

From Papyrus to the Virtual Elephants in Rooms

Joyce (Jo) Lam & Keith Schengili-Roberts



Joyce (Jo) Lam

“Enable Information”
joyce.lam@enableinfo.com

MSc Information Design and
Development

Information Architect at AMD

Toronto Past President of
Society for Technical
Communication

Member of Information 4.0
Consortium

Member of Women in Tech

Member of Women in STEM



Keith Schengili- Roberts

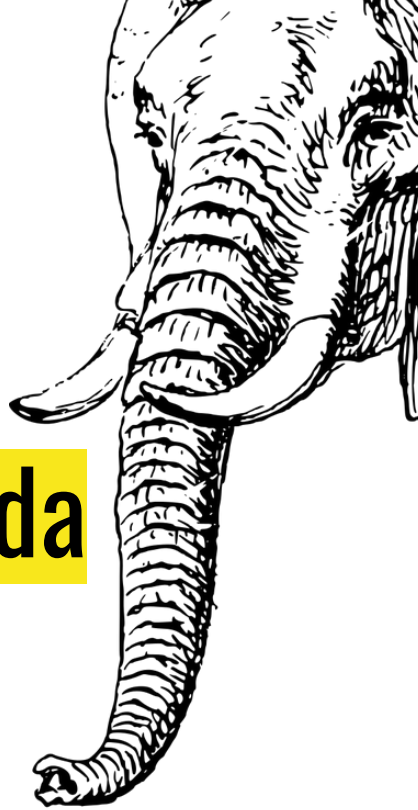
"DITAWriter"
keith@ditawriter.com

Information Architecture
Instructor at the
University of Toronto

Technical Documentation
Manager at AMD

Member of OASIS DITA
Technical Committee

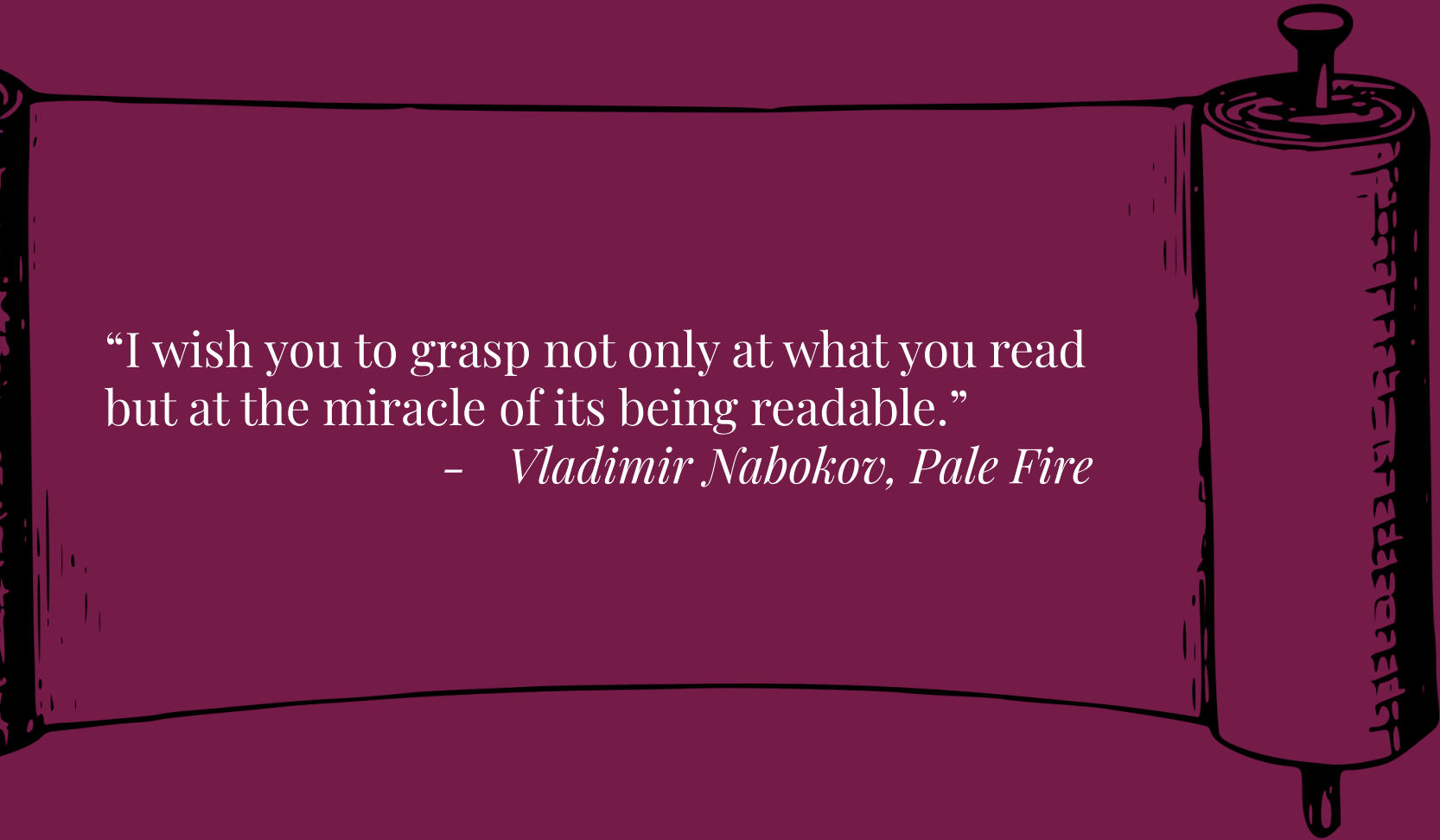
Agenda



- Reading is not Natural
 - Instinct and Pursuit
 - Technology of Documentation
 - New Modes for Old Brains
-

“I wish you to grasp not only at what you read
but at the miracle of its being readable.”

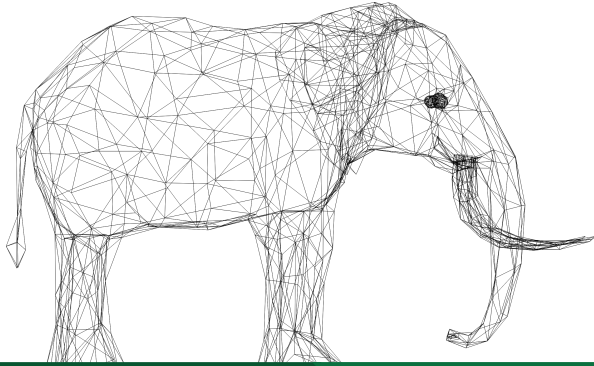
– *Vladimir Nabokov, Pale Fire*



Reading is not Natural



Reading needs Systematic Training to Recognize Patterns



Reading is not an innate ability, but rather it is an artificial skill that requires systematic instructions and practice to learn, improve, and maintain.

— *Sousa, 2005*

Basic Features

Identify lines, contours, and shapes.

Characters

Combine basic features.

Words

Combine characters.

Sentences

Combine words.

Paragraphs

Combine sentences.

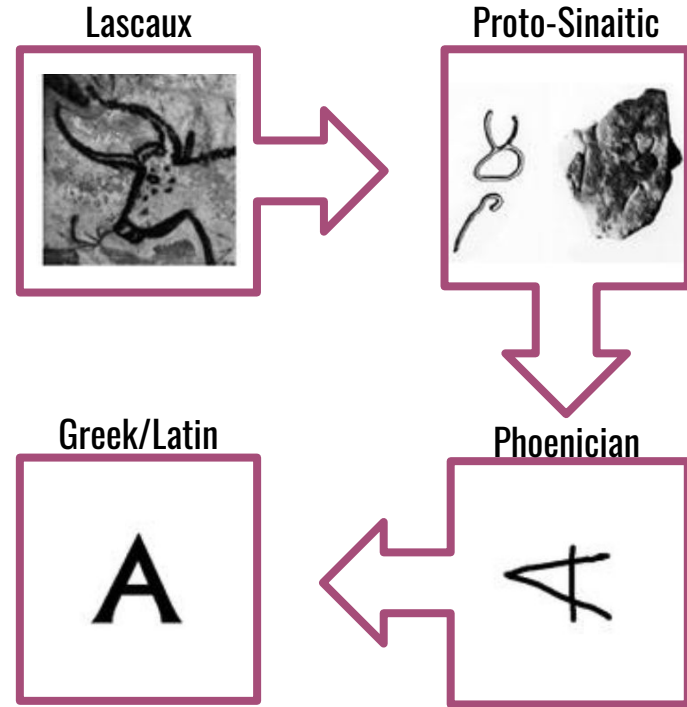
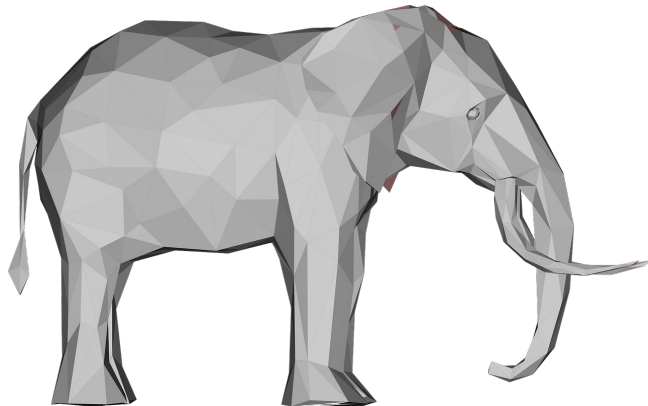
— *Johnson, 2014*

Our Brain is Wired for Pattern Recognition

Ancient scribes adapted lines and strokes to represent figures.

Most characters in worldwide alphabetic systems consist of sets of no more than three lines or strokes.

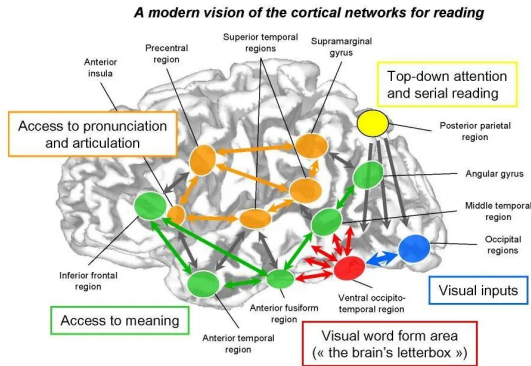
All writing systems link back to sound. Our neurons have been “recycled” for use in reading, assembling strokes into letters, letters into words, words to meaning (Dehaene, 2009).



Images from *Reading in the Brain*,
Stanislas Dehaene, 2009

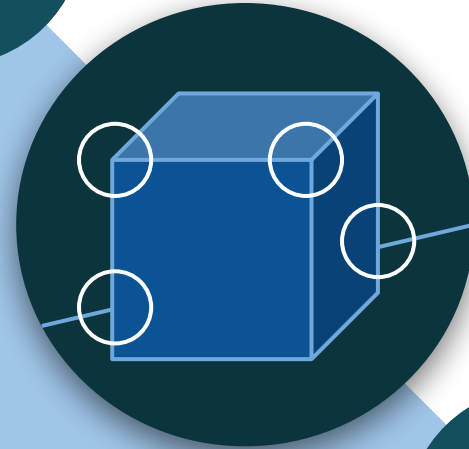
Featured-driven Reading

Whenever one object intersects with another it usually produces a “T-junction” and our visual system has neurons adapted to recognizing these types of intersections (Dehaene, 2009).



Imagens

The mental representation of visual information. Consists of **objects**, **natural groupings of objects**, **whole parts of objects**, and **spatial information** (Johnson, 2014).



Logogens

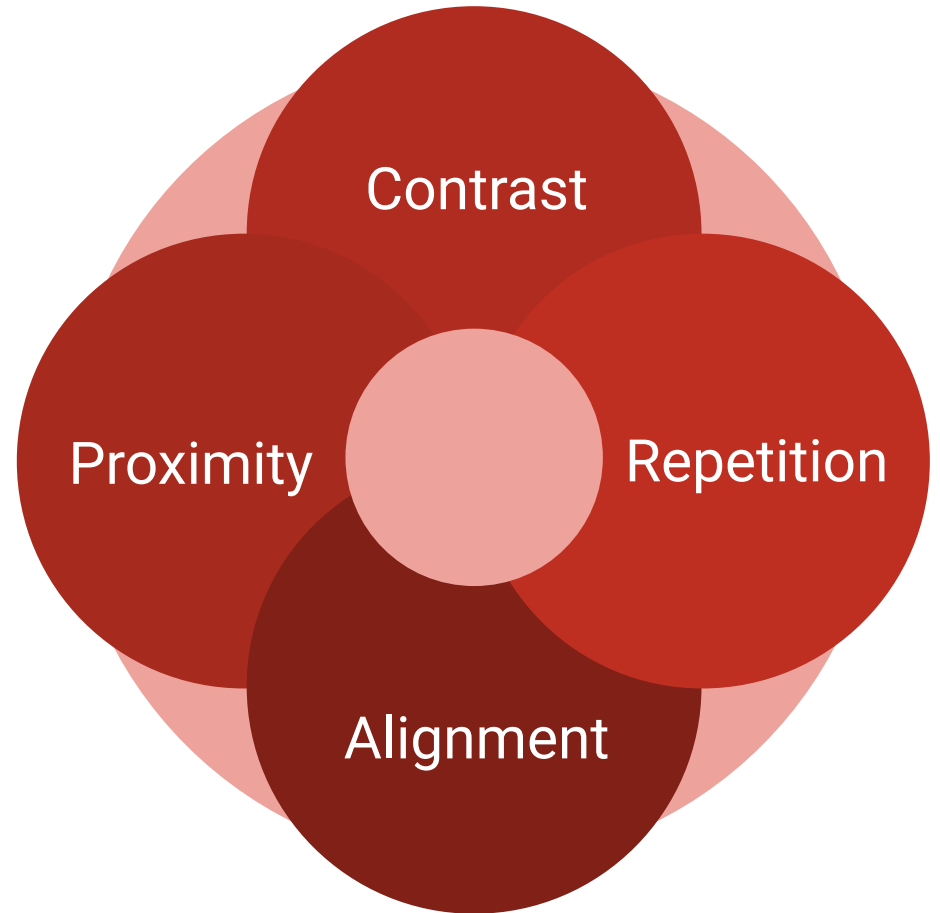
The mental representation of language information. Provides support for reading, writing, understanding, and logical thought (Johnson, 2014).

Filters for the Brain

Structured presentation can help us scan and extract desired information more quickly and easily (Johnson, 2014).

Our vision and perception has evolved to see and seek structure. Gestalt's principles help describe what does or does not draw attention, which is a crucial consideration for preattentive processing and ease of search (Koffka, 1935).

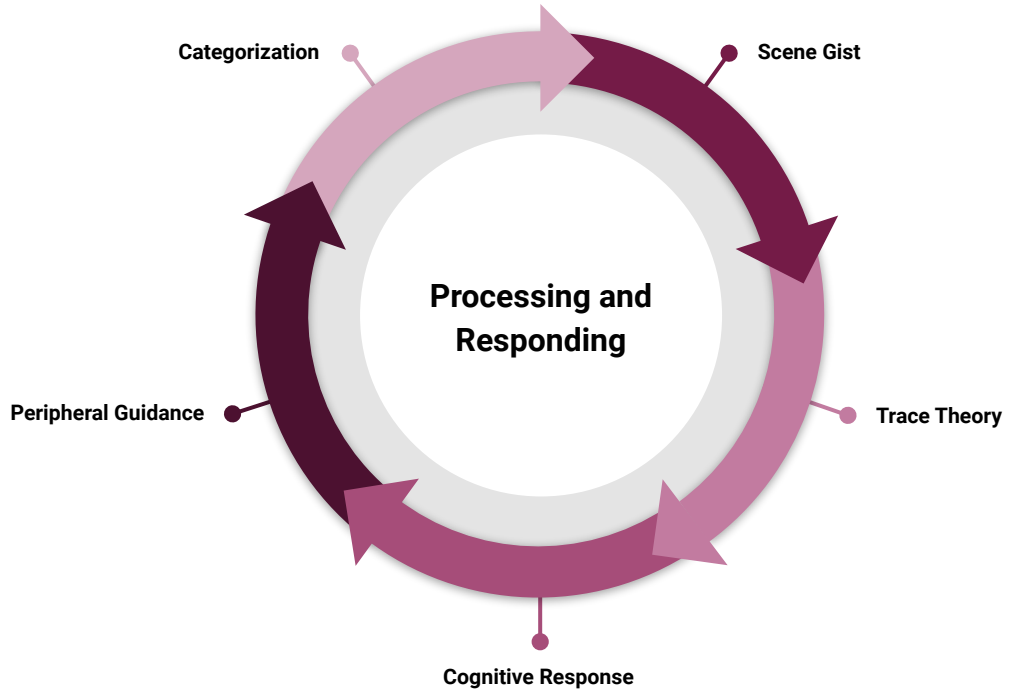
Anything the brain processes faster than 10 msec per item is considered preattentive (Treisman & Gormican, 1988), which helps determine the priority of the information to fixate your attention to.



Seeing is a Pattern

Seeing is a series of repeating cognitive processes that constructs the eye movement control loop, which consists of eye movements, information extraction, perception, interpretation, and planning the next eye movement (Ware, 2012).

To prepare for the next eye movement, the primary visual cortex employs a heuristic strategy by focusing on the next visually strong target (Wolfe & Gancarz, 1996).

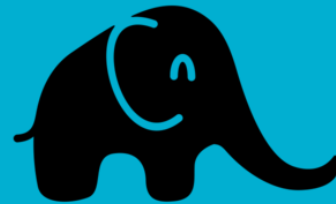


Reducing Cognitive Load

- Technical communicators have a responsibility to reduce the cognitive load for user.
- Technical communication is our pursuit for **simplifying the processes of perceiving and understanding information** – perhaps even on an instinctual level.



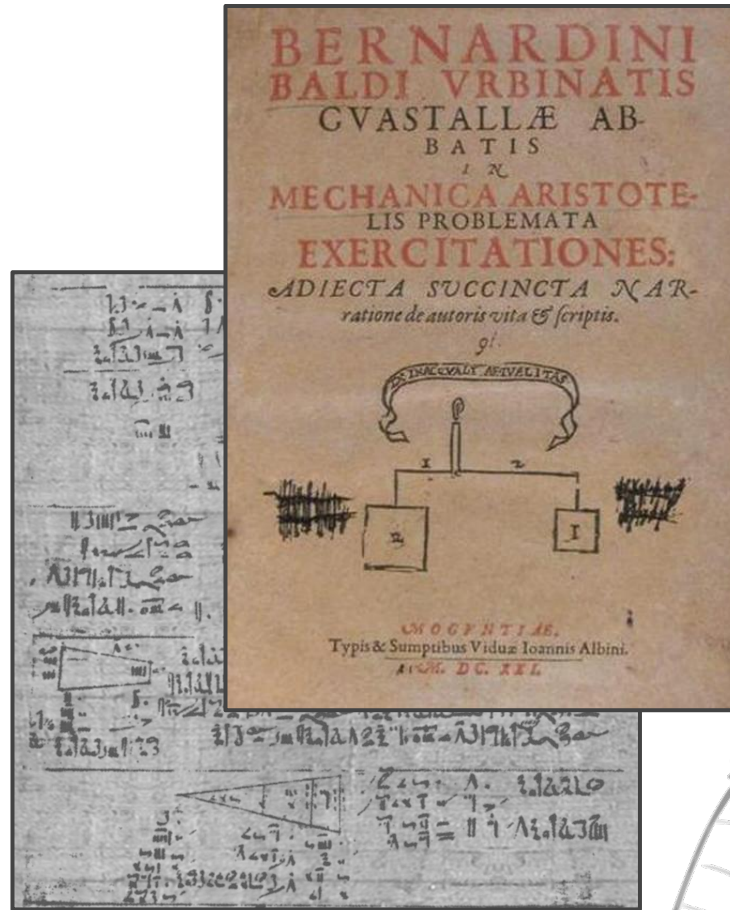
Instinct and Pursuit



The First Instructional Texts

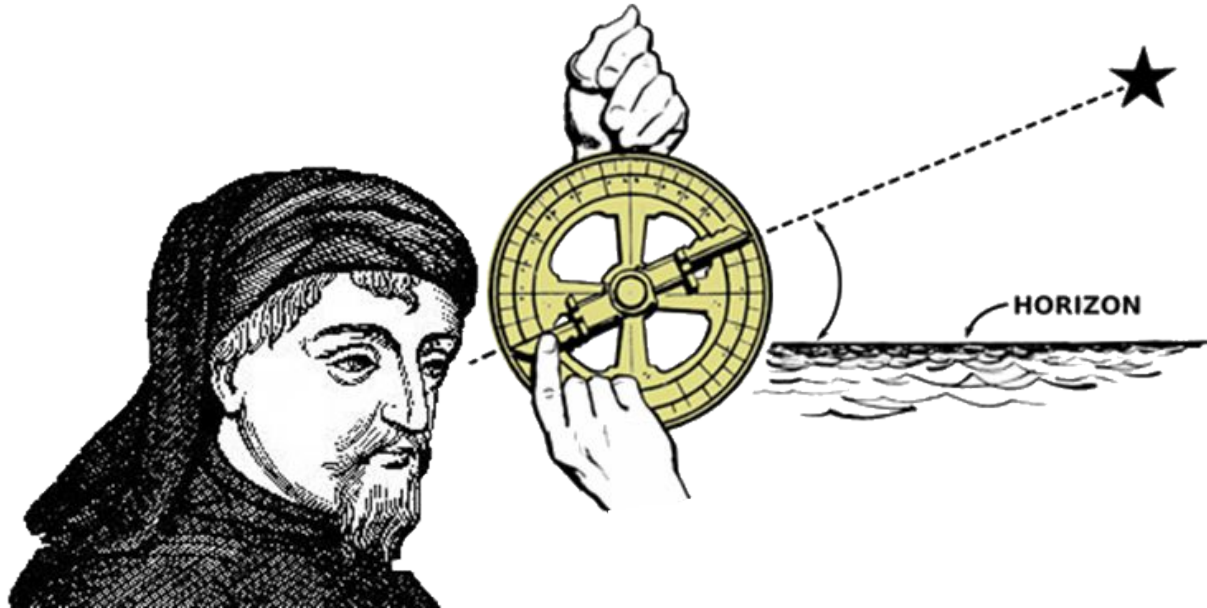
There has always been a need to instruct people on how to do practical things, including:

- Ancient Egyptian medical and mathematical texts
- Certain Ancient Greek philosophical works
- Roman engineering texts
- Medieval instructional publications



The First Instruction Manual in English

In English, the first instruction manual is considered to be Chaucer's "A Treatise on the Astrolabe".



Self-improvement through Instruction

In the Western medieval period the vast majority of texts produced were religious in nature (sermons, prayers, the lives of saints, etc.).

From then until the early 20th century, manuals tended to focus on techniques, often including aspects of self-improvement.

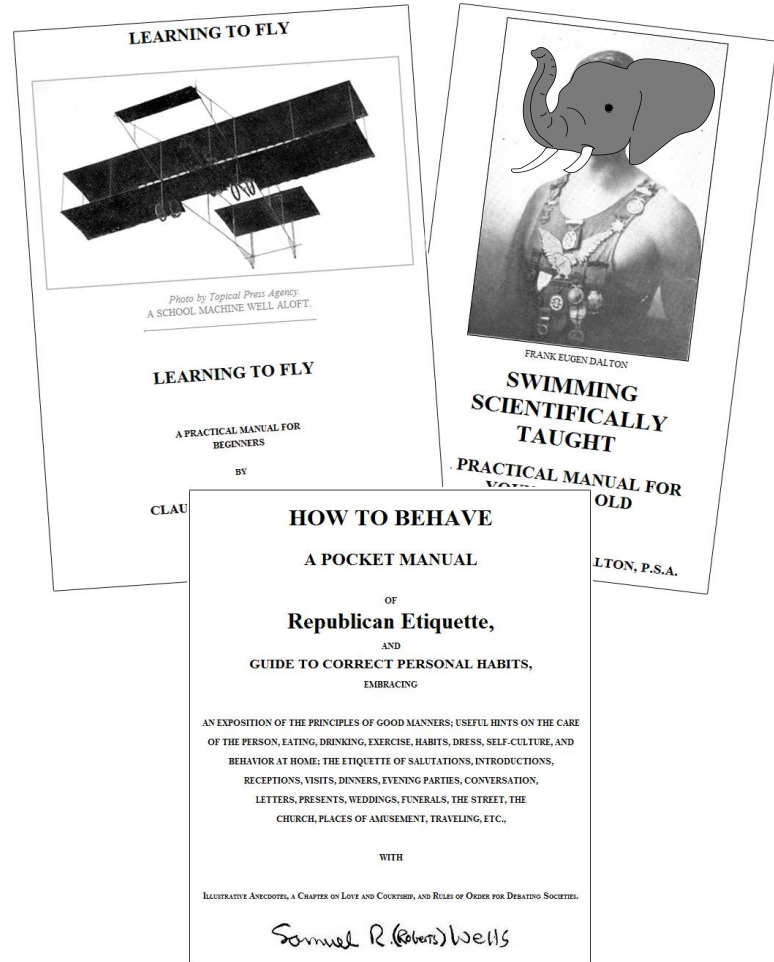
This paved the way for texts that taught other techniques—not just medical, mathematical, or engineering—including subjects like cooking and combat.



19th and Early 20th Century Instructional Texts

Instructional texts became much more prevalent beginning in the Victorian era at the dawn of industrialization.

Tended to contain lots of anecdotes, were often prescriptively moralistic, and came with suggestions on how to become a better person who could fight/fly/swim/behave.

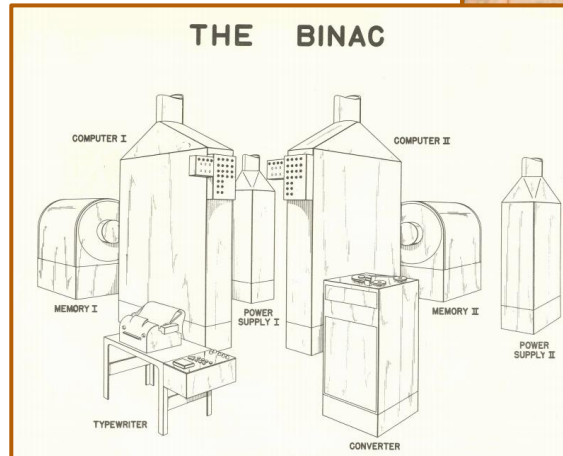
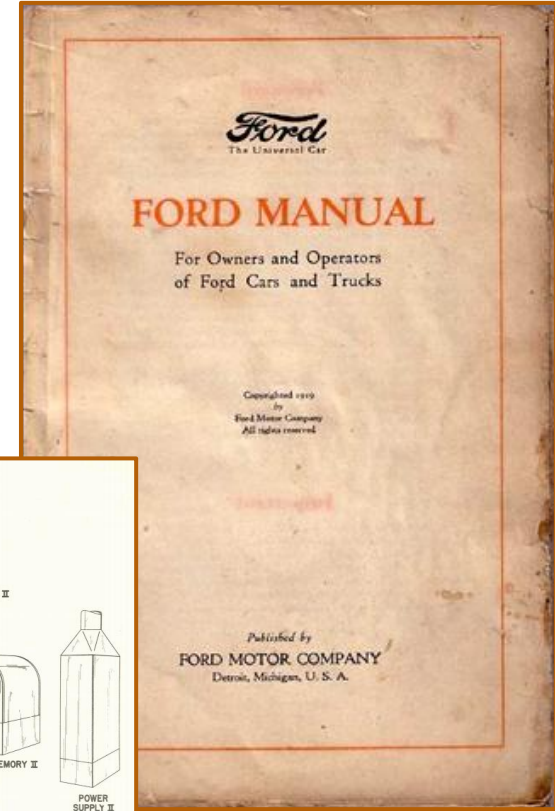


Technology of Documentation



Technical Writing and the 20th Century

- Technical writing as we know it started in early-mid 20th century; focused on instructing people how to work with technology
 - Ford Model T manual (1919) is a good example of this
 - Joseph Chapline wrote first computer manual for BINAC (1949)
- Back when technical documentation could best be characterized as a cottage industry, manuals were typically created as one-offs, written by experts and usually aimed at experts
- They tended to idiosyncratic and often one-offs



Beginning of Structured Authoring

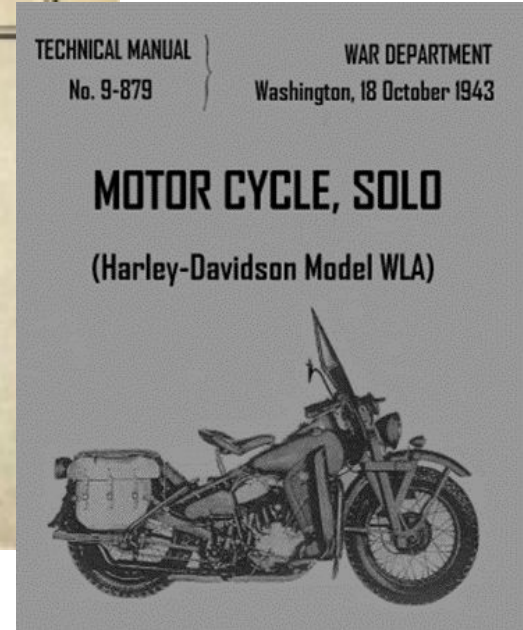
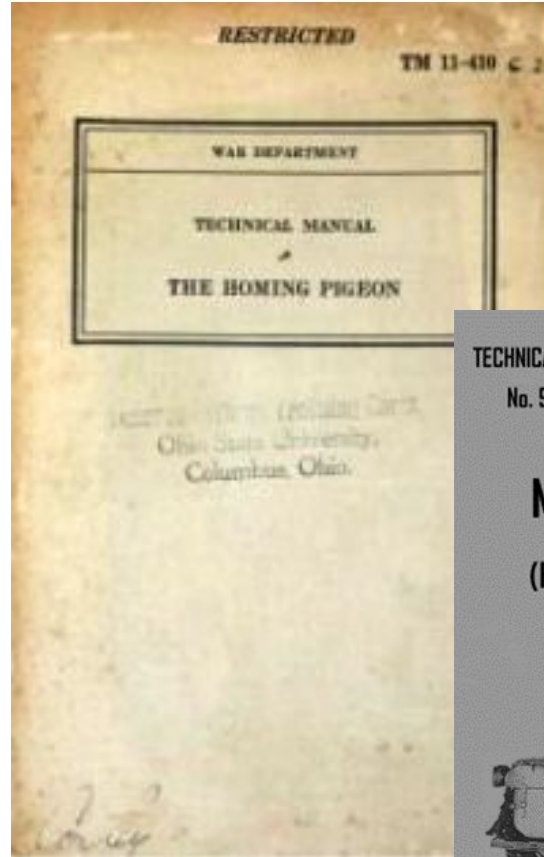
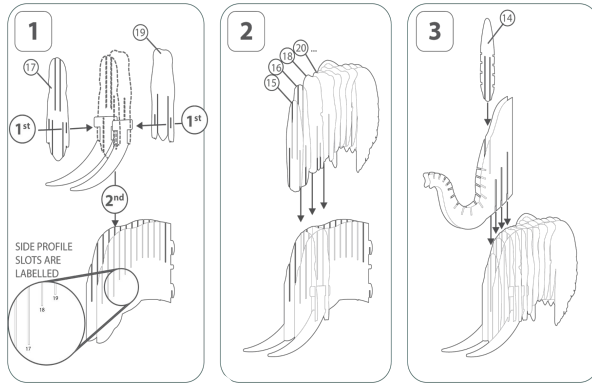
- Much of structured content ultimately derives from studies done in the U.S. Army and Navy from the 1950's and 1960's seeking more effective documentation.
- When seeking information quickly, users do not want a story (narrative), they want just the information they need to do the job.
- Typed, modular information helps accomplish this.



Rise of the Technical Writer

- This changed with the advent of consumer products and military equipment too complex to be understood at a glance.
- Non-SME writers began to enter the mix, and technical writing was recognized as a distinct profession during WWII.

Small Elephant Assembly



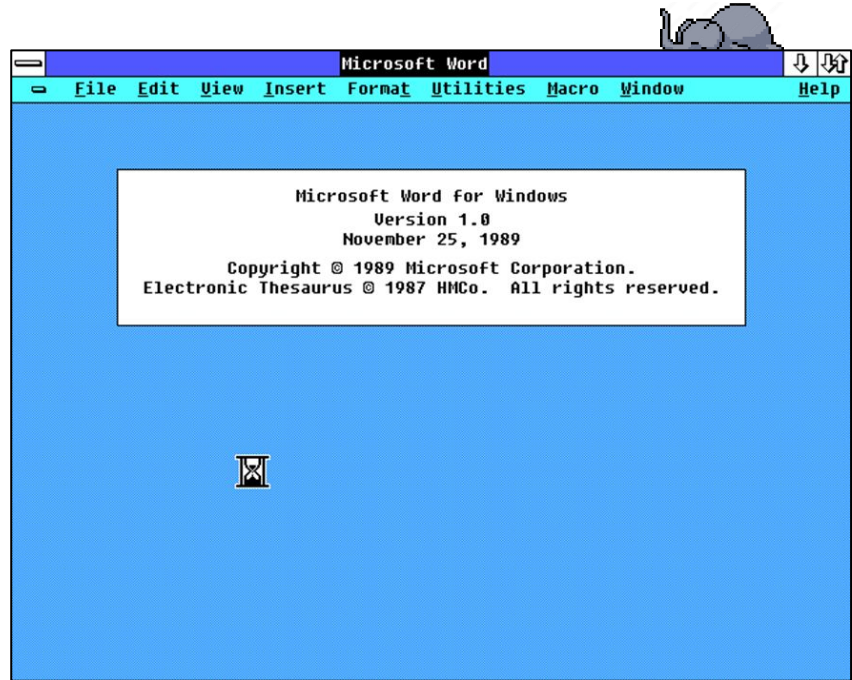
Software and Growth in Technical Documentation

- Growth of technical writing developed alongside growth of software industry.
- DeskTop Publication (DTP) programs made it possible to writers to create documents.
- More documents were produced, often in volume.
- Back in the day, software documentation was considered the "bulk" of the product; it made up most of the box that shipped with the product.



The Rise of Desktop Publishing (and Unstructured Content)

- Non-structured, but allowed for the creation of content quickly and easily.
- Over time, it was also possible to output content to various formats, including PDF and Web .
- This is the beginning of the unstructured vs. structured content dichotomy: ease-of-use + low overhead vs significant value add but higher upfront processing.



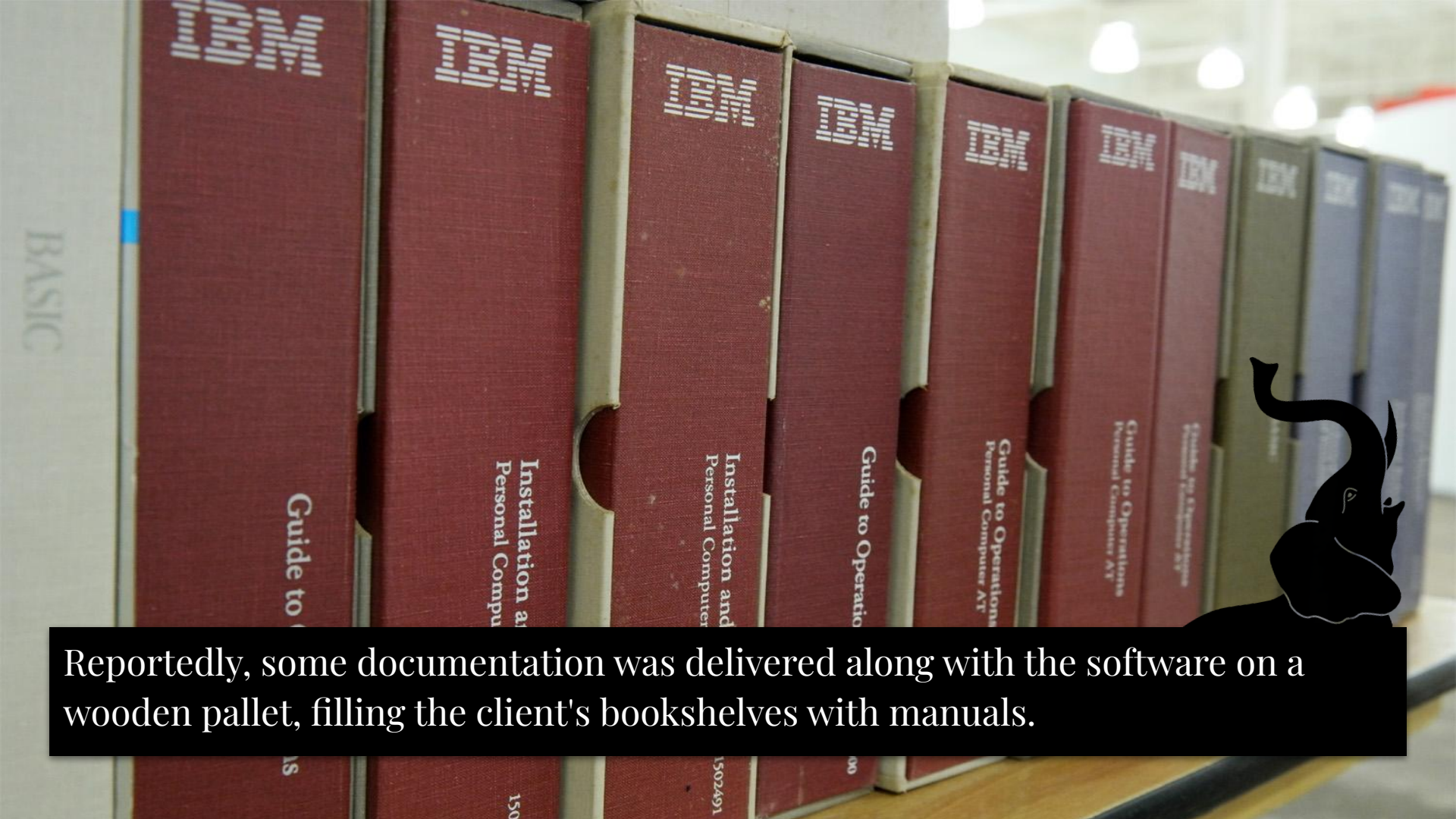
The Author as a Young Tech Writer

This is where Keith enters the picture.

He originally helped create manuals that were printed and put in a big box.



Keith from Delrina Company Photo



Reportedly, some documentation was delivered along with the software on a wooden pallet, filling the client's bookshelves with manuals.

The Weight of Documentation

In the 1980s, A Boeing 747 required the following in terms of documentation:

- 129,000 parts, each requiring extensive documentation
- 31,000,000 pages of docs
- @ 10 minutes per page, would take a single technical writer 216,000 24hr days to write

- Total weight of 31 Million pages of docs: 450 tons
- Total weight of a Boeing 747: 437.5 tons

- With digitization and PDFs, this same amount of information could fit on a few CD-ROMs
- However, there was still the task of organizing this info...



90s-era IBM Faced Two Documentation Issues

- In the 90s, IBM documentation teams were facing two main issues:
 - Content needed to be tailored for the web
 - Too many output formats, needing lots of tools
- IBM opted to create an XML-based, topic-typed documentation standard, called "DITA"
 - 1 Topic = 1 Web page
 - Made it an open standard so that they could benefit from sharing info with their business partners
- DITA Open Toolkit opened up multi-channel publishing

Structured authoring, when done right, focuses on the needs of the reader, providing information when and where needed.

TECHNOLOGY REVIEW



Deborah S. Ray
and Eric J. Ray
Editors

DITA: An XML-based Technical Documentation Authoring and Publishing Architecture

Michael Priestley, Gretchen Hargis, and Susan Carpenter

This column examines emerging technologies of interest to technical communicators to help them identify those that are worthy of further investigation. It is intended neither as an endorsement of any technology or product, nor as a recommendation to purchase. The opinions expressed by the column editors are their own and do not represent the views of the Society for Technical Communication. All URLs and site contents were verified at the time of writing.

The Darwin Information Typing Architecture (DITA) is a technical documentation authoring and publishing architecture that is based on principles of modular reuse and extensibility. This article discusses how DITA affects how we write, how we design, and how we process technical documentation, and what benefits the DITA approach can deliver that traditional documentation strategies cannot.

Over the past few years, XML (Extensible Markup Language) has gained popularity in the technical writing profession by offering us a

logical and fairly straightforward framework for developing structured information. For technical communicators, XML promises capabilities to separate form from content; to use specific, customized markup to describe content; and to use a standard solution without depending on proprietary tools or formats. The promised result of XML is documentation that is reusable in any medium, useful for specialized tools and for our customers, and interchangeable without depending on a particular authoring environment.

XML in and of itself, however, has not, to date, been a panacea in our quest to achieve these goals; instead, we often still struggle to develop processes that realize the potential of XML. In this article, we introduce the Darwin Information Typing Architecture (DITA), which provides technical communicators with an XML-based architecture for authoring, producing, and delivering technical information.

As you'll see, DITA goes further than other currently available solutions by allowing us to easily create highly specialized structure and content, yet still retain interchangeability

and reuse of the content and process. As a result, DITA helps solve current problems in information development, including those of information reuse and information delivery in multiple media (single-sourcing), and helps us maximize the potential of XML for technical communicators.

In the following sections, we provide a brief overview of terms and concepts; describe the promise of XML and its shortcomings; and describe how DITA addresses content-, design-, and process-related problems.

A BRIEF INTRODUCTION TO MARKUP LANGUAGES AND XML

If you are already familiar with XML DTDs and XSLT as used for documentation, you can skip to the next section. Otherwise, read on for a brief introduction to the principles of markup languages in general, and XML and related standards in particular.

A *markup language* is a set of start and end tags you can use to "mark up" text with additional information about your content—for example, `<xmp>` the xmp tag set tells processes that this text is part of an example `</xmp>`. This information can be used for

- ◆ Displaying the text, to apply different fonts and styles to different types of information
- ◆ Processing the text, to extract particular subsets of the information for particular uses
- ◆ Searching the text for particular kinds of information

XML is a standard for defining markup languages. XHTML (Extensible HTML) is an example of an XML-compliant markup language, as are WML (Wireless Markup Language) and DocBook. XML is a streamlined version of SGML (Standard Generalized Markup Language), an older and broader standard for defining markup languages.

New Modes for Old Brains



Macro Patterns



**Systematic
Labelling**

**Modular,
Topic-based
Architecture**

+

**Constrained
Writing
Environments**

**Separation of
Content and
Form**

**Structured
Authoring**

A standardized, methodological approach to content creation.

*The DITA Style Guide:
Best Practices for Authors
(Self, 2011)*

Micro Patterns



Jakob Nielsen, 1998

Small groups of words that can be skimmed by a person to get a clear idea of the content of a Web page. This included article headlines, page titles, subject lines and e-mail headings. Such phrases also may be taken out of context and displayed on a directory, search result page, bookmark list, etc.

Anil Dash, 2002

Small information chunks that can stand alone or be used in a variety of contexts, including instant messages, blog posts, RSS feeds, and abstracts.



Rob Hanna, 2017

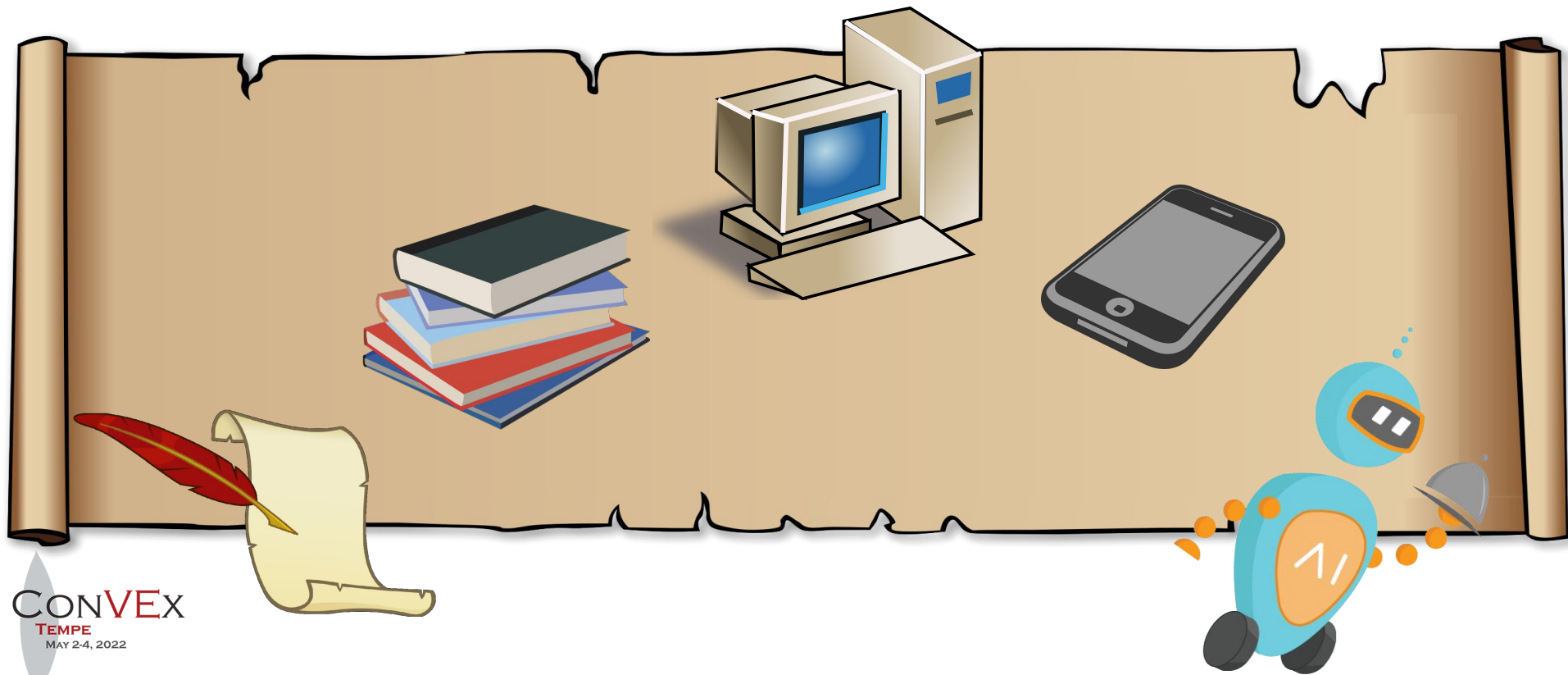
Content that is about one primary idea, fact, or concept, it is easily scannable, labeled for clear identification and meaning, and appropriately written and formatted for use anywhere and any time it is needed.

“The goal is not to produce clear-concise simple information, but rather to produce information within the situation that is easy to use.”

— *Dr. Alex Vakaloudis*



From Papyrus to Smart Environments



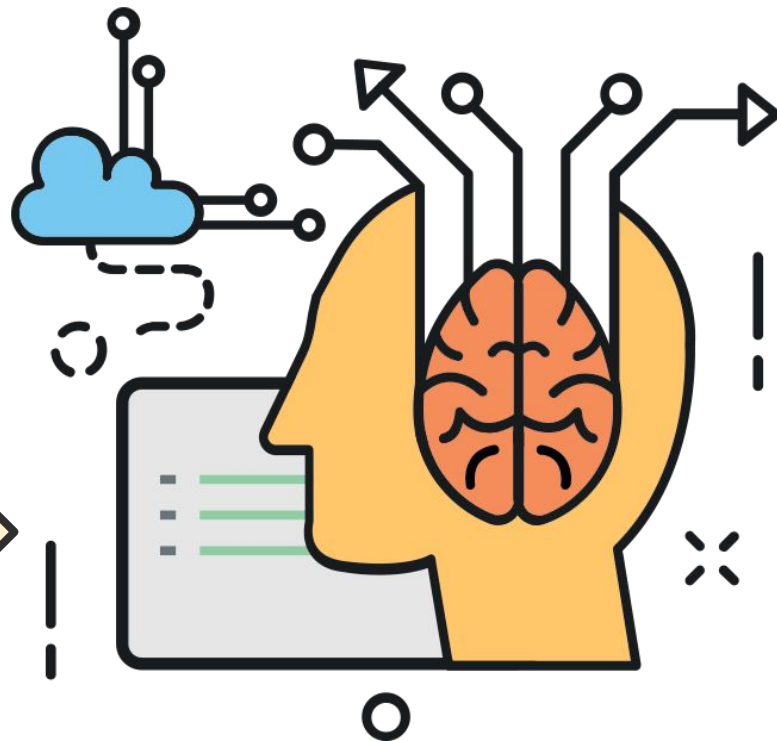
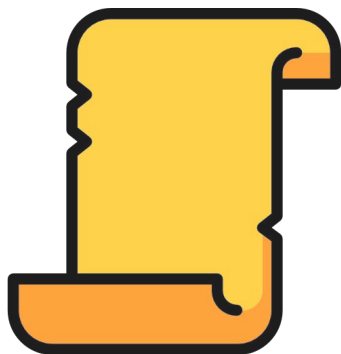
CONVEX
TEMPE
MAY 2-4, 2022

CIDM

Beyond the Point of User's Cognitive Saturation

“Simple problems have become automated to the point that only the most critical decisions are presented to the human for supervisory decisions. Each decision a human makes carries higher weight, and mistakes tend to have greater consequences. The ability of an automated system... **depends greatly on communicating appropriate data to the user.**”

— Vakaloudis, 2020



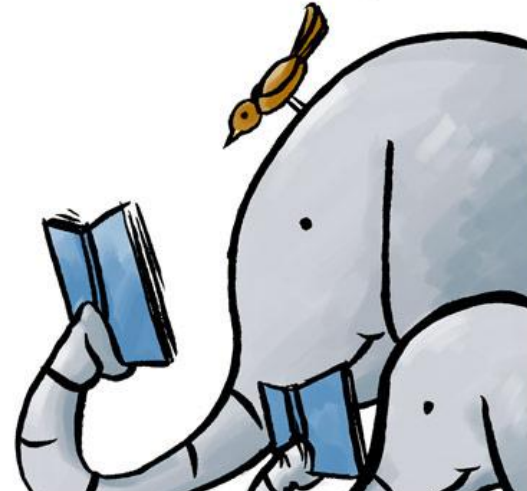
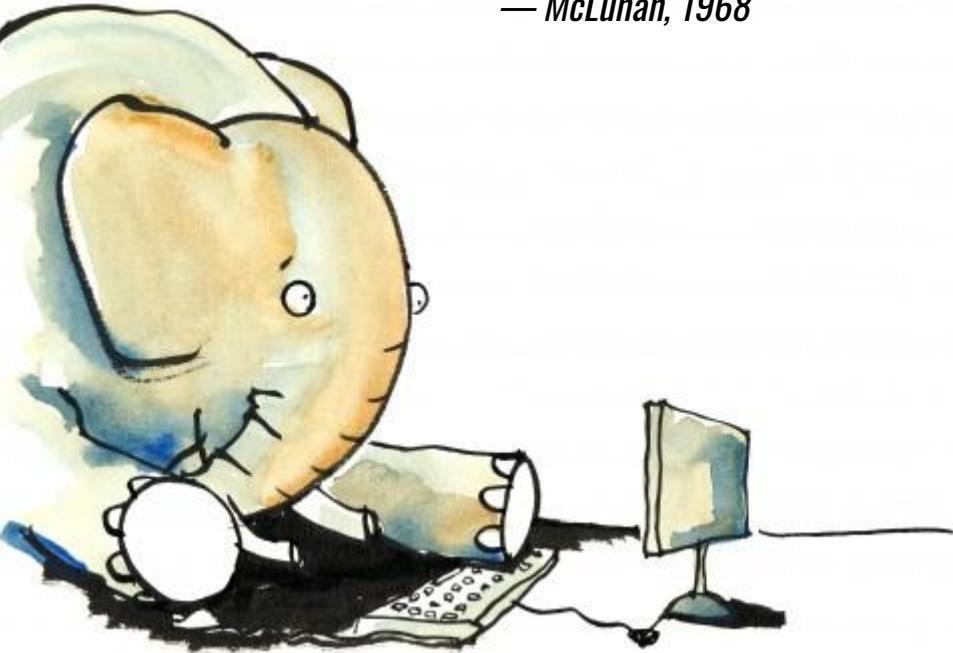
The Reader's Abilities Have Not Changed

“When you give people too much information, they instantly resort to pattern recognition.”

— *McLuhan, 1968*

Thus we seek out only patterns we recognize and—can label.

No matter what point in history we are at, our solutions are always human-focused.





There is actually only ONE elephant...

...and it's still in the room.

From Papyrus to the Virtual Elephants in Rooms



Joyce (Jo) Lam

“Enable Information”

joyce.lam@enableinfo.com



Keith Schengili-Roberts

“DITAWriter”

keith@ditawriter.com