

---

# How the Internet will determine the future of publishing!

---

**Giordano Beretta**

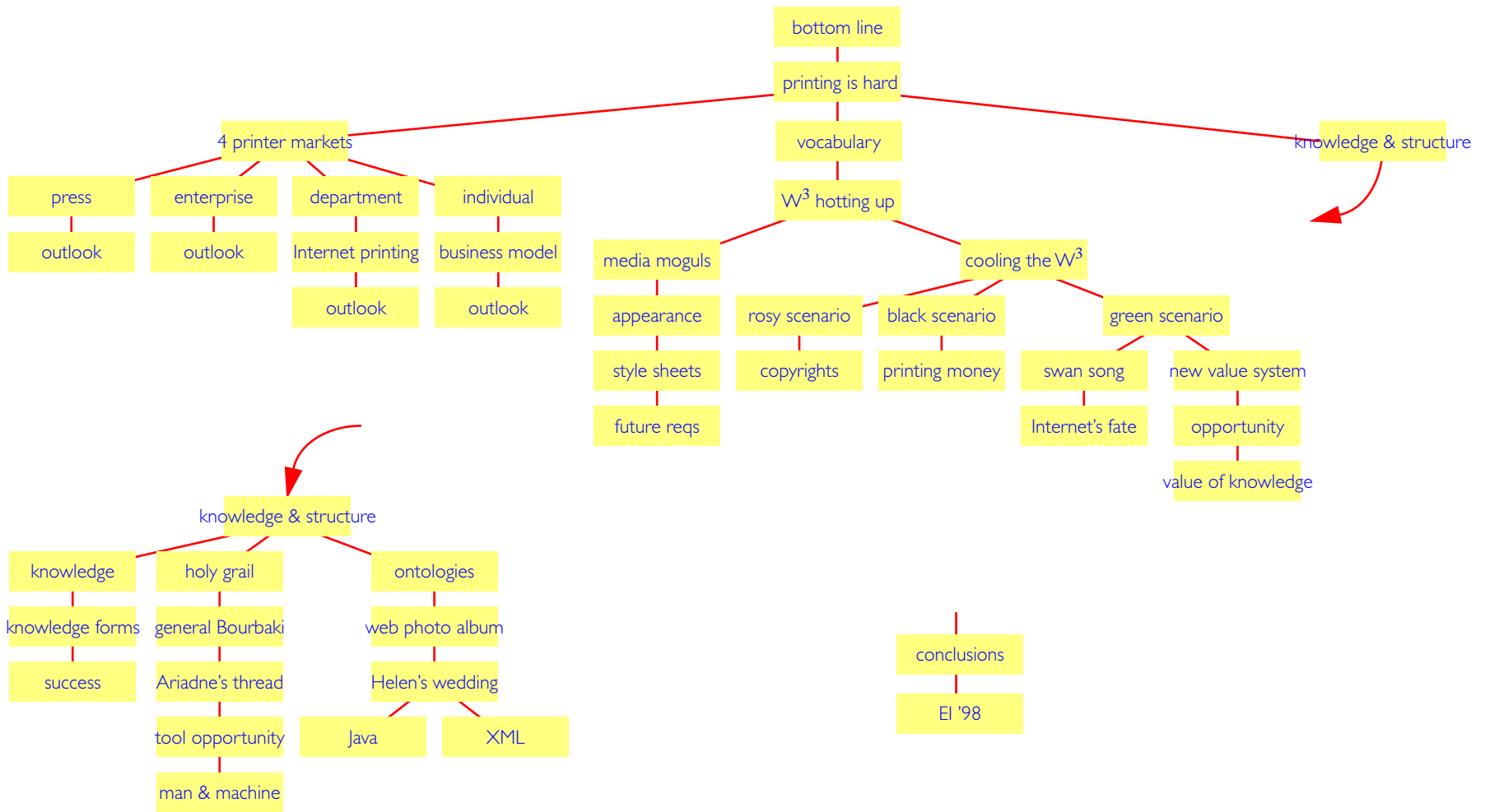
**Hewlett-Packard Laboratories  
Imaging Technology Department**

**1501 Page Mill Road  
Palo Alto, CA 94303**

[http://www.hpl.hp.com/personal/Giordano\\_Beretta/](http://www.hpl.hp.com/personal/Giordano_Beretta/)



# Outline



# The bottom line

3



- World Wide Web is the hot new publication medium
- Paper is the best medium to present written information
- To own the digital printing market you have to be the *best* in printing information off the Web
- Web printing is hard because Web pages are linked but disconnected: poor structure
- William Shakespeare, Romeo and Juliet, act II, scene II:  
*But, soft! what light through yonder window breaks?  
She speaks yet she says nothing: what of that?*

# Why printing on WWW is hard

4



## *Bottom line*

- Web printing is hard because in traditional printing
  - the author decides contents, structure, and appearance
- ... while on the Web
  - the author decides contents and structure
  - the reader decides the appearance
- Web designers optimize material for on-line viewing
  - high-concept design
  - attention grabber
  - multimedia
  - dynamic

# Four printer markets



- **Commercial presses**
  - offset lithography (planographic), gravure (intaglio), letterpress (relief)
  - web, sheet-fed, multicolored, arbitrary finishing, trade printers
  - Heidelberg, Roland, Xeikon, Indigo, ...
- **Enterprise (production) printing systems**
  - Electrophotography, solid-ink jet
  - Scitex, Tektronix, Xerox, ...
- **Department (workgroup) printers**
  - Brother, Canon, Hewlett-Packard, IBM, Lexmark, NEC, Tektronix, Xerox, ...
- **Individual (desktop) printers**
  - Canon, Epson, Hewlett-Packard, Lexmark, NEC, ...

# Press



## *Four printer markets*

- Typical performance: over 10,000 forms per hour
- Requires high capital investment
- Complex make-ready
- Maximum flexibility
- Optimal for long runs



# Outlook

7



## *Four printer markets — Press*

- Direct to press
- Shorter runs become viable
- Personalized printing
- Lowest cost per page



# Enterprise (production) printers



## *Four printer markets*

- Typical performance: over 200 pages per minute
- Medium capital investment
- Run by trained operator
- Application: print on demand



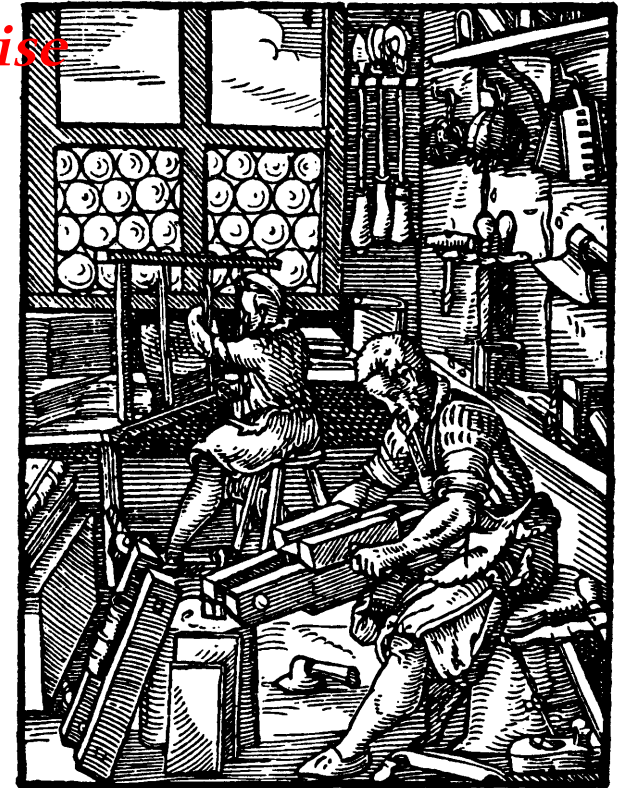
# Outlook



## Four printer markets — Enterprise

- Combination with e-commerce can open new markets
- Copyright problem (branding, DOI<sup>1</sup>)
- Challenges:
  - workflow
  - finishing
  - paper stock

## Der Buchbinder.



Ich bind allerley Bücher ein/  
Geistlich vnd Weltlich/groß vnd klein/  
In Perment oder Bretter nur  
Vnd beschlags mit guter Clausur  
Vnd Spangen/vnd stempff sie zur zier/  
Ich sie auch im anfang planier/  
Etlich vergüld ich auff dem schnitt/  
Da verdien ich viel geldes mit.



1. Digital Object Identifiers: electronic marking, system to track goods in digital commerce

# Department (workgroup) printer

10



## *Four printer markets*

- Typical performance: over 30 pages per minute
- Low capital investment
- Shared in a workgroup
- Internet printing



# Internet printing

11



## *Four printer markets — Department*

- Print to URL
  - challenge: across firewalls
- HTML queue status from URL
  - challenge: printer vendors customize HTML for product differentiation
- Web point and print
  - challenge: hot link in HTML view for automatic downloading of drivers
- For more information see <http://www.pwg.org/>
  - Printer Working Group — “make printers and the applications and operating systems supporting them work together better”

# Outlook

12



## *Four printer markets — Department*

- Remote diagnostics and servicing now common on enterprise printers will be standard on these printers thanks to Internet
  - firewalls protect manufacturers from competitors
- Driver and color management problems
  - Windows NT 5.0 will have better integrated drivers ...
  - ... and color management
- Battlefield for market-share wars

# Individual printers



## *Four printer markets*

- Typical performance: over 2 pages per minute
- Cost is typically expensed
- On network, but for personal use
  - different from personal printer, which is local on a parallel I/O interface port
  - Internet connection can be used for gathering marketing info



# Business model

14



## *Four printer markets — Individual*

- Many inexpensive printers
  - no margin on device, razor-blade model
  - profit on consumables
- Idle most of the time
  - low volume use
  - obsolete before broken
- Motivate users to print a lot
  - bundle project kits with printers
  - make project kits available over the Internet
  - how can consumer be enticed to use them?
- Keep introducing new models with incremental improvements

# Outlook

15



## *Four printer markets — Individual*

- Highest cost per page
  - source for lucrative consumable business
  - exploit vanity factor
- Problems of printing HTML information must be solved
  - cascading style sheets
- End-users will not calibrate their devices for color reproduction
  - color fidelity is an oxymoron in this market segment
  - ... yet Internet shopping requires predictable color reproduction
  - new concept: color integrity

# The hotting-up of the Web

16



*Søren Kierkegaard*

- Birth of the tabloid press in the 1850s
  - Hubert Dreyfus, UC Berkeley
- Leveling of information
  - everybody is interested in everything
  - nothing is too trivial or too important
- Accumulate information—postpone decisions indefinitely
  - nobody takes action, nobody is responsible for truth
  - no risk in action: there is no mastery, just gossip
- Æsthetic sphere of existence
  - inability to distinguish between trivial and important leads to boredom
  - 1996: standardization efforts to control appearance

# The hotting-up of the Web



## Marshall McLuhan: media hot and cool

- Mass media: everybody becomes involved in them at the same time

hot	cool
low participation	high participation
excludes	includes
radio	telephone
movie	early TV
high definition	low definition
phonetic alphabet	ideogrammatic chars
paper (for writing)	stone
lecture	seminar
book	dialogue
glasses	dark glasses

hot	cool
specialist	generalist
detrribalizes	retribalizes
intensely filled-in city	casually structured town
city slicker	rustic
Calvin Coolidge	F.D.R.
bomb scare	war
self help books	aphorisms
high literacy culture	non-literate culture
waltz	twist
hot jazz	cool jazz

# Media conglomerates

18



## *The hotting-up of the Web*

- Maturity of Internet allows big players to enter the market directly
  - no pioneer phase
  - very high entry price
- Hybrid magazine/news-service/television
  - Web sites organized like TV channels
  - Web sites organized like magazines
  - same with TV tie-ins
  - ephemeral (paper as archiving medium)
- High-concept designs
  - alliances with Hollywood, Digital Coast
  - buy contents in Multimedia Gulch & Silicon Alley
- Media: amalgam of forms blurred under epistemological pressures

# The appearance problem

19



## *The hotting-up of the Web*

- Designers work with space
- 15 years ago a similar problem occurred with the advent of WYSIWYG text editors:
  - a document that looks well on the screen does not necessarily print well
  - different fonts are more readable at 72 dpi than at 600 dpi
  - printed documents easy to browse by thumbing...
  - ...on-line document easy to browse with nesting levels
- Documents are frequently repurposed and their style should be easy to modify with global operations
- Solution: *style sheets*
  - Example: Tioga editor in Cedar

# Cascading style sheets



*The hotting-up of the Web*



20

- In-line style commands, no printing support
- Formatting model: box oriented
- Tools currently available mostly only for professionals, like FrameMaker 5.5, Interactor 1.1, DreamWeaver, StyleMaker, Symposia doc+ 3.0, PageSpinner, Sheet Stylist, Cascade, QuickSite 2.0
- Most people's browsers are pre-CSS
- For more information see <http://www.w3.org/Style/css/>

# Future requirements

21



## *The hotting-up of the Web — CSS*

- Support in major authoring tools for Web content creators
- Specific support for various targets
  - fast or slow link view
  - PDA view
  - hard copy
  - read or browse
  - different environment capabilities
  - different audiences
  - ...
- Define concept of a page
  - positioning
  - isolated text (widows and orphans)
  - fonts

# Cooling the Web

22



## *The hotting-up of the Web*

- Kierkegaard: Flatness is overcome by creating opportunities for vertical activities
  - from the æsthetical sphere, people evolve to the ethical sphere
  - people do not just accumulate information but take action and make commitments
- McLuhan: Instead of saving work, labor-saving devices permit everybody to do their own work
- Opportunity: cool the Web
- Three scenarios for the Web's future as a cool medium
  - technologist's view: rosy scenario
  - economist's view: black scenario
  - humanist's view: green scenario

# Rosy scenario (technologist's view)

23



## *Cooling the Web*

- Uniform Intel hardware architecture
  - PC's are a commodity
  - affordable
- Uniform software
  - Windows NT, CE, and 98 platforms
  - middleware / customized applications
- Uniform Internet carrier infrastructure
  - Who will be the next Ma' Bell?
  - Worldcom? British Telecom?
- Systems have become a commodity: they are *substitutable*
  - the accumulation of wealth will come from the naming of things rather than the making of things

# Show stopper: intellectual property

24



## *Cooling the Web — Rosy scenario*

- Concepts behind copyrights
  - Ideas and facts
  - Expression of ideas
  - Fair use
- Copyright holder in a collective work
  - *Tasini vs. New York Times* (based on section 201 (c) of the federal copyright law)
  - author retains the rights to the individual article and may license it to an on-line publisher without permission from or payment to the publisher
  - Steve Lohr's article in *New York Times* of August 14, 1997
  - <http://www.nwu.org/nwu/tvt/tvtrule.htm>
- Collections can be created on the fly by weaving a thread
- Conclusion: What counts is what can be done with an idea



# Inlining, trust

25



## *Cooling the Web — Rosy scenario*

- Inlining: include an image on somebody else's server
  - considered antisocial because contents is used for commercial profit without establishing a formal relation
  - Georgia law (1996) requires explicit permission to include a link to another site on a Web page
- Trust: how can I know you are who you say you are?
  - ID cards help solidifying strong society & good behavior by giving authorities ways for holding people accountable for their actions
  - identifications permit trust & commerce, which benefit all members of society
  - digital certificates:
    - who is the authority?
    - how do you enforce it?
    - only Georgia has a law

# Black scenario (economist's view)

26



## *Cooling the Web*

- 20 years of social polarization
- Career self-reliance, virtual corporations
- R&D evolved into A&D
- The anorexic company: no slack for emergency



# Printing money

27



## *Cooling the Web — Black scenario*

- Baby-boomers must save their own money for retirement
- Due to low interest rates the only long-term savings options are 401(k) plans investing in the stock markets
- Demands for stock is so high, that companies can print stock certificates for M&A
- When baby-boomers cash in to retire, highly leveraged anorexic conglomerates collapse like houses of cards
- Conclusion: The Web allows individuals to make a living as independent professionals — they will have to buy tools

# Green scenario (humanist's view)

28



*Cooling the Web*

- McLuhan's global village
- The physical location is no longer important
  - you no longer have to be in a specific geographic place for your specialty
- In the village participation is high and organization is low
  - this is the formula for stability in organizations
- *Communities of practice, extended knowledge networks (EKN)*
- Next: why a global village, new value systems

# Global village: after a swan song

29



## *Cooling the Web — Green scenario*

- Cultures dawn by bloating up before disintegrating
  - with the increase of communication speed, powers can extend their influence from the *center* to a *margin* that is increasingly farther away...
  - ...until the communications channels can no longer sustain growth
  - the speed of electronic communication is such that our specialist and fragmented civilization of center-margin structure is suddenly reassembling itself into an organic whole when the communications channels get clogged
- Arnold Toynbee: “A disintegrating civilization purchases a reprieve by submitting to forcible political unification in a forcible state”
- We are in a hectic period of M&A and at the same time it becomes ever harder to find executives that can manage these large and complex companies

# Internet's fate

30



## *Green scenario — Swan song*

Are there consequences if Ma' Bell's successor would collapse?

- The global village can route TCP/IP and HTTP like the UUCP community could route e-mail
- Domain name servers are not vital
  - California phone system works although half of the numbers are unlisted
  - epidemic algorithms for replicated database maintenance (Demers et al. 1989)
- Similarly for the other elements of the technical infrastructure

# New value system for the post-modern



## Cooling the Web — Green scenario

- The value system of the *nomad* in the global village is very different from that of the modern *citizen*
- Commercial success: Think locally — act globally
- Lewis Lapham’s series of antonyms:

Citizen	Nomad
authority	power
achievement	celebrity
doubt	certainty
science	magic

Citizen	Nomad
experience	innocence
build	wander
happiness	pleasure
literature	journalism
civilization	barbarism
will	wish

Citizen	Nomad
peace	war
history	legend
argument	violence
art	dream
agriculture	banditry
politics	prophecy

# Opportunity in the post-modern age

32



## *Cooling the Web — Green scenario*

- Søren Kierkegaard: provide tools for people to evolve from the æsthetical to the ethical sphere
- Empower individuals to emancipate from levelled world to strong identities
- Help people to emancipate from accumulating information to taking action and making commitments
- Provide tools to distill information into knowledge
- Knowledge: justified true belief, mastery



HP Laboratories

# Value of knowledge



## *Cooling the Web — Green scenario*

- There are no enforceable copyrights in the global village
  - Tasini vs. New York Times: publish early and publish often
  - <http://www.nwu.org/nwu/tvt/tvtrule.htm>
- What counts is what can be done with an idea



# Knowledge and structure

34



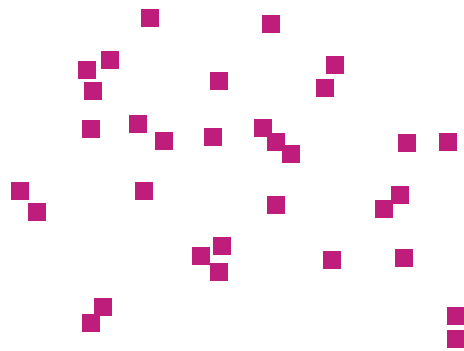
## *Definitions*

- Doug Engelbart: categorization of evolving knowledge
- Recorded dialog
  - databases, diaries, notes, address books, captions
- External intelligence
  - ontologies, guide books
- Knowledge products
  - Web photo albums, scrap books

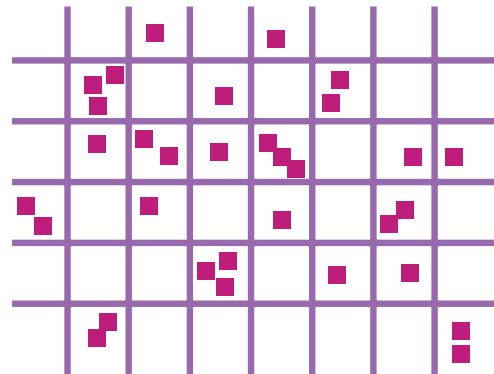
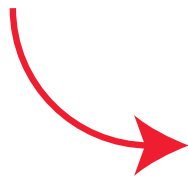
# Categorization of evolving knowledge



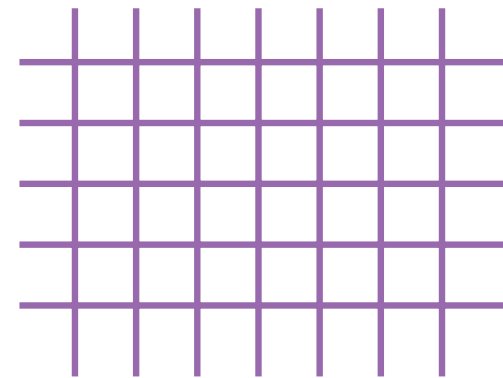
## Knowledge and structure — Definitions



information  
recorded dialog



knowledge



structure  
external intelligence



# Definition for knowledge

36



## *Knowledge and structure — Definitions*

Simple definition:

Entity  $A$  knows  $p$  if and only if

- 1  $A$  believes that  $p$
- 2  $A$  is justified in believing that  $p$
- 3  $p$  is true

- Informally: justified true belief
  - belief is based on experience, skills, and mastery
  - knowledge in its strict form of propositional knowledge—how to do things—rather than an accumulation of facts
  - skills and mastery ensure that the belief is acquired by experience, ruling out that a belief meeting these conditions is possessed by sheer accident

# Tacit knowledge

37



## *Knowledge and structure — Definitions*

Ikujiro Nonaka: Knowledge creating company

Tacit knowledge...

- ...consists of mental models, beliefs, and perspectives so ingrained that we take them for granted, and therefore cannot easily articulate them
- ...is deeply rooted in action and in an individual's commitment to a specific context

# Explicit knowledge

38



## *Knowledge and structure — Definitions — Nonaka*

- Explicit knowledge is formal and systematic; it can easily be communicated and shared in product specifications or a scientific formula or a computer program

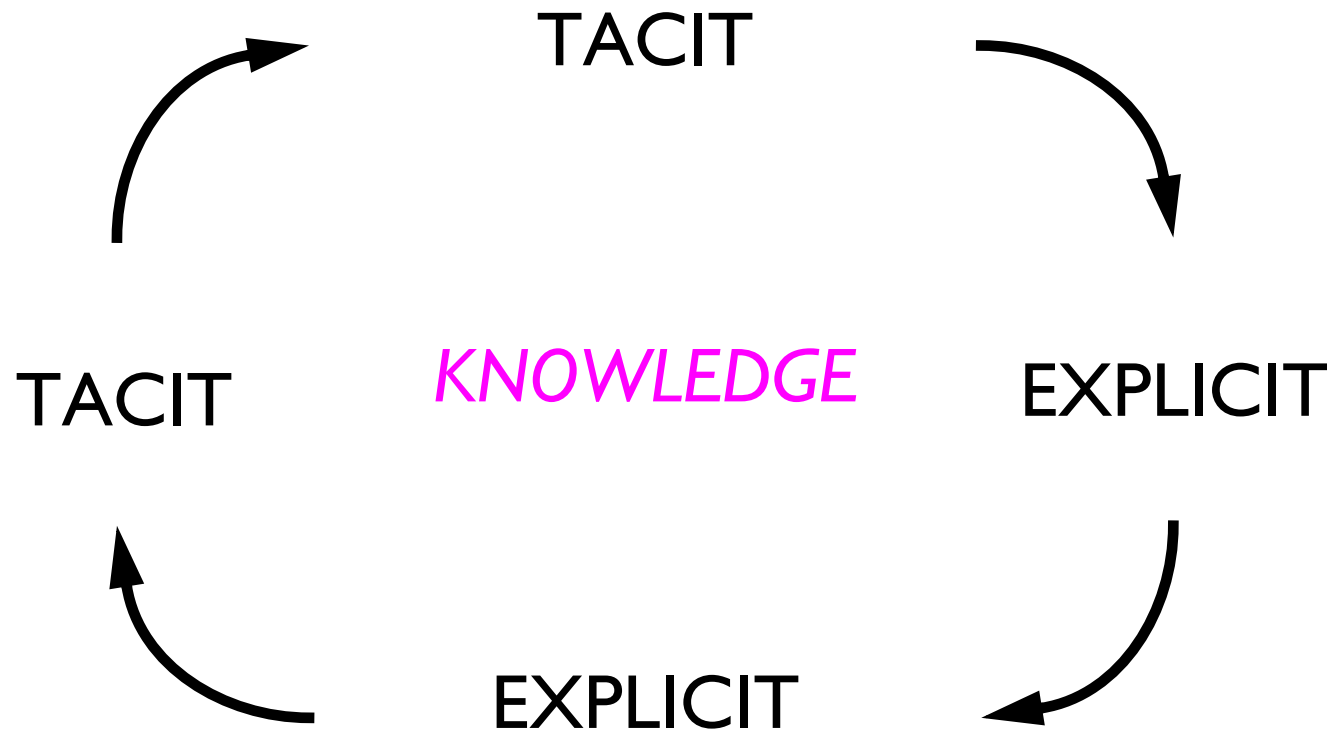


# Knowledge creating organization

39



*Knowledge and structure — Definitions — Nonaka*

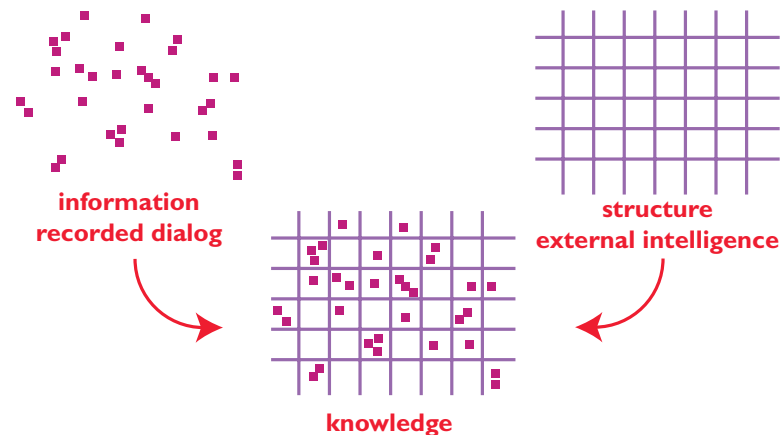


# Structure



## *Knowledge and structure*

- Quality of knowledge can be measured by how effectively it is communicated
- Effective communication requires clear organization
- Clear organization is achieved by introducing good structures



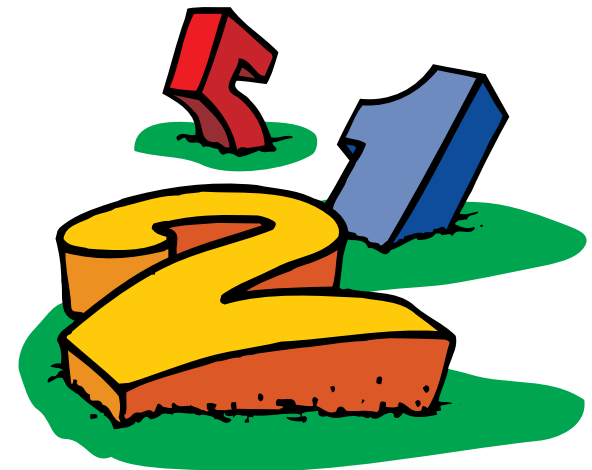
# General Nicolas Bourbaki

41



## *Knowledge and structure — Structure*

- Graphs are too weak: only one relation
- New interpretation of mathematics after 1935
- Relational construct: a set with relations
- System of axioms represents properties of constructs
- Mathematical creativity: find new constructs by defining maps that preserve the relations
- Two-step approach
  - find a good system of axioms
  - find a good isomorphic construct



# Ariadne's thread

42



*Knowledge and structure — Structure*

- For Leibniz, the author finds a linear order in a set of information and creates knowledge in the form of a linear thread

*The true method must provide us with a filum Ariadnes, that is to say a kind of sensitive and coarse means that guides the mind, in the same way as lines drawn in geometry and the type of operations that are prescribed to apprentices in Arithmetic. Without that our mind would not know how to go along a long path without straying.*

Gottfried Wilhelm Leibniz

# Software tool opportunity

43



## *Knowledge and structure — Structure*

- Mathematically: lattices
- Algorithms that manage structures and can traverse them
- Hard design problem: define the best partition between human and tool systems
- Human system is good at
  - categorization (find correct class)
  - associations (story telling, Ariadne's thread)
- Tool system
  - enabling technology: data structures, ADT, OOPS
  - editors for the systematic definition of structures (ontology editors)

# Humans and machines



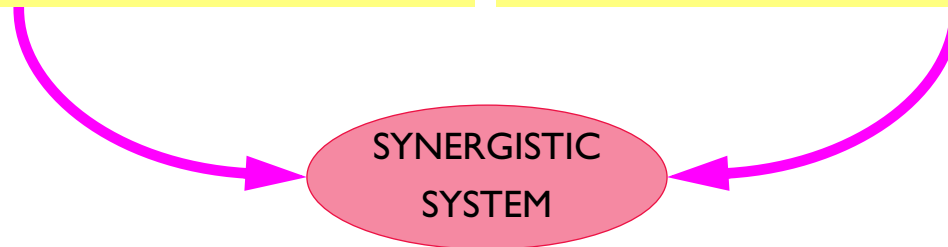
## *Knowledge and structure — Structure*

### *Human system*

categorization	paradigms
organization	customs
procedures	attitudes
methods	emotions
associations	moods
narrations	language
culture	learning
values	knowledge

### *Tool system*

ontology editors	authentication
Web browsers	watermarking
on-line databases	fingerprinting
Java beans	mobile systems
Internet	networks
global file systems	servers
data structures	brokers
nanocommerce	agents



**facilitates the creation of knowledge**

# Ontology



## *Knowledge and structure*

45

- Tool for systematically specifying a structure
- An ontology is a specification of a conceptualization
  - <http://www-ksl.stanford.edu/kst/what-is-an-ontology.html>
- Ontologies provide a vocabulary for representing and communicating knowledge about some topic and a set of relationships that hold among the terms in that vocabulary
- Stanford University KSL Ontolingua Server provides a distributed collaborative environment to browse, create, edit, modify, and use ontologies
  - <http://www-ksl.stanford.edu/>

# Using the ontology



## Knowledge and structure

- Example of an ontology:
  - <http://www-ksl.stanford.edu/>
  - go to Ontolingua Server
  - load the ontology Wedding-Pics
  - created by John Tillinghast
- Once the ontology has been specified, the ontology editor with its inference engine is no longer necessary
- A specific very simple application can be written that allows to author and read a set of Ariadne's threads based on a recorded dialog



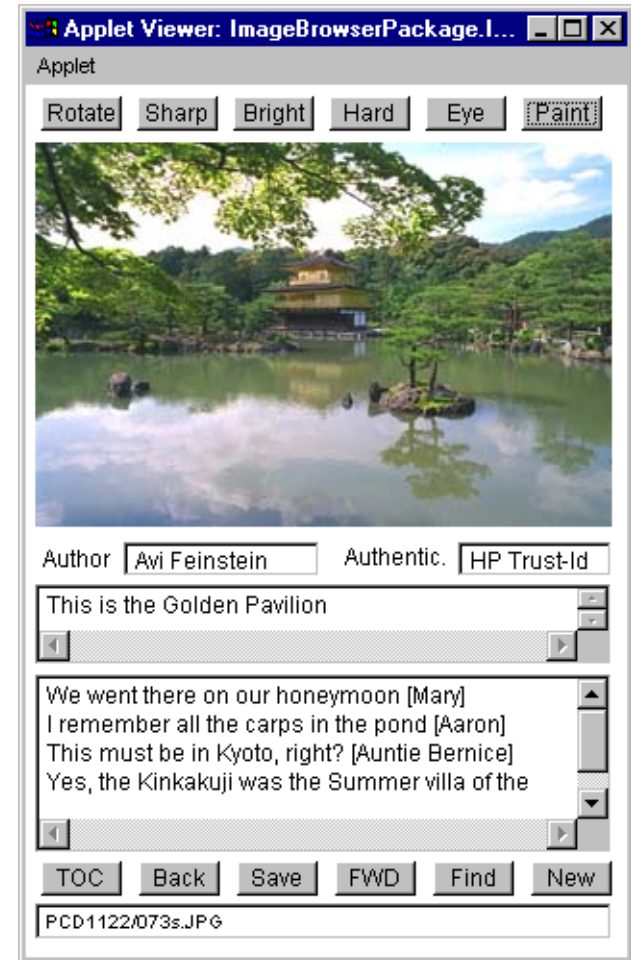
# Java applets



## Knowledge and structure — Using the ontology

47

- Ontolingua frames can be translated into Java classes
- Java classes can be encapsulated into Beans
- A visual editor can be used to rapidly build a browser as an applet that represents some specific knowledge



# Extensible Markup Language (XML)

48



## *Knowledge and structure — Using the ontology*

- A data format for structured document interchange on the Web
  - a metalanguage to let authors design their own markup language
  - an abbreviated version of SGML tailored to Web applications
- More powerful than HTML
  - extensibility: can define new tags and attribute names
  - structure: document structure can be nested (tree instead of list)
  - appearance: Document Style Semantics & Specification Lang. (DSSSL) instead of CSS
  - validation: a grammar can be supplied for structural validation (optional)
- Applications of XML
  - information discovery can be tailored to individual users
  - different views of the same data can be presented to different users
  - processing load can be transferred from server to client
  - mediation between heterogeneous databases



## *Knowledge and structure — Using the ontology*

- A community of practice (see global village, 28) develops & agrees on a Document Type Definition (DTD)
- Information components are stored in a repository
- Server creates documents (with DTD) on the fly, client takes into account user preferences and viewing options
- Embedded Java applet makes document active
- Microsoft's XML parser for Internet Explorer 4.0
  - <http://www.microsoft.com/standards/xml/xmlparse.htm>
  - other structured document specialists: Adobe, ArborText, Chrystal, DataWare, Documentum, Folio, Inforium, Inso, OpenText, PIT, Texcel, Vignette, ...

# Conclusions



## *The Web as a cool or a hot medium*

- People will print because paper is a convenient medium

Segment	Medium	Key market	Enablers
press	hot	low per page cost	customizing, flexibility, bindery
enterprise	hot	print-on-demand	cascading style sheets, e-commerce
department	cool	community of practice	ODB + XML + Java structure/knowledge networks
individual	hot	consumer; convenience vanity factor	fancy consumables rapid product cycles

- Traditional contents has no value in the global village
  - in the e-media the container is gone
  - copyrights protect an expression, not the idea
  - collections can be created on the fly from components by weaving a thread

# Conclusions



## *Hot & cool markets*

Hot	Cool
consumer	professional
followers	early adopters, visionaries
meme complex	community of practice extended knowledge network
brand management	integrity
entertainment value	added value
charge what the market will bear	try to break even
high-concept	tool's power
profits on quantity	profits on quality

# Conclusions



*Paper is a convenient knowledge vector*

- Knowledge tools will produce tailored documents on the fly
- Externalizing knowledge is very hard
- E-commerce will allow people to earn a living by making their knowledge explicit with a new generation of tools
- Of the large companies only Microsoft has a business model fit for the global village
  - produce middleware for a large army of independent applications customizers
  - others: solution oriented, cannot enter in many vertical markets because of inertia

# Conclusions



## *Historical perspective*

- Evolution of e-publishing

Decade	Buzz word	Example	Key technology
70s	WYSIWYG	Bravo, Word	GUI
80s	object based, OLE	Tioga, FrameMaker	API, foundation classes
90s	Web, dynamic hypermedia	ODB + XML + Java	knowledge networks

- Max Frisch: “Technology is the knack of so arranging the world that we do not have to experience it”