## Cost-Benefit Analysis for the Design of Personal Knowledge Management Systems



16.06.2008 / ICEIS 2008 @ Barcelona, Spain.

<Ⅲ

#### Disclaimer



#### This presentation is not intended to replace the paper. Get these slides from <a href="http://pubs.xam.de">http://pubs.xam.de</a>



During the last decade, Knowledge Management is mached a consolidated status as a manageploiting personal information from different everyday applications through common semantics-based metadata. Later on, such personally created, collected,





- Introduction to Personal Knowledge Management (PKM)
- Comparision to Organisatinal KM (OKM)
- Processes in PKM
- Challenges for Re-Using Personal Knowledge
- Quantification of Costs and Benefit
- Conclusions



The most important contribution of management in the 20th century was to increase manual worker productivity fifty-fold. The most important contribution of management in the 21st century will be to increase knowledge worker productivity – hopefully by the same percentage. [...] The methods, however, are totally different from those that increased the productivity of manual workers.

Peter F. Drucker, 1958

#### Today: Communication of Men and Machines





#### Today: Communication of Men and Machines





#### **Organisational Perspective on KM**



#### Personal Perspective on KM





#### Definition of PKM



European Guide to good Practice in Knowledge Management, EUROPEAN COMMITTEE FOR STANDARDIZATION. **CWA 14924** (CEN Workshop Agreement), **ICS** 03.100.99. 2004.

- Personal KM: A set of concepts, disciplines and tools for organizing often previously unstructured knowledge, to help individuals take responsibility for what they know and who they know.
- Organizational KM: Unlike personal KM, which centres on the individual, organizational KM depends upon an enterprise-wide strategic decision to actively manage knowledge through a range of processes, tools and people.

### SECI model (Nonaka and Takeuchi 95)

- Socialisation
   talk, chat
- Externalisation
  - create, write
- Combination
  - collect, copy & paste
- Internalisation
  - search, read



#### **SECI** Today





#### Re-Use of Knowledge Increases Productivity

Today in most organizations, reuse is addressed only at the institutional level, if at all. But it stands to reason that the most effective knowledge workers reuse their own knowledge all the time. If I'm productive as an author, as I write this book I should be able to easily find and reuse previous passages I've written on the subject, perhaps with slightly different language to avoid plagiarizing myself and my publishers (I won't comment on how well I actually do this). A productive lawyer would index and rapidly find all the opinions and briefs he's ever written, and routinely reuse them for new clients. But while we know this is true, organizations have yet to help knowledge workers execute this sort of reuse. If they were smart, they'd make it easier-and would also provide taxonomies, training, role models, and encouragement.

#### Initia King for a Living First Set Better ferla mance and Resolts from Hanwiedge Workers Thomas H. Davenport

ARVARD ENGINEES SENSOR FREE

#### Re-Use Your Own Knowledge

 $\rightarrow$  Note-taking is communication with yourself





Cost of Externalisation  $(C_E)$ 

Cost of Retrieval (C<sub>R</sub>)

Benefit of use (B)

- Total costs  $C = C_{E+} C_{R}$
- Benefit B?
- $\rightarrow$  Personal KM is always gambling:
  - $\rightarrow$  Will I ever need this knowledge again? In what context?
  - $\rightarrow$  Is it cheaper to re-create the knowledge?  $C_C < C_{E_+}C_R$
  - → What value will it have? How much effort is it worth to structure and formalize?  $B > C_E + C_E$
  - $\rightarrow$  Should I try to search my PKM system now?

#### Knowledge Processes



Technology is common in the domain of knowledge distribution, but it rarely enhances the process of knowledge use. Distribution delivers knowledge to the potential user's desktop but cannot dictate what he or she does with it thereafter. It would be interesting to envision technologies that help to manage personal knowledge as it applies to decisions and actions, but beyond the very rudimentary "personal information managers" that allow searching of unrelated bits of information, little progress has been made toward "personal knowledge managers." On the



Working Knowledge: How Organizations Manage What They Know Thomas H. Davenport und Laurence Prusak, 1998, Mcgraw-Hill Professional, S. 142





# Challenges for Re-Using Personal Knowledge



#### Challenge: Ease of Use ( $\rightarrow$ costs!)

#### Ontology Editor (Protégé)

#### Sticky Note (PostIt<sup>™</sup>)

ject Edit Window OWL Help				
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	🕅 A R 🎦 🖻 🎘	III III		
OWLClasses P Properties	Forms Individuals	Metadata		
	C iso:Address (type-owt)C	lass)		
server merarchy V C A LU,	C iso:Address (type=owh	Class)		
C)iso Address	Contraction of the second			CX
CisoAngle	Language and		53 7205	
C iso:AppSchInfo	Name Labels		Annotations	V C, + -
© iso:AscTypeCd	Iso'Address		Property	Value
🔆 iso:AttributeTyp	And the second	DI	so:stereotype DataTy	pe
© iso:Binary	Documentation			
C ISO:BrowGraph	location of the responsible i	individual or		
Ciso:CharSatCd	organization			
Cien Cita				
CisoCitation	Accessed in the second		ALCING M	At much Thinks I All
iso:ClasscationCd	Asserted Infetred		ALCIASS	At own: Thing All
© iso:Consts	Asserted Conditions	C C + -	× P Properti	esaVC+-
ତ iso:Continfo		NECESSARY & SUFFICIE	NT	Name
©iso:Contact		NECESSA	RY DisprostCo	nde <sup>O</sup>
Siso:ContentTypCd	(a) owl Thing		D iso:admin	Area
iso:DQElement	Iso:adminArea ≥ 0		D iso:eMailA	dd <sup>O</sup> bb
Ciso:Duscope	lso:adminArea ≤ 1	C C	D iso:delPoi	nt <sup>o</sup>
C iso DatasetSeries	iso:city 2 0		D iso:countr	0
© iso:DatatypeCd	iso:city ≤ 1	5	D iso:city O	
© iso:Date	liso:country ≥ 0	5		
© iso:DateRef	S iso.country ≤ 1			
© iso:DateTime	iso:sMailAdd > 0		- I Standard	
© iso:DateTypCd	iso postCode 2 0		Disj. V	C + H - X
🤤 iso:DigTranOps 🔤	S iso postCode ≤ 1			
C iso:DimNameTypCd				
Ciso:Dictorco	S iso:postCode (type=ow	l:DatatypeProperty)		
C iso Distrib	<u>1</u>			CX
© iso:Distributor	Name Labole Foute	alant Dranartine	Annotations	VC
© iso:EllParas	Indine Labers Equiv	alean Properties	Annotations	VL
© iso:EvalMethTypeCd	iso:postCode		Property	Value
© iso:ExtEleInfo				
© iso:Extent	Documentation			
e iso:FeatTypList	ZIP or other postal code	ZIP or other postal code		
Siso:FeatureTyp				
Ciso Generichame	1.			
C iso GeoObiTypCd	Same content of	Taul Association		
© iso:GeometObis	Domain defined	Range	[])	Allows multiple values
© iso:Ident	Domain u V +	- String		harmon Frankland
© iso:ImgCondCd	10 million	• <b>String</b>	U	inverse Functional
🕒 iso:InifTypCd	C/Iso.Address	XMI Schema Data	shme	
Siso:KeyTypCd		Time ostitutid Date		
Siso:Keywords		string	•	
Siso MainErrac Cd				
Ciso Mainfinfo				
- 100.mail.min				
C iso:MdExtinfo				



#### Challenge: Scalability (quantity+time)

#### Ontology Editor (Protégé) v

	m Files Protege_2.0%so-metadata.pp	rj, OWL File	J ~ `	
roject Edit Window OWL Help			-	
OWLClasses P Properties	AR P Forms	Metadata	S	searc
sserted Hierarchy V C × IV	C iso:Address (type=owt:Class)			
ow Thing	C iso:Address (type=owl:Class)	e		inring
C Iso:Address			uu	
🤤 iso:Angle	Name Labels			
C iso:AppSchInfo	Land Lands			
C iso:AttributeTyp	ISO:Address		con	
- © iso:Binary	Documentation			
iso:BrowGraph	location of the responsible individu	ualor		
O iso:CellGeoCd	organization	Section 1		
► © iso:Cita				
- Çiso:Citation	Asserted Inferred		At Class	N
© iso:ClasscationCd				
C iso:Continto	Asserted Conditions	C C	PII Pro	pertie
© iso:Contact		NECESSARY & SUFFIC	ENT	Na
- 🥥 iso:ContentTypCd	() owl Thing	NECESS	D iso:po	Iminărea O
Ciso:DQElement	liso:adminArea≥ 0		E D iso:et	fallAdd <sup>o</sup>
O Iso:DuScope	iso:adminArea ≤ 1		C D iso:de	elPoint <sup>O</sup>
© iso:DatasetSeries	iso:city≥ 0		D iso:cc	untry
🔘 iso:DatatypeCd	iso:city≤1	-	D ISO:CI	An I
😳 iso:Date	so country ≤ 1	-	<u>-</u>	
Siso:DateRef	(≥) iso:delPoint≥ 0		<u> </u>	
- Ciso:DateTime	liso:eMailAdd ≥ 0		C Die	VC+H-Y
© iso:DigTranOps	l≥ iso:postCode≥ 0			
🕒 😉 iso:DimNameTypCd	AS Iso postcode 1	L		
Siso:Dimen	S iso:postCode (type=owl:Datat	ypeProperty)		- ^ •
© iso:Distance				CX
© iso:Distributor	Name Labels Equivalent P	Properties	Annotations	VCI
© iso:EllParas	Contrast Con		Bronortu	Volue
Siso:EvalMethTypeCd	lisopostCode		Property	value
C iso Extent	Documentation			
© iso:FeafTypList	7IR or other postal code			
© iso:FeatureTyp	Lir bi oner postar code			
©iso:Format				
Giso:GenOhiTupCd				1.000M 2850 205
© iso:GeometObis	Domain defined	Range		Allows multiple values
≻©iso:Ident	Domain u V + -	String	•	Theorem Eulerional
e iso:ImgCondCd	C) iso Address			
U ISO:InitTypCd		XML Schema Dat	tatype	
C iso:Keywords		string	-	
© iso:Lineage				
🔘 iso:MainFreqCd				
C iso:Maintinfo				
The second	8			

Which system would you prefer for full-text search, aggregate queries, re-use (e.g. copy&paste) ?

link Hageman

#### Sticky Notes (PostIt<sup>™</sup>)

#### Challenge: Loss of Knowledge

- Loss of structure/semantics
  - Conversion of KI between applications
  - Between communication partners
    - Internalisation: KIs come with certain structure (email, documents, WWW)
    - Externalisation: Publish knowledge (email, documents, WWW)





Original idea: M. Boettger, 2005, PKM and ``cues to knowledge"



- No structure at all: very high internalisation costs
- Even with very high formality, some internalisation costs remain (e.g. time to read and understand)
- Externalisation costs roughly proportional to effort spend
  - we assume the effort has some effects
- → there might be a sweet spot for the total costs

#### We need a way to quantify costs

#### Unified Knowledge Model (UKM)

**Definition**: A knowledge item / is the smallest unit of content in the UKM.

- A knowledge item is either
- a snippet of content which can contain something between a single word up to a sentence, or
- a knowledge item is a statement
   (/x /x /) between other knowledge items

Aggregate queries and semantic queries (reasoning) can retrieve more knowledge than put in!



sk Hageman

landia Stern

Over the lifetime of a PKM system
 A set of knowledge items x created

Cost Model

– A set of tasks *t* performed

## Gain G = Benefit - Costs $G = \sum_{t} B(t) - (\sum_{x} C_{E}(x) + \sum_{t} C_{R}(t))$



#### Value Of Knowledge (Benefit)

- How to measure value?
  - Time needed to re-create the knowledge?
  - The value of knowledge does not exist as such (Iske and Boekhoff, 2002)
- → Change of value in the world resulting from the action taken because of the knowledge
- $\rightarrow$  In practice: knowledge item has value 1 or 0

<ul> <li>□</li> </ul>	t
Use 🧔	

Benefit of use (B)



#### Cost Model: Externalisation

- Write something down
   Create a knowledge item
- Connect knowledge items

   Create links between KI
- Structure knowledge
  - Split one KI into several smaller, connected KIs
     e.g. format text into pragraphs, headlines; mark something in bold
- Formalise knowledge
  - Assign formal type to KIs
  - Assign formal semantics to links



Cost of Externalisation  $(C_E)$ 

#### **Cost Model: Externalisation Operations**

- Add/delete/update content
  - Proportional to number of added symbols
- Add/delete/update formal statement
  - Varies with semantic consequences of statement (changes in a type hierarchy require more thinking time than putting a note in a folder)



Cost of Externalisation  $(C_E)$ 



#### Cost Model: Retrieval per Task

Information Retrieval Process (Bates 2002):

- Searching (query  $\rightarrow$  results)
  - Cost of formulating a query
- Browsing (scanning lists)
  - Cost of evaluating a result
- Following Links

Personal X Information 14 Retrieval

Cost of Retrieval (C<sub>R</sub>)

$$C_R(t) = C_{ql}(t) + k(t)e$$

number of

by query

items retrieved

- Use/consume/read/transform knowledge item  $B(t) = p_t k(t)$ 
  - Cost of use proportional to size
  - The only process step that can bring value

Precision of search (probability an item has value=1)

**Results: text or formal statements** 

#### Complete Cost Model

$$G = \sum_{t} (B(t) - C_R(t)) - \sum_{x} C_E(x)$$
  
= 
$$\sum_{t} k(t) (p_t - e) - C_{ql}(t)$$
  
= 
$$\tau = \tau = \tau = \tau$$

 $- \tau_{content_a} - \tau_{content_d} - \tau_{stmt_a} - \tau_{stmt_d}$ 

$$\tau_{op} = n_{op} c_{op}$$



#### How To Get Value

- 1. Externalise only relevant knowledge
- At the right degree of formality → formal queries can save a lot of time
- 3. Search at the right moment
- 4. Invest in restructuring/formalisation



#### Conclusions

8 O

- 1. Important factors for cost/benefit ratio are:
  - Granularity (size) of knowledge items
  - Degree of formality
- 2. Look at the complete process (externalisation, retreival)
- 3. Look at indivdual processes/incentives first, then at the team/enterprise/community
- 4. Future of knowledge society depends on ability to further decrease costs of knowledge management
   → requires more semantics
  - THANK YOU.
- Future work: Tool for Semantic PKM, see <u>http://cds.xam.de</u>

Get these slides from <a href="http://pubs.xam.de">http://pubs.xam.de</a>







#### Prototype for Semantic PKM

- http://cds.xam.de
- Looking for private beta users, send me an email



#### Comparison



	Organisational KM	Personal KM		
Perspective	Enterprise, top-down	Individual, bottom-up		
Changes	Fluctuation of employees	Change of employer		
Goal	Increase productivity			
Degree of Formality	explicit ("publication")	informal ("note")		
Context	Job	Job and private		