



Om märkspråk, kunskap, och ordning

Bidrag till kursen "Den sociala ordningen och
kunskapens ordning"

Hasse Haitto
HasseH@enigma.com

enigma

UNLOCKING THE MOTION OF INDUSTRY

Vad är märkord?

enigma

- **Märkord** är markeringar som omsluter sitt innehåll (engelska: *tags*) "taggar"
- Typiskt utseende på uppmärkt text:
`<titel>Rubriktext</titel>`
- Märkorden skapar en ordning, en struktur utöver innehållet
- En fördel om uppmärkningen är oberoende av program, dvs läsbar av människor (och program)

Typografi och märkord

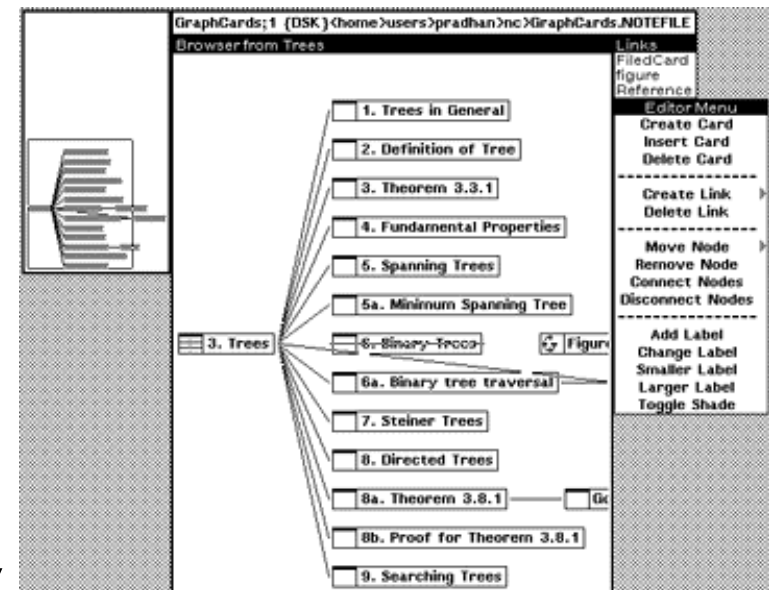
enigma

- Typografiska konventioner som rubriker, fotnoter, anföringstecken osv är en begreppsapparat som vuxit fram under århundraden
- *Innebörden* av märkorden skiljer beroende på tillämpning och epok: ursprungligen ett sätt att koda formatering ("växla till fet stil här")
- Idag används märkord snarare för att skapa en *logisk ordning* över stoff (formatering kan ju skapas genom att tolka märkorden)
- Utveckling motsvarande typografins verktyg och arbetssätt återstår för att fånga upp kunskap (jfr gIBIS)

Hypertext och hypermedia

enigma

- Def: Hypertext är (maskinstödda alt. elektroniska) referenser eller länkar mellan information
- Några hållpunkter i hypertextens historia:
 - 1933-45 Vannevar Bush (memex)
 - 60-talet bl.a. Ted Nelson och Douglas Engelbart NLS/Augment), Brown University/FRESS
 - 70-talet: ZOG/KMS
 - 80-talet: NoteCards, Intermedia, HyperCard
 - 90-talet: HTTP + HTML + NCSA Mosaic = WWW



Några konstateranden

enigma

- Vi har inte (än) uppnått allmän funktionalitet som fanns redan t.ex. i NLS/Augment för tre decennier sedan, t.ex.
 - Adressering/indexering
 - Separation av länkar från adressering
 - ePost integrerad med övrig funktionalitet
 - (Hierarkisk) filtrering av innehåll
 - Videokonferens
 - ...

Douglas Engelbart

enigma

- Läs mer på WWW t.ex.
 - <http://www.bootstrap.org/>
 - <http://sloan.stanford.edu/MouseSite/1968Demo.html>



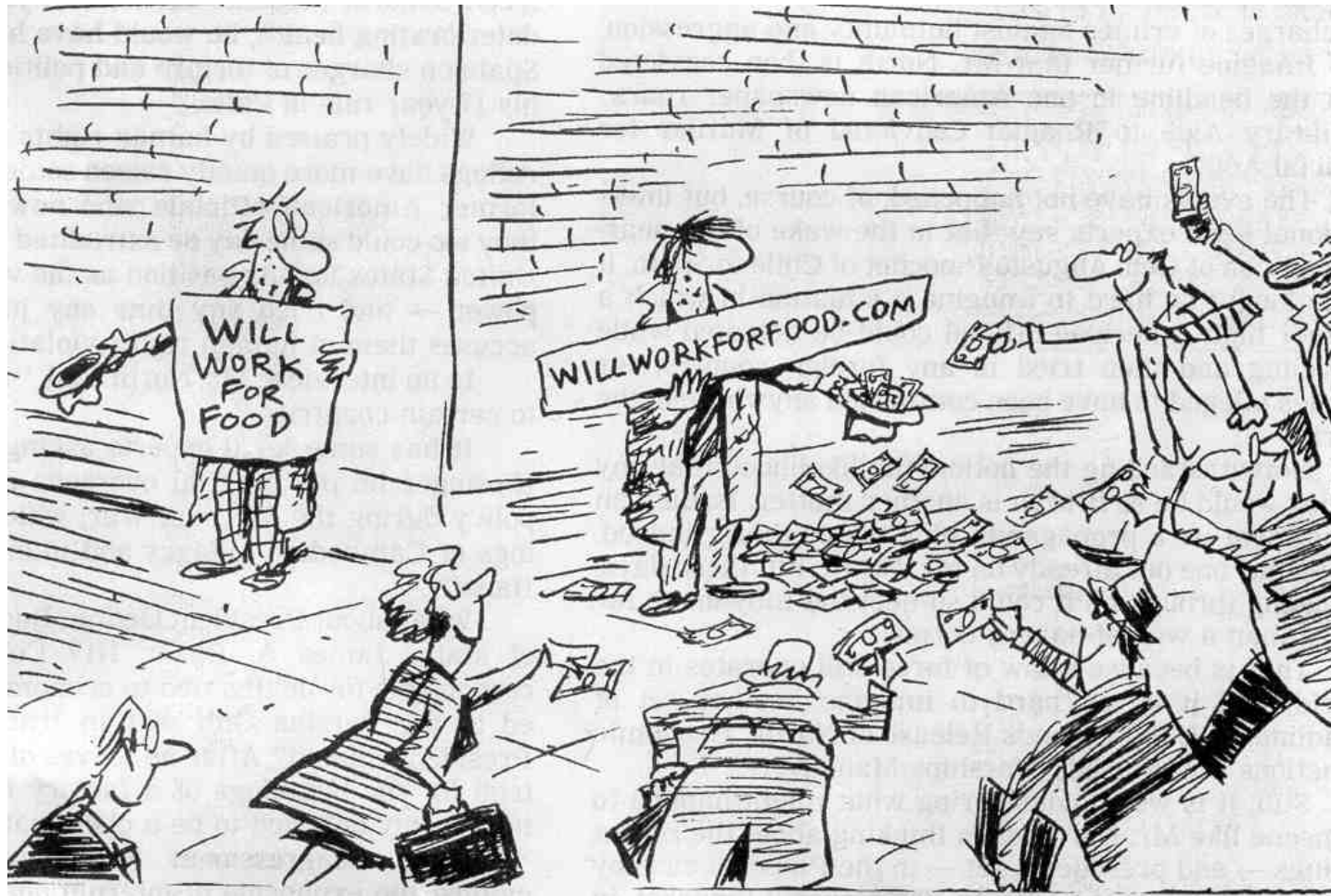
Betydelsen av standarder

enigma

- **Manu'tius, Aldus**, eg. *Aldo Manuzio*, f. ca 1450, d. 1515, italiensk humanist, verksam som boktryckare i Venedig från 1490-talet. M. satsade hängivet på att ge ut den klassiska grekiska och latinska litteraturen, och många texter trycktes först av honom. [...]. **Som redaktörer anlätades framstående lärda, t.ex. Erasmus**. De behändiga men inte särskilt billiga böckerna trycktes i en ny stil, **kursiv**, som baserats på modern humanisthandskrift. (Källa: *Per S. Ridderstad*, NE)
- Standarder väsentliga för såväl format som innehåll!
- Kodning av kunskap—en garant för vederhäftighet och kvalitetssäkring av innehållet

Drivkrafter och resurser för kunskapsarbete?

enigma



Dana Summers
The Orlando Sentinel
Tribune Media Services

Industrins behov...

enigma

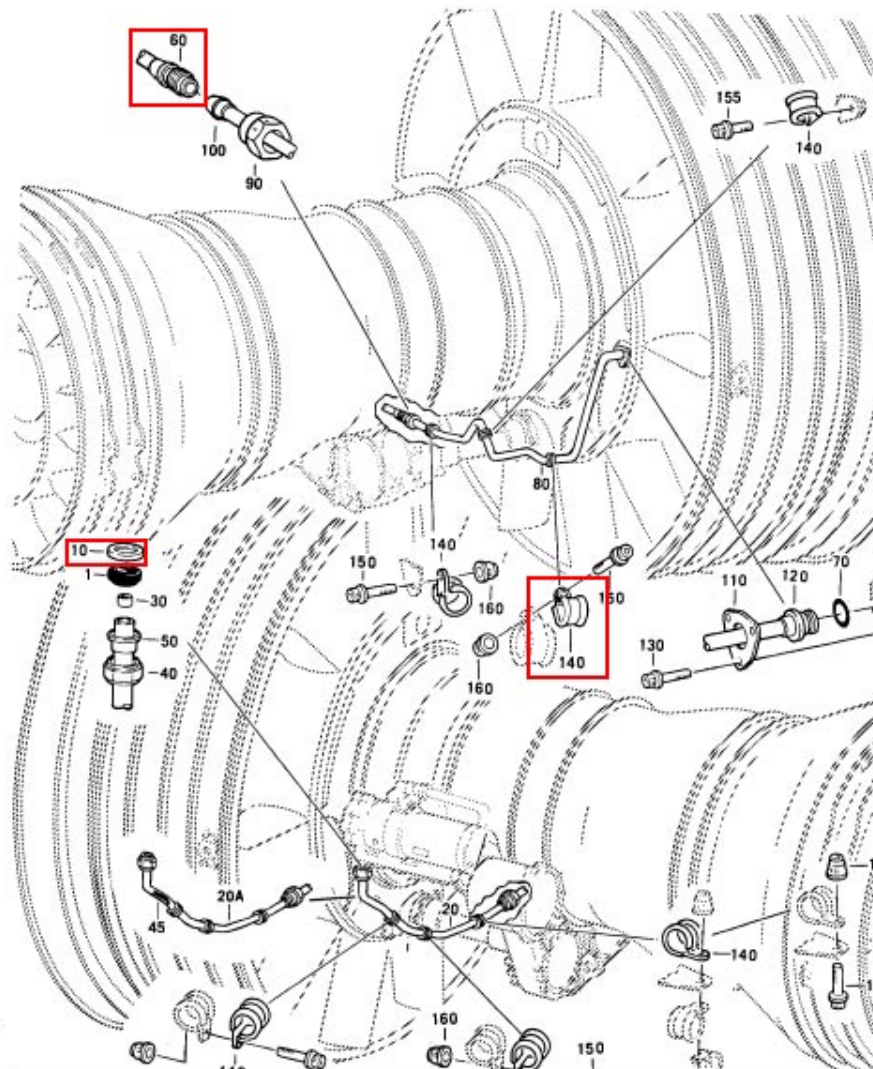


FIG-ITEM	PART NUMBER	NOMENCLATURE
	1234567	
1	389114	PACKING-PREFORM
10	354118	RETAINER-PKG
20	51T374	TUBE-ASSY OF, OIL, COOLER INLET (LS05)
	OR	AS3214-02 WIRE
20A	53T108-01	TUBE-ASSY OF, OIL, COOLER INLET (LS05) pre * SB PW4ENG 79-17
30	1A8508	.LINER-REINFORCING, TUBE
40	225875	.NUT-TUBE CPLG
45	52T967	.STOP-NUT pre * PW4ENG 79-17
50	546690	.FERRULE-TUBE FITTING
60	ST1400-12	.COUPLING-TUBE
70	AS3209-116	PACKING-PREFORM
140	ST1435-12	CLAMP-LOOP, CUSHION
150	MS9556-08	BOLT-MACH,DH .190-32X.438

...och lösningar

enigma

- Komplexa problem att lösa: T.ex. Ett flygplan har tiotusentals delar och kilometervis av ledningar (och det finns många leverantörer av reservdelar)
- Uptime/downtime/JIT
- eCommerce, B2B, vortals/Exchanges/auctions ("letsbuyit")
- Suggestive selling, transactive content
- Service-bulletiner, COCs, Green Pages
- Anpassning av innehåll för "operatörer"

Kodning med hjälp av SGML

enigma

- Standard Generalized Markup Language (1986), en ISO-standard sprungen ur GenCode och IBMs GML
- Märkorden är valfria
- SGML-standarden beskriver bl.a.
 - en syntax (grammatik)
 - en kärna av funktionalitet
 - tillval till standarden
- Tyvärr för komplex för de flesta...

Grundläggande SGML

enigma

- **Märkord** omsluter element och skapar en *hierarki*
- **Element** identifierar koncept
- **Attribut** är information knuten till element:

```
<titel id="rub1">Rubriktext</titel>
```

- Alla dokument hör till någon sorts överordnad *klass*
- Klassen kännetecknas av gemensamma element och attribut

Efterföljare till SGML

enigma

- "HTML isn't a little brother of the Standard Generalized Markup Language (SGML), it's more of a Forrest Gump-like cousin"

— Sean McGrath, *PC Magazine*, juni 1996

- Extensible Markup Language (XML), 1998
 - Framtagen av SGML-specialister
 - En rejäl förenkling av SGML med små tillägg
 - En specifikation från W3C, ej ISO, dvs helt andra premisser för dess fortsatta utveckling
 - Ett stort antal andra specifikationer som komplement (se www.w3.org)

Kodning av kunskap: Länkar

enigma

- Vannevar Bush: "As We May Think" 1945, Atlantic Monthly: Associativt tänkande
- Hyperlänkar och märkspråk:
 - SGML:
 - ID/IDREF(S): inom samma dokument
 - TEI extended pointers: en konvention för att uttrycka länkar mellan olika dokument inkl. stöd av struktur
 - HyTime—extremt välutvecklat stöd för att länka till/från i princip vad som helst, när som helst (standarden själv är en tillämpning av SGML)
 - XML:
 - Xlink och XPointer



Kodning av kunskap: Ordförteckningar m.m.

enigma

- Vanligen representerade i en alfabetisk, hierarkisk ordning, med synonymer och referenser ("se även...") för att skapa flera åtkomstvägar till samma eller *besläktad* information
- En sorts karta över kunskap, uttryckt i form av *koncept* eller *ämnen*, med hänvisningar till deras *förekomster* och *associerad* information
- "Handgjorda" register oftast mest värdefulla (jfr Vannevar Bushs Path Blazer, en sorts kunskapare)
- Behovet att kunna kombinera olika register en av de bakomliggande drivkrafterna för *Topic Maps*

Kodning av kunskap: Topic Maps

enigma

■ Topic Maps

- En ISO-standard 1999 (uttryckt m.h.a. SGML)
- Finns nu även i en XML-version

■ Kodifierar begrepp som

- Ämne (topic) inkl. namn, sorteringsbegrepp, och typ
- Förekomst (roll och slag)
- Besläktade koncept
- Modellen omfattar även omfång, publika beteckningar och metadata

Representation av Topic Maps

enigma

- Representation av ämnen och förekomster är skilda (jfr HyTime)
- Besläktade koncept (topic associations) är *länkar* mellan koncept och är *oberoende* av de informationsresurser som ev. är förekomster av dessa ämnen—dvs topic maps kodifierar kunskap och har ett egenvärde utöver materialet de används för
- Topic Maps kan därför påläggas olika material
- Separation av adresseringen (kontra länkning) gör det möjligt att applicera Topic Maps på material t.ex. på CD eller WWW

"Life is full of connections; knowledge and creativity thrive on them, encyclopaedias, dictionaries and other reference works need to be able to exploit them—and topic maps are the tool that enable them to do so, efficiently and reliably."

—Steve Pepper

- Se även www.ontopia.net och sök på WWW...