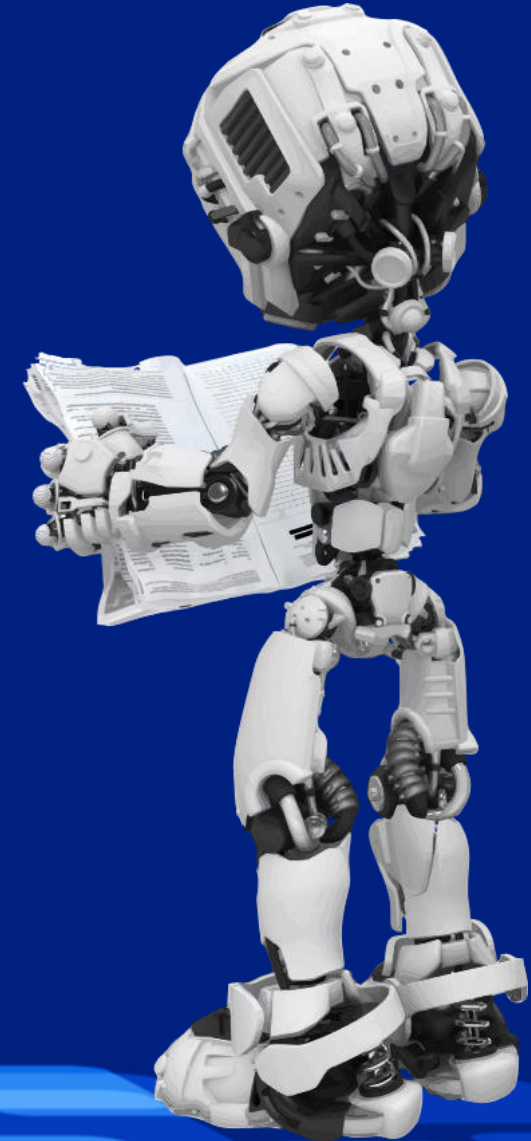


Microcontent

content that is

- about one primary idea, fact, or concept
- easily scannable
- labelled for clear identification and meaning, and
- appropriately written and formatted for use anywhere and any time it is needed.

It's not microcontent just because it's small





Microcontent

1. Purpose

Microcontent must be about only one subject

2. Function

Microcontent must be typed to identify user intent

3. Structure

Microcontent must use predictable patterns and language

4. Context

Microcontent must be easily relatable to other content



Conclusions

Our content isn't ready for generative AI technology to work effectively for enterprise information

While LLMs are getting better at refining general knowledge, they will never possess the subject matter expertise about your products to create better technical content

Fine tuning your model will improve responses and reduce hallucinations but it will never eliminate them

Organizations will need to evaluate the ROI for content transformation versus costs for fine tuning and coping with the consequences of hallucinations

A microcontent strategy can improve performance for both humans and machines



A well-fed robot ...

Is a smarter, more
helpful robot.



Does your content speak robot?



Follow this QR code or visit
<https://pcas.info/ai-ready> for your
free AI-Readiness Assessment.

