

The Science of Information Delivery

Robert Dale, CTO, Arria NLG

Summer School on NLG, Summarisation, and Dialogue Systems | 24th July 2015

© 2015 ARRIA NLG plc.

The Take-away

- Natural Language Generation is just a means to an end
- That end is Information Delivery
- NLG practitioners are really Information Delivery practitioners



Agenda

- What is Information Delivery?
- The Elements of a Science of Information Delivery

Caveats

- Primary focus is on text rather than voice
- Orientation is towards engineering rather than cognitive modelling
- Disposition is towards **precision NLG**

Bookcases

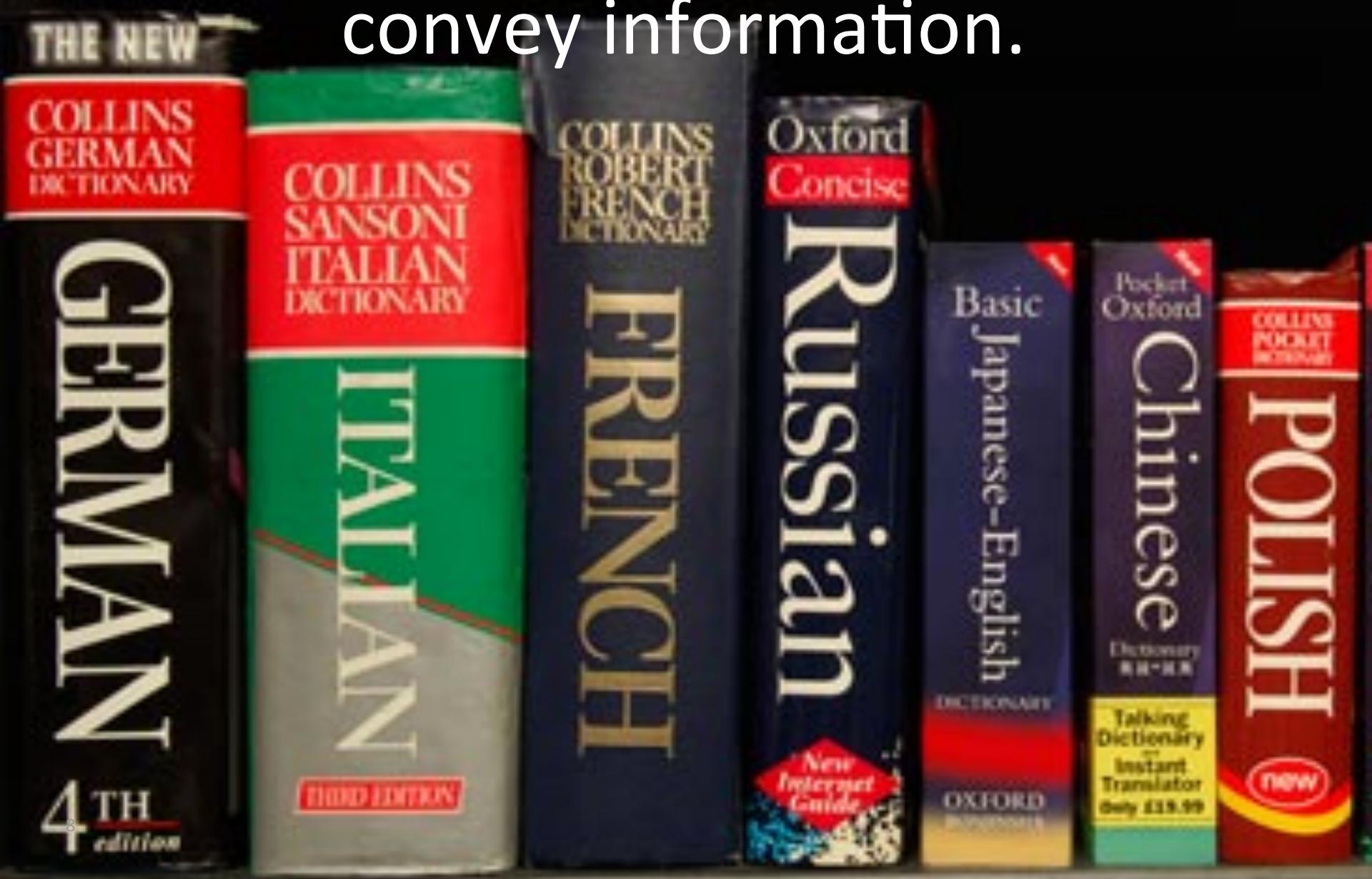


What is Information Delivery?

A Definition

- Information delivery is the process of providing the right information to the right audience at the right time and in the right manner

Language is not the only way to convey information.





If the only tool you have
is a hammer you tend to
see every problem as a nail

- Abraham Maslow

Modalities

- A modality is a particular way in which information is encoded for presentation
- The sensory modalities:



Visual



Auditory



Gustatory

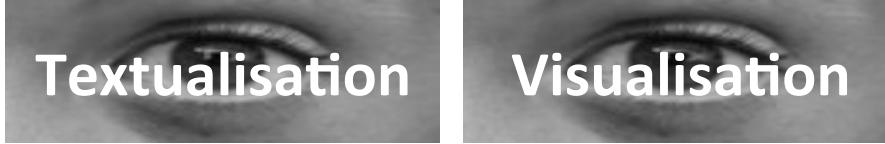


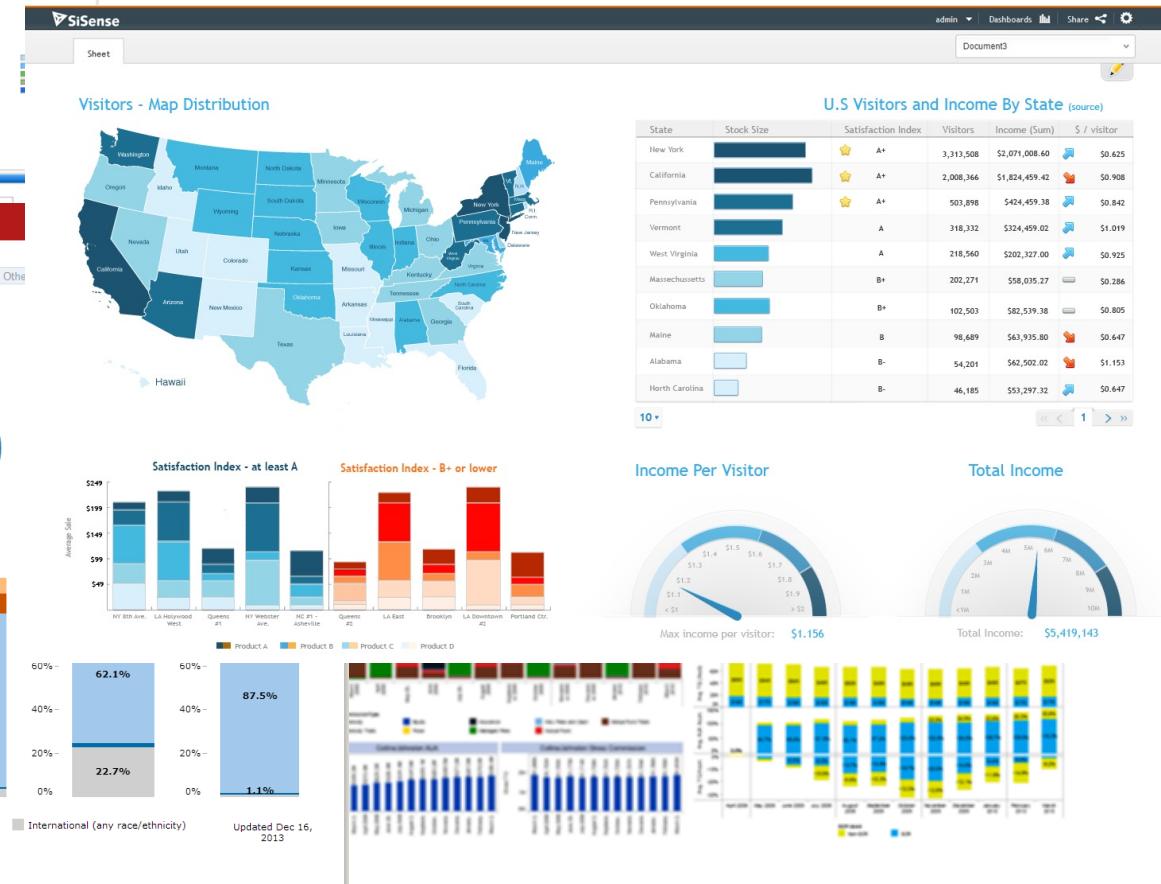
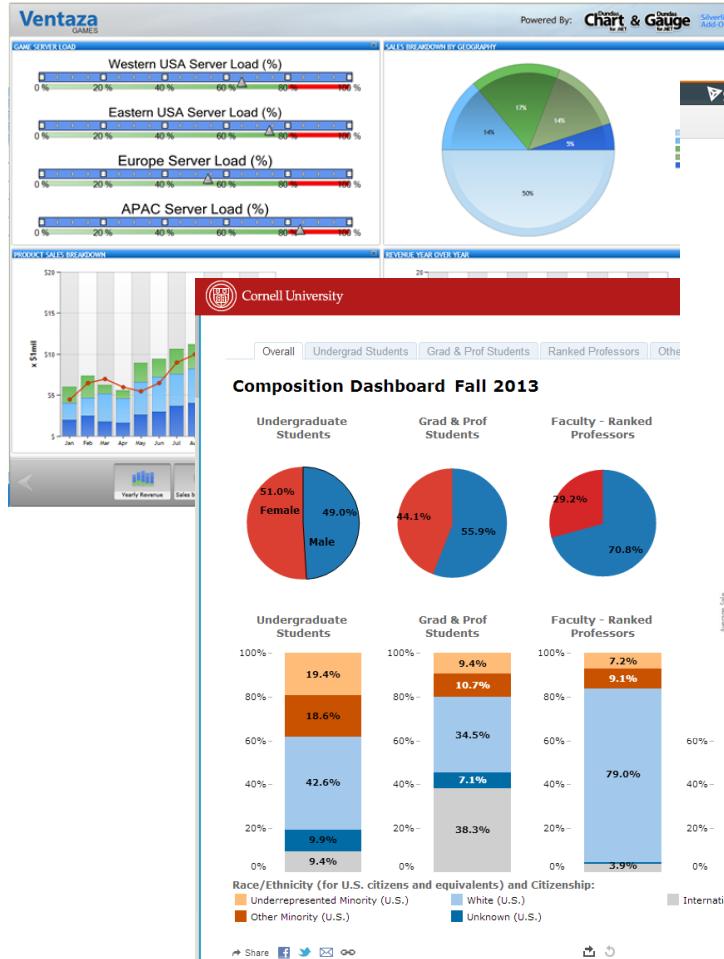
Olfactory



Tactile

NLG's Nearest Neighbour





An Old Meme

**One Look Is Worth
A Thousand Words--**

One look at our line of Republic, Firestone, Miller and United States tires can tell you more than a hundred personal letters or advertisements.

**WE WILL PROVE THEIR VALUE
BEFORE YOU INVEST ONE DOLLAR
IN THEM.**

Ever consider buying supplies from a catalog?

What's the use! Call and see what you are buying. One look at our display of automobile and motorcycle accessories will convince you of the fact.

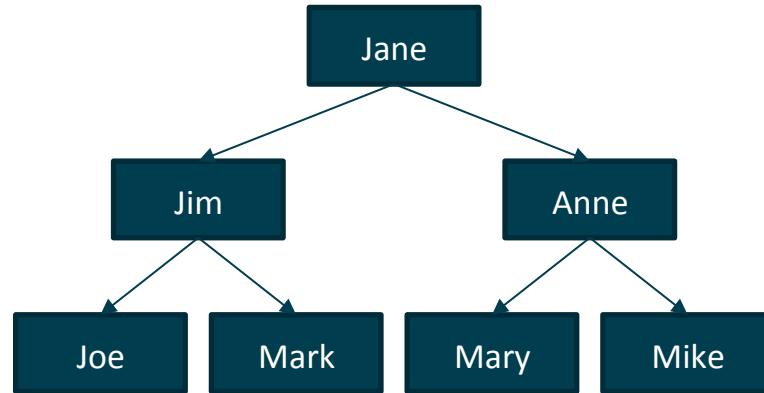
**THAT WE HAVE EVERYTHING FOR
THE AUTO**

Piqua Auto Supply House

133 N. Main St.—Piqua, O.

Who Works for Who?

- Jane is Jim's boss.
- Jim is Joe's boss.
- Anne works for Jane.
- Mark works for Jim.
- Anne is Mary's boss.
- Anne is Mike's boss.





Choosing Text or Graphics

- Language is the most elaborate, complete, and widely shared system of symbols that we have available
- Images are better for:
 - spatial structures
 - location
 - detail
- Words are better for representing:
 - procedural information
 - logical conditions
 - abstract verbal concepts
- The general view: providing information in more than one medium of communication will lead to better understanding

Recent research shows the integration of verbal and visual elements is better than text separated from visual elements as follows

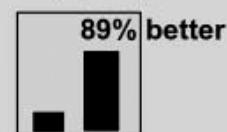
in one group of studies
adding visuals to words improved learning by 23%

Mayer, 2001, 74



in another group of studies
adding visuals to words improved transfer of learning by 89%

Mayer, 2001, 76

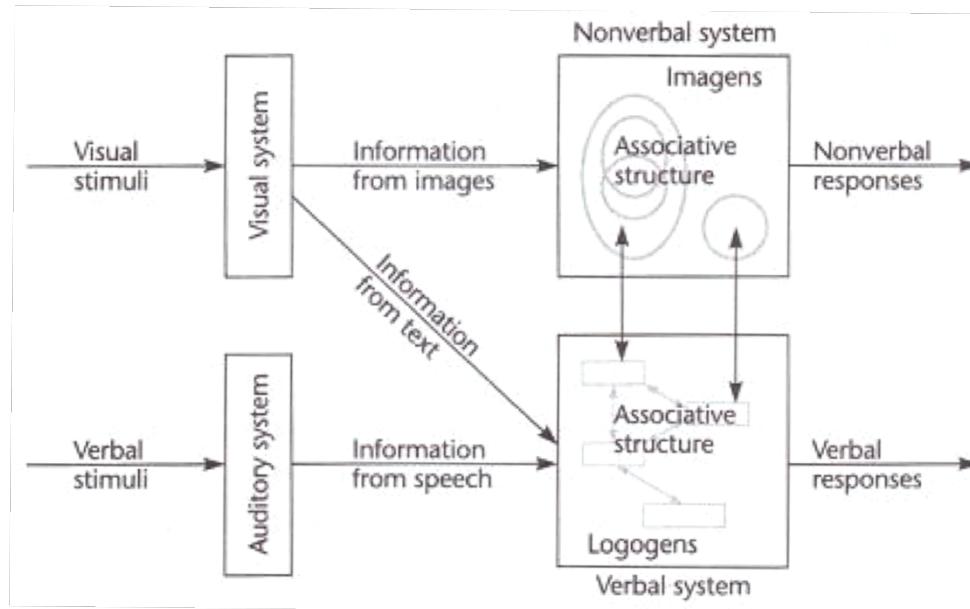


in other words integrated visual-verbal language
improved communication and learning by large, measurable amounts

Mayer, Richard E. 2001, Multimedia Learning, Cambridge, Cambridge University Press

Dual Coding Theory

- Visual and verbal information is stored in different systems having different characteristics



Other Pragmatic Considerations

Some quick arguments for text over graphics:

Language ...

- downloads and renders faster
- is easy to abstract to a database or CMS
- is searchable and indexable
- can be selected, copied and pasted
- is portable across platforms

A picture
is worth a
thousand words.
but it uses up
three thousand
times the memory

As the Chinese say,
1001 words
is worth more than a picture.
– John McCarthy



Language is Always Embodied

Visual Context

Textualisation

Text is a Visual Information Channel

- Formatting and typography matters
- Font choice matters
- Line length matters
- Point size matters

CPA this week was 16.0% higher than the benchmark at \$35.95 and has seen an increase of \$9.78 (up 37.4%) since last week. Since the last reporting period, the campaign has seen an increase in CPA of \$9.20 (up 34.4%).

CPA was 16.0% higher than the benchmark at \$35.95 this week.

- Since the last reporting period, the campaign has seen an increase in CPA of \$9.20 (up 34.4%).
- Since last week, the campaign has seen an increase in CPA of \$9.78 (up 37.4%).

Acme Insurance Advertising Campaign

Periodic Report for July 27, 2014

Observations

CPA was 17.8% lower than the benchmark at

\$25.47 this week and has seen an increase of \$0.73 (up 3.0%) since last week. Since the last reporting period, the campaign has seen a decrease in CPA of \$1.06 (down 4.0%). Spend is up 15.8% to \$69,047.10. Current conversion rate of 10.3% is 2.8% above campaign benchmark.

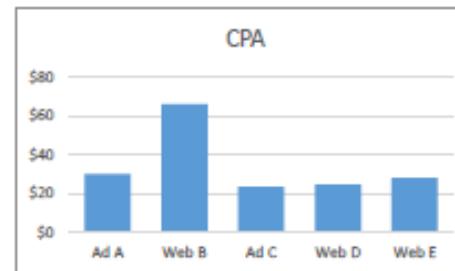
The efficient CPM (\$0.25) resulted in higher volumes of impressions for Website E but has decreased the conversion rate. Website E generated a CPA of \$21.73.

86.2% of visitors left the Acme site after viewing only the home page.

Highlights

- Website D is the best performing site with a CPA of \$21.75 and a conversion rate of 9.40%.
- Website B generated the highest amount of traffic to the website, accounting for 55.6% of the traffic overall.
- Viewers from Website E viewed the most number of pages (2.6 on average).

Overall Campaign Performance To Date				
Total Conversions	CR	CPA	Spend	CTR
51,734	9.47%	\$29.21	\$1,511,170	0.05%



Things that could be improved

- Website B is the worst performing site with a CPA of \$47.63 and a conversion rate of 5.90%.
- Viewers from Ad Partner C viewed the least number of pages (1.5 on average).
- Ad Partner C generated the lowest amount of traffic to the website, accounting for only 1.0% of traffic overall.

Possible Optimisations

Site Level: Across sites, Website B is performing poorly in terms of CPA at \$47.63. Other things being equal, directing that expenditure to Website D would result in a CPA of \$25.21 (down \$1.66) and a conversion rate of 12.19% (up 1.86%).

Tactics within Sites: Other things being equal, reallocating resources from Ad Partner C's Look-a-Like Targeting to their CPA Program would result in a CPA of \$18.56 (down \$4.52) and a conversion rate of 53.09% (up 30.19%).

Creatives within Tactics: Website D's Site Retargeting using Home creatives is not performing well. Moving these resources to their Site Retargeting using Many Ways Icon would result in a CPA of \$13.52 (down \$0.58) and a conversion rate of 32.09% (up 4.25%).

Creative Sizes within Tactics: Ad Partner A's Contextual Filtering using Home at 728x90 is not performing well. Moving these resources to their Contextual Filtering using Home at 300x600 would result in a CPA of \$4.24 (down \$32.96) and a conversion rate of 42.29% (up 40.04%).

Acme Insurance Advertising Campaign Periodic Report for July 27, 2014. Observations: CPA was 17.8% lower than the benchmark at \$25.47 this week and has seen an increase of \$0.73 (up 3.0%) since last week. Since the last reporting period, the campaign has seen a decrease in CPA of \$1.06 (down 4.0%). Spend is up 15.8% to \$69,047.10. Current conversion rate of 10.3% is 2.8% above campaign benchmark. The efficient CPM (\$0.25) resulted in higher volumes of impressions for Website E but has decreased the conversion rate. Website E generated a CPA of \$21.73. 86.2% of visitors left the Acme site after viewing only the home page. Highlights: Website D is the best performing site with a CPA of \$21.75 and a conversion rate of 9.40%. Website B generated the highest amount of traffic to the website, accounting for 55.6% of the traffic overall. Viewers from Website E viewed the most number of pages (2.6 on average). Things that could be improved: Website B is the worst performing site with a CPA of \$47.63 and a conversion rate of 5.90%. Viewers from Ad Partner C viewed the least number of pages (1.5 on average). Ad Partner C generated the lowest amount of traffic to the website, accounting for only 1.0% of traffic overall. Possible Optimisations: Site Level: Across sites, Website B is performing poorly in terms of CPA at \$47.63. Other things being equal, directing that expenditure to Website D would result in a CPA of \$25.21 (down \$1.66) and a conversion rate of 12.19% (up 1.86%). Tactics within Sites: Other things being equal, reallocating resources from Ad Partner C's Look-a-Like Targeting to their CPA Program would result in a CPA of \$18.56 (down \$4.52) and a conversion rate of 53.09% (up 30.19%). Creatives within Tactics: Website D's Site Retargeting using Home creatives is not performing well. Moving these resources to their Site Retargeting using Many Ways Icon would result in a CPA of \$13.52 (down \$0.58) and a conversion rate of 32.09% (up 4.25%). Creative Sizes within Tactics: Ad Partner A's Contextual Filtering using Home at 728x90 is not performing well. Moving these resources to their Contextual Filtering using Home at 300x600 would result in a CPA of \$4.24 (down \$32.96) and a conversion rate of 42.29% (up 40.04%).

Acme Insurance Advertising Campaign Periodic Report for July 27, 2014.

Observations:

CPA was 17.8% lower than the benchmark at \$25.47 this week and has seen an increase of \$0.73 (up 3.0%) since last week. Since the last reporting period, the campaign has seen a decrease in CPA of \$1.06 (down 4.0%). Spend is up 15.8% to \$69,047.10. Current conversion rate of 10.3% is 2.8% above campaign benchmark. The efficient CPM (\$0.25) resulted in higher volumes of impressions for Website E but has decreased the conversion rate. Website E generated a CPA of \$21.73. 86.2% of visitors left the Acme site after viewing only the home page.

Highlights:

Website D is the best performing site with a CPA of \$21.75 and a conversion rate of 9.40%. Website B generated the highest amount of traffic to the website, accounting for 55.6% of the traffic overall. Viewers from Website E viewed the most number of pages (2.6 on average).

Things that could be improved: Website B is the worst performing site with a CPA of \$47.63 and a conversion rate of 5.90%. Viewers from Ad Partner C viewed the least number of pages (1.5 on average). Ad Partner C generated the lowest amount of traffic to the website, accounting for only 1.0% of traffic overall.

Possible Optimisations:

Site Level: Across sites, Website B is performing poorly in terms of CPA at \$47.63. Other things being equal, directing that expenditure to Website D would result in a CPA of \$25.21 (down \$1.66) and a conversion rate of 12.19% (up 1.86%).

Tactics within Sites: Other things being equal, reallocating resources from Ad Partner C's Look-a-Like Targeting to their CPA Program would result in a CPA of \$18.56 (down \$4.52) and a conversion rate of 53.09% (up 30.19%).

Creatives within Tactics: Website D's Site Retargeting using Home creatives is not performing well. Moving these resources to their Site Retargeting using Many Ways Icon would result in a CPA of \$13.52 (down \$0.58) and a conversion rate of 32.09% (up 4.25%).

Creative Sizes within Tactics: Ad Partner A's Contextual Filtering using Home at 728x90 is not performing well. Moving these resources to their Contextual Filtering using Home at 300x600 would result in a CPA of \$4.24 (down \$32.96) and a conversion rate of 42.29% (up 40.04%).

Acme Insurance Advertising Campaign

Periodic Report for July 27, 2014

Observations

CPA was 17.8% lower than the benchmark at

\$25.47 this week and has seen an increase of \$0.73 (up 3.0%) since last week. Since the last reporting period, the campaign has seen a decrease in CPA of \$1.06 (down 4.0%). Spend is up 15.8% to \$69,047.10. Current conversion rate of 10.3% is 2.8% above campaign benchmark.

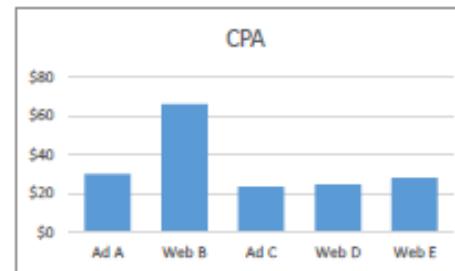
The efficient CPM (\$0.25) resulted in higher volumes of impressions for Website E but has decreased the conversion rate. Website E generated a CPA of \$21.73.

86.2% of visitors left the Acme site after viewing only the home page.

Highlights

- Website D is the best performing site with a CPA of \$21.75 and a conversion rate of 9.40%.
- Website B generated the highest amount of traffic to the website, accounting for 55.6% of the traffic overall.
- Viewers from Website E viewed the most number of pages (2.6 on average).

Overall Campaign Performance To Date				
Total Conversions	CR	CPA	Spend	CTR
51,734	9.47%	\$29.21	\$1,511,170	0.05%



Things that could be improved

- Website B is the worst performing site with a CPA of \$47.63 and a conversion rate of 5.90%.
- Viewers from Ad Partner C viewed the least number of pages (1.5 on average).
- Ad Partner C generated the lowest amount of traffic to the website, accounting for only 1.0% of traffic overall.

Possible Optimisations

Site Level: Across sites, Website B is performing poorly in terms of CPA at \$47.63. Other things being equal, directing that expenditure to Website D would result in a CPA of \$25.21 (down \$1.66) and a conversion rate of 12.19% (up 1.86%).

Tactics within Sites: Other things being equal, reallocating resources from Ad Partner C's Look-a-Like Targeting to their CPA Program would result in a CPA of \$18.56 (down \$4.52) and a conversion rate of 53.09% (up 30.19%).

Creatives within Tactics: Website D's Site Retargeting using Home creatives is not performing well. Moving these resources to their Site Retargeting using Many Ways Icon would result in a CPA of \$13.52 (down \$0.58) and a conversion rate of 32.09% (up 4.25%).

Creative Sizes within Tactics: Ad Partner A's Contextual Filtering using Home at 728x90 is not performing well. Moving these resources to their Contextual Filtering using Home at 300x600 would result in a CPA of \$4.24 (down \$32.96) and a conversion rate of 42.29% (up 40.04%).

Robert Horn's Information Mapping

'Structured writing': specific presentational mechanisms for different information types

The most common information types are:

- Procedure
- Process
- Principle
- Concept
- Fact
- Structure
- Classification

What you'll learn using informationmapping!

Reorganization of the company

As you well know, our company has had considerable difficulty this year with the stalled economy, inflation, with the labor difficulties we've had at several plants, and with the development of our new products, especially in the home products system. This situation has forced the management team to assess our entire company and its operations with a view to finding a better way to organize it for improved profits and long-term efficiency.

Some of our departments have been growing and shrinking without much rhyme or reason, and before this occasion we have not made the effort to take a really hard look at what we were doing. Instead, we were at a standstill, and there will be the aim of eliminating duplication when we could and pulling together groups that belong together functionally.

Now we are announcing a major reorganization to take effect on 18 January. We will announce the details on 12 January including dates when new managers will hold meetings with various employees to whom the information is pertinent. We will also, at that time, distribute a complete schedule setting forth who will be working for whom. In the meantime, we are announcing the following changes so the managers in charge of the affected divisions and departments can prepare for the reorganization.

Janice Moreland will move from Vice President for Research to Vice President for Operations. Jack Spotter will be the new head of the Research Department, moving from his position as Assistant Vice President for Operations. Marilyn Belt will become Director of the new Home Products Division, which used to have only project status. These changes in department managerial positions will take place on 5 January. Current Assistant Directors will remain in their positions at that time unless otherwise notified.

Then, on 12 January, changes at the level of Assistant Directors will go into effect. The staffs of these departments will be informed in an email notification from their managers about whether they will be moving with their current managers or staying in their current departments. In most cases there will be no change, as we are trying to keep as many departments intact, with experienced staff, as possible.

before

Reorganization of the company

Background

As you well know, our company has had considerable difficulty this year with:

- the stalled economy
- inflation
- labour difficulties at several plants, and
- development of our new products, especially in the home-products systems.

The management team decided to reassess our entire company and its operations and find a better way to organize it for improved profits and long-term efficiency.

As of January 5, Home Products will be created as a new division

New division

Management changes

The new management positions are described in this table:

Name	Previous position	New management position
Janice Moreland	VP, Research	VP, Operations
Jack Spotter	Assistant VP, Operations	Director, Research
Marilyn Belt	Assistant to the Plant Manager	Director, Home Products Division (new division)

Notification to staff

The managers of these departments will inform their staffs via email if they will be moving or staying in their current departments. In most cases there will be no change. We are trying to keep departments intact, if possible.

Effective dates

The effective dates for these changes are listed below:

Date	Description
January 5	Changes in Corporate Officers and Division Chiefs.
January 12	Changes in Assistant Directors positions and announcement of details of reorganization.
January 18	Reorganization takes effect.

after

Acme Insurance Advertising Campaign

Periodic Report for July 27, 2014

Observations

CPA was 17.8% lower than the benchmark at

\$25.47 this week and has seen an increase of \$0.73 (up 3.0%) since last week. Since the last reporting period, the campaign has seen a decrease in CPA of \$1.06 (down 4.0%). Spend is up 15.8% to \$69,047.10. Current conversion rate of 10.3% is 2.8% above campaign benchmark.

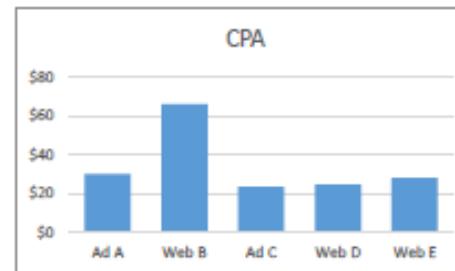
The efficient CPM (\$0.25) resulted in higher volumes of impressions for Website E but has decreased the conversion rate. Website E generated a CPA of \$21.73.

86.2% of visitors left the Acme site after viewing only the home page.

Highlights

- Website D is the best performing site with a CPA of \$21.75 and a conversion rate of 9.40%.
- Website B generated the highest amount of traffic to the website, accounting for 55.6% of the traffic overall.
- Viewers from Website E viewed the most number of pages (2.6 on average).

Overall Campaign Performance To Date				
Total Conversions	CR	CPA	Spend	CTR
51,734	9.47%	\$29.21	\$1,511,170	0.05%



Things that could be improved

- Website B is the worst performing site with a CPA of \$47.63 and a conversion rate of 5.90%.
- Viewers from Ad Partner C viewed the least number of pages (1.5 on average).
- Ad Partner C generated the lowest amount of traffic to the website, accounting for only 1.0% of traffic overall.

Possible Optimisations

Site Level: Across sites, Website B is performing poorly in terms of CPA at \$47.63. Other things being equal, directing that expenditure to Website D would result in a CPA of \$25.21 (down \$1.66) and a conversion rate of 12.19% (up 1.86%).

Tactics within Sites: Other things being equal, reallocating resources from Ad Partner C's Look-a-Like Targeting to their CPA Program would result in a CPA of \$18.56 (down \$4.52) and a conversion rate of 53.09% (up 30.19%).

Creatives within Tactics: Website D's Site Retargeting using Home creatives is not performing well. Moving these resources to their Site Retargeting using Many Ways Icon would result in a CPA of \$13.52 (down \$0.58) and a conversion rate of 32.09% (up 4.25%).

Creative Sizes within Tactics: Ad Partner A's Contextual Filtering using Home at 728x90 is not performing well. Moving these resources to their Contextual Filtering using Home at 300x600 would result in a CPA of \$4.24 (down \$32.96) and a conversion rate of 42.29% (up 40.04%).

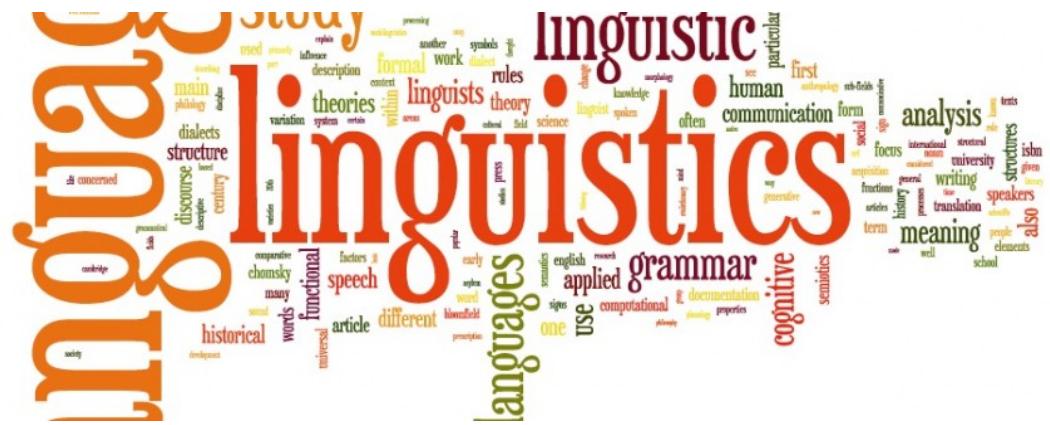
Current Research into the Asset Management Product Space

**Current Research
into the
Asset Management
Product Space**

The Elements of a Science of Information Delivery

1. Linguistics

- The scientific study of language and its structure, including the study of grammar, syntax, and phonetics



Key Contributions

- Philosophy of language: speech act theory, sense and reference ...
- Conversational analysis
- Syntax
- Formal semantics
- Discourse and pragmatics
- Phonology: prosody, intonation, ...
- Stylistics

TRAITS OF **GOOD WRITING**

IDEAS

A thought or suggestion as to a possible course of action

VOICE

What notes the writer's personal style, as all his or her feelings and convictions come out through the words

SENTENCE FLUENCY

Accurate and rapid facility in using a variety of different sentence patterns in a single writing activity

CONVENTIONS

spelling, punctuation, capitalization, grammar, and paragraphing to enhance readability

WORD CHOICE

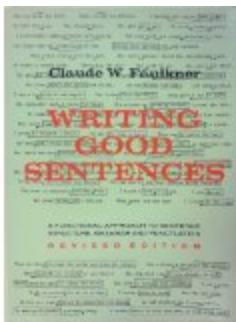
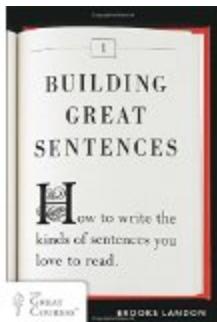
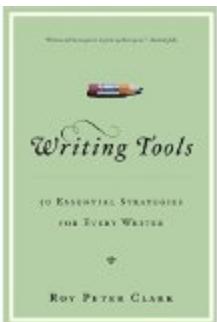
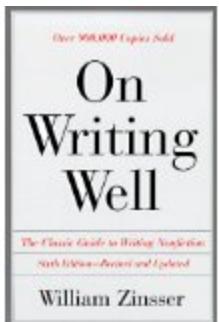
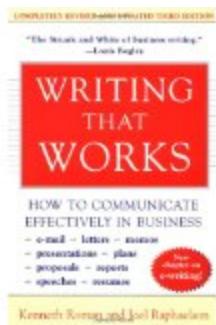
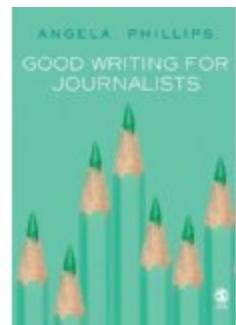
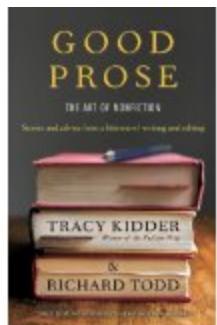
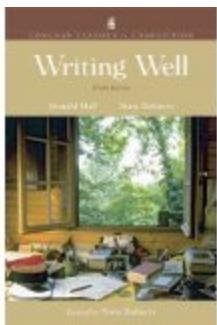
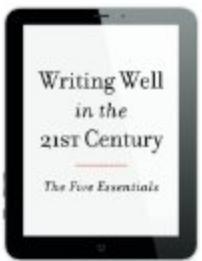
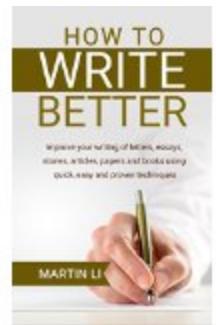
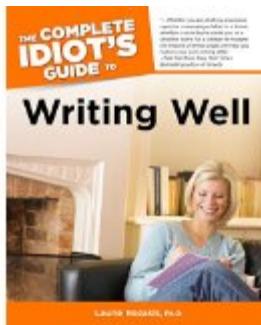
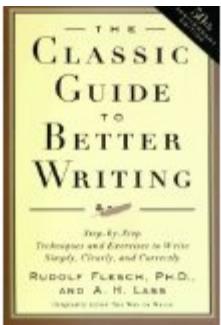
Diction ("a saying, expression, word"), in its original, primary meaning

PRESENTATION

The manner or style in which something is given, offered or displayed

ORGANIZATION

The structure or arrangement of related or connected items



An Old Classic



An Element of Style

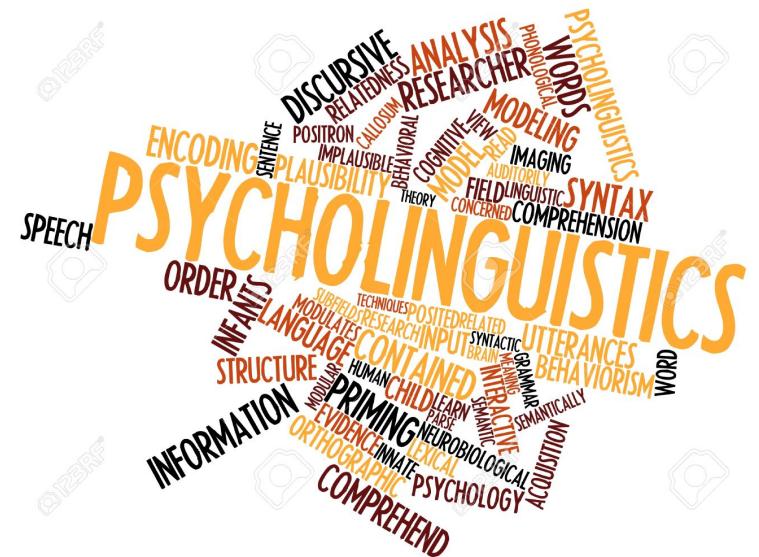
- Vigorous writing is concise. A sentence should contain no unnecessary words, a paragraph no unnecessary sentences, for the same reason that a drawing should have no unnecessary lines and a machine no unnecessary parts. This requires not that the writer make all his sentences short, or that he avoid all detail and treat his subjects only in outline, but that he make every word tell.

An Element of Style

- Vigorous writing is concise.
- A sentence should contain no unnecessary words,
a paragraph no unnecessary sentences,
for the same reason that
a drawing should have no unnecessary lines
and a machine no unnecessary parts.
- This requires
not that the writer make all his sentences short,
or that he avoid all detail and treat his subjects only in outline,
but that he make every word tell.

2. Psycholinguistics

- The study of the relationships between linguistic behaviour and psychological processes



Key Contributions

- An understanding of human processing limitations
 - Memory constraints
 - Management of ambiguity and complexity
- How context impacts interpretation
 - Most sentences are ambiguous in ways we don't even notice
- How the human language production mechanism works
- Readability studies

3. The Psychology of Perception

- Perception is the process of recognizing and interpreting sensory stimuli



Key Contributions

- Understanding how perception works
 - Leads to rules for information presentation
- Understanding the limits of the human perceptual system
- Understanding legibility



4. Graphic Design

- Graphic design is the process of visual communication and problem-solving through the use of type, space, image and colour

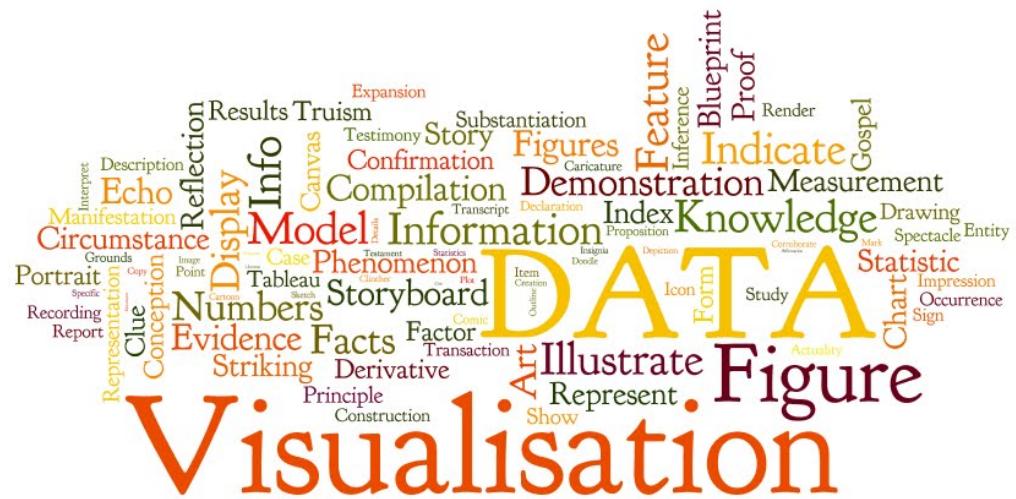


Key Contributions

- The use of visual presentation to enhance legibility and readability
- The role of typography
- The role of branding: identity reinforcement through logos, colours and text
- The effects of physical placement

5. Visualisation

- Data visualization describes any effort to help people understand the significance of data by placing it in a visual context

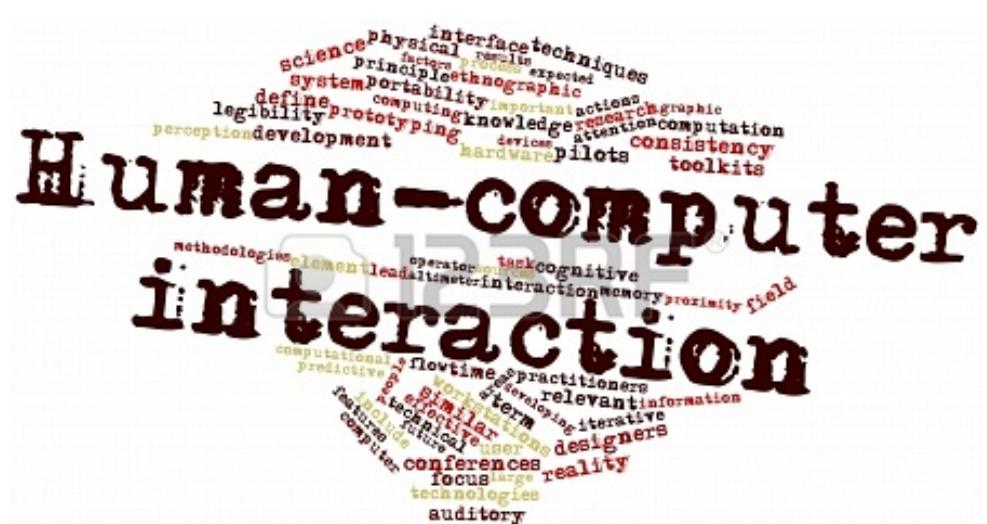


Key Contributions

- Classification of data types
- Guidelines on choice of visualisations

6. Human–Computer Interaction

- Human–computer interaction (HCI) researches the design and use of computer technology, focusing particularly on the interfaces between people (users) and computers



Key Contributions

- User-centred thinking
- A mind-set: seeing a document as an interface to information

Where Can the Study of Information Delivery Take Us?

Starting with:

- Linguistics
- Psycholinguistics
- The Psychology of Perception
- Graphic Design
- Visualisation
- Human-Computer Interaction

... Create this:

<http://neomam.com/interactive/13reasons/>

Readings for Information Delivery 101

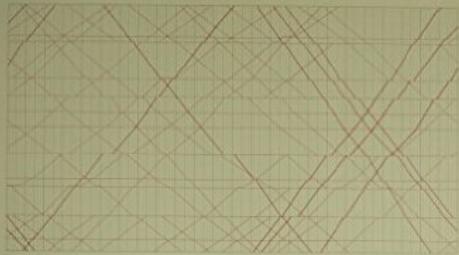
THE WALL STREET JOURNAL.
**GUIDE TO
INFORMATION
GRAPHICS**

**THE DOS & DON'TS
OF PRESENTING
DATA, FACTS,
AND FIGURES**

DONA M. WONG

"INVALUABLE." —HOW DESIGN

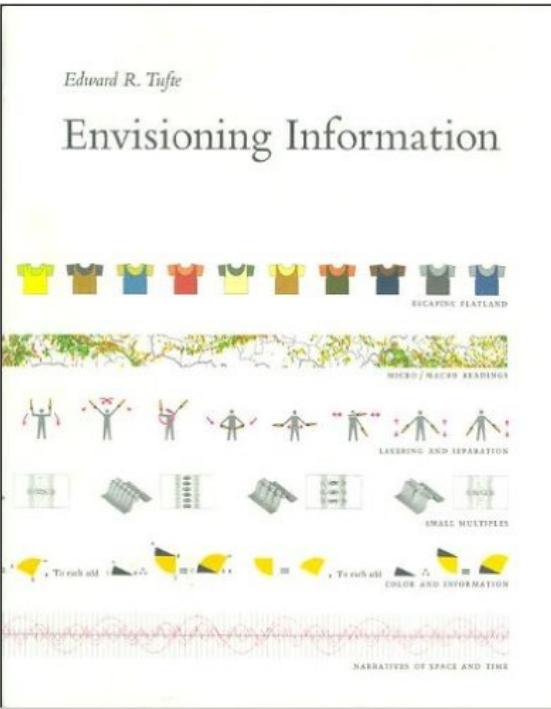




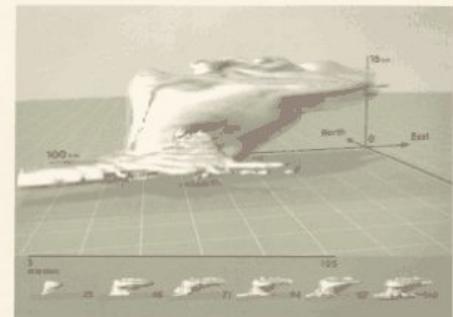
SECOND EDITION

The Visual Display of Quantitative Information

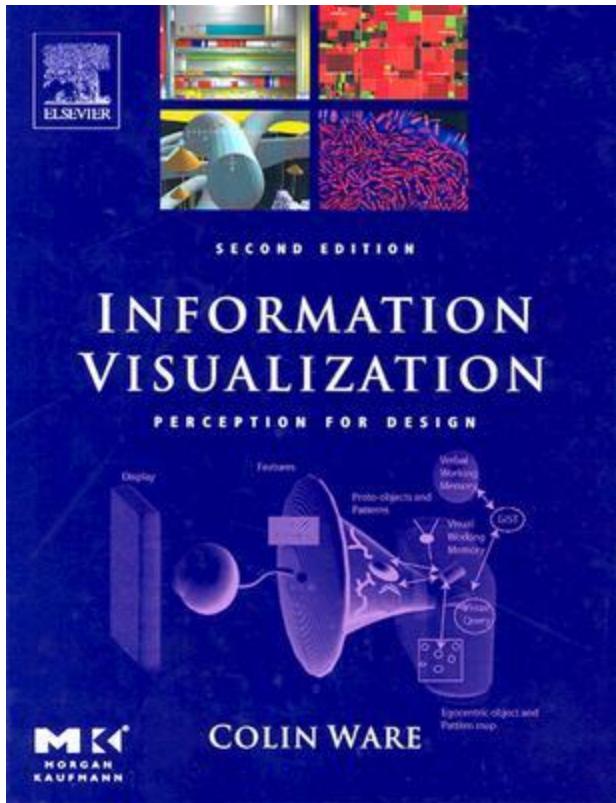
EDWARD R. TUFTE



EDWARD R. TUFTE
VISUAL EXPLANATIONS



IMAGES AND QUANTITIES, EVIDENCE AND NARRATIVE



Robert E. Horn

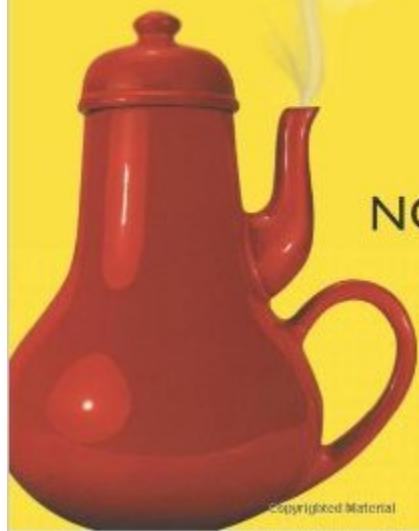


Visual Language

Global Communication
for the 21st Century

Copyrighted Material
REVISED & EXPANDED EDITION

The DESIGN
of EVERYDAY
THINGS



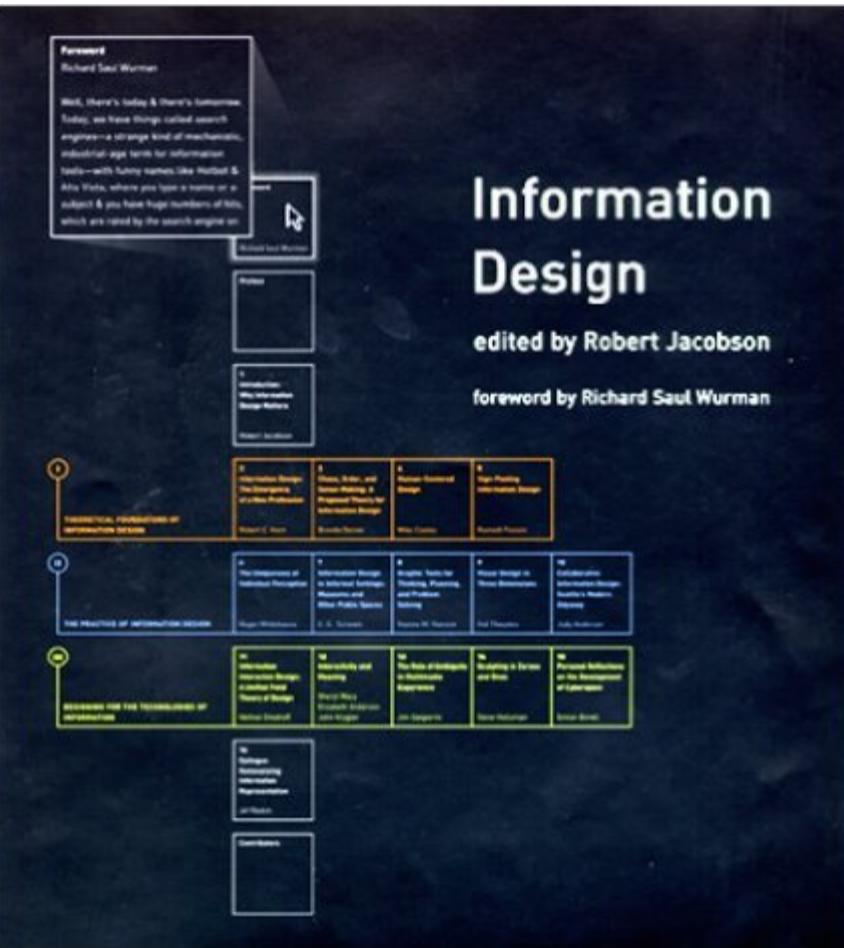
DON
NORMAN

Copyrighted Material

Information Design

edited by Robert Jacobson

foreword by Richard Saul Wurman



Key Messages

- Text is not the only way of delivering information
- Even when text is the best way, it is always presented in a visual context

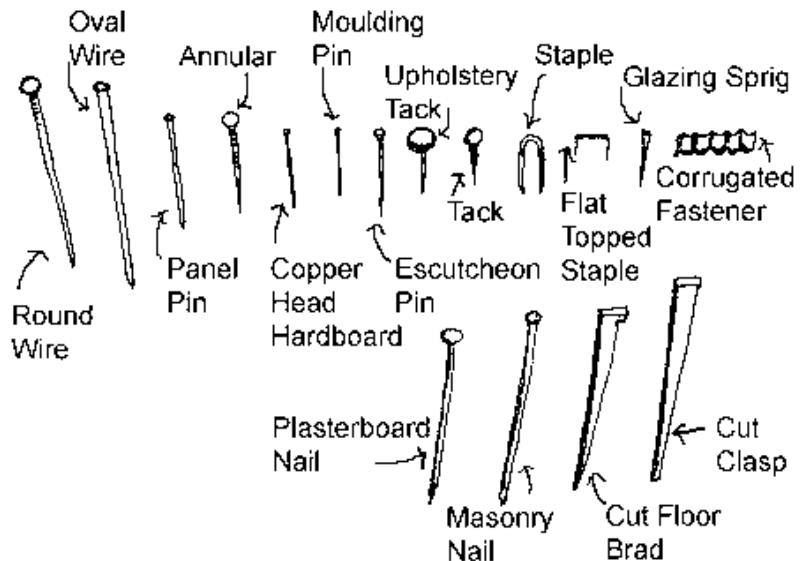
Key Messages

- Know when you shouldn't use a hammer



Key Messages

- Know when you shouldn't use a hammer
- If you do use a hammer, use the right nails for the job



Key Messages

- Know when you shouldn't use a hammer
- If you do use a hammer, use the right nails for the job
- Care about how the end result looks



A collection of various hammers and mallets, including claw hammers, ball-peen hammers, and large wooden mallets, scattered on a light-colored surface. The tools are made of wood, metal, and plastic.

Questions?