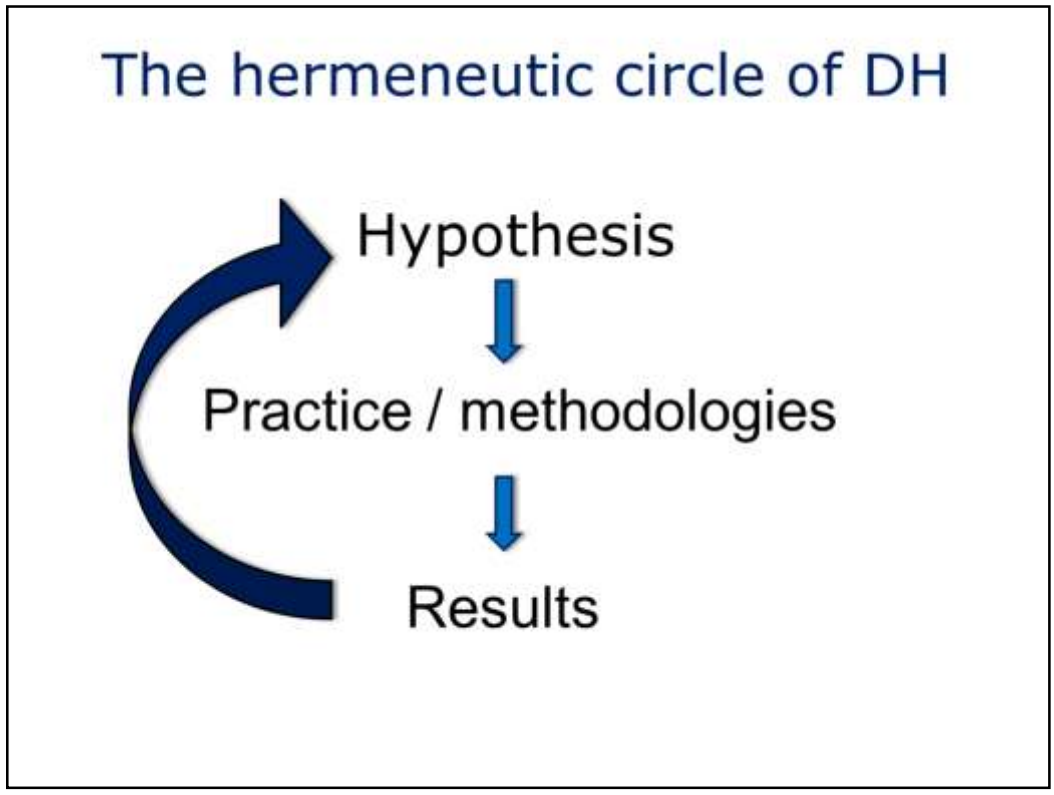
The socio-cultural foundations of the scholarly digital edition

Domenico Fiormonte, Università Roma Tre

La Sapienza DigiLab – Rome 3-5 December 2014

(domenico.fiormonte@uniroma3.it)

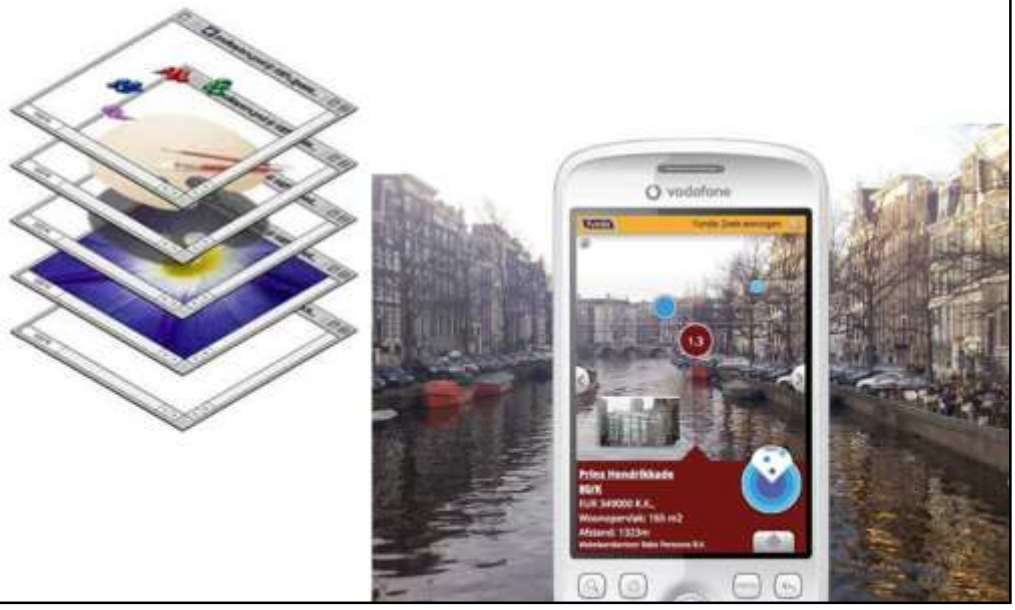


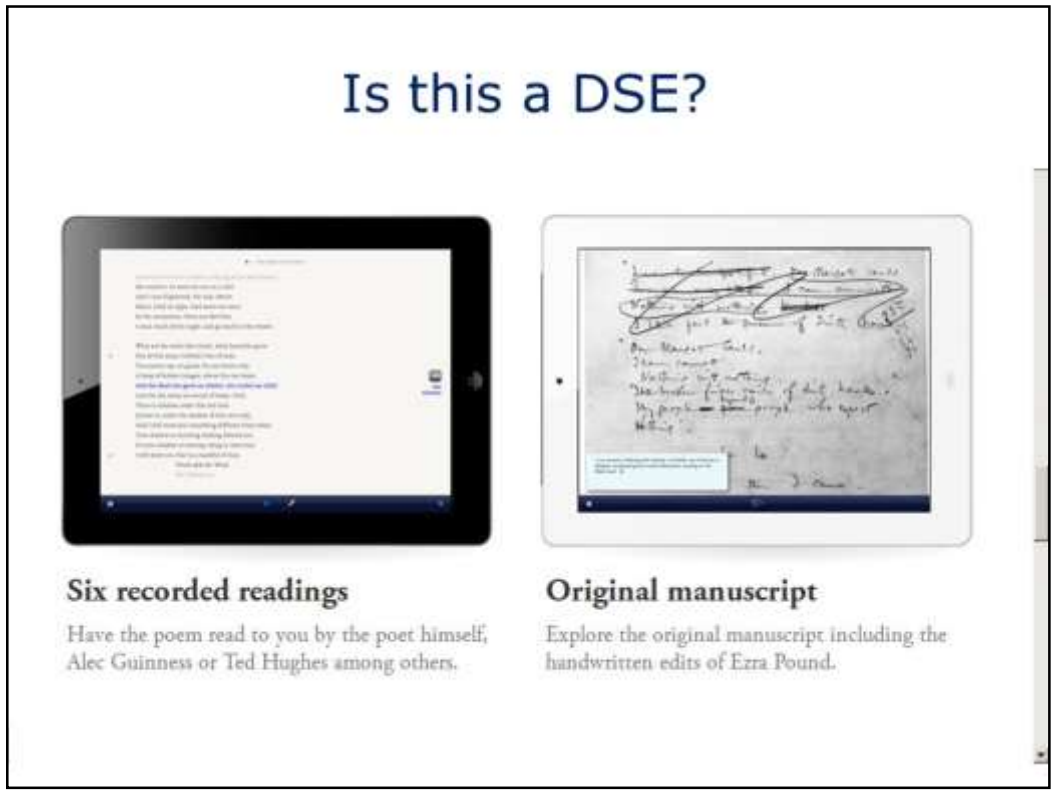
Il circolo ermeneutico delle DH: i risultati ridefiniscono costantemente gli obiettivi e le ipotesi iniziali

I. The scenario

(Or some preliminary facts that we tend to remove…)

Where can we find philology today?

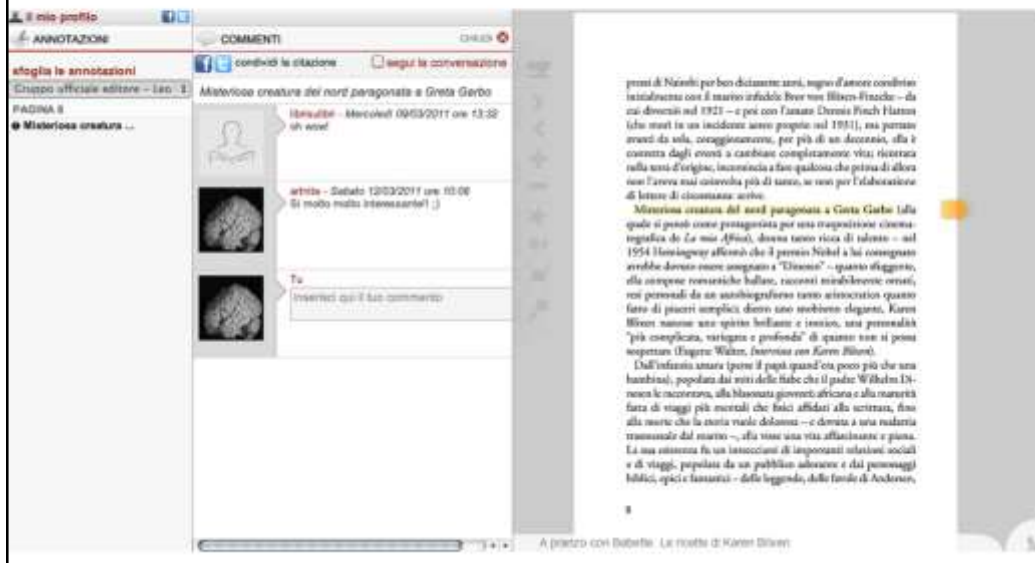




http://www.touchpress.com/titles/thewasteland/

http://youtu.be/rlhosnfP-Jw

And how about the web 2.0?



E-philology or cultural informatics?

- "In the process of the translation from physical and electronic media technologies to software, all individual techniques and tools that were previously unique to different media "met" within the same software environment. This meeting had fundamental consequences for human cultural development and for the media evolution. It disrupted and transformed the whole landscape of media technologies, the creative professions that use them, and the very concept of media itself" (Manovich, *Software culture*)

Il remix è "la logica culturale del capitalismo globale" un'estetica della variazione, della stratificazione e della riscrittura ove è sempre possibile aprire una finestra sul processo delle "forme variabili in continua mutazione".

In my view, this ability to *combine previously separate media techniques represents a fundamentally new stage in the history of human media, human semiosis, and human communication*, enabled by its "softwarization." I describe this new stage in media evolution using the concept of *hybridity*. In the first stage, most existing media were simulated in a computer and a number of new types of media that can only be realized in a computer were invented. In the second stage, these simulated and new mediums started exchanging properties and techniques. To distinguish these processes from more familiar remixes, I introduce the new term *deep remixability*. Normally a remix is a combination of content from a single medium (like in music remixes), or from a few mediums (like Anime Music Video works which combine content from anime and music video). However, the software production environment allows designers to remix not only the content of di!erent media types, but alsocombine previously separate media techniques represents a fundamentally new stage in the history of human media, human semiosis, and human communication, enabled by its "softwarization." I describe this new stage in media evolution using the concept of hybridity. In the first stage, most existing media were simulated in a computer and a number of new types of media that can only be realized in a computer were invented. In the second stage, these simulated and new mediums started exchanging properties and techniques. To distinguish these processes from more familiar remixes, I introduce the new term deep remixability. Normally a remix is a combination of content from a single medium (like in music remixes), or from a few mediums (like Anime Music Video works which combine content from anime and music video). However, the software production environment allows designers to remix not only the content of different media types, but also their fundamental techniques, working methods, and ways of representation and expression. their fundamental techniques, working methods, and ways of representation and expression.

(Manovich, Software takes command, p. 46)

Code

- "Codes—by name and by matter—are what determine us today, and what we must articulate if only to avoid disappearing under them completely. They are the language of our time precisely because the word and the matter code are much older…"
- "Today, technology puts code into the practice of realities, that is to say: it encodes the world" Friedrich Kittler

Il nodo della questione: siamo noi a scrivere il codice o è il codice a produrre noi?

II. The power of representation(s)

(Introducing the problem of representation in culture and technology)

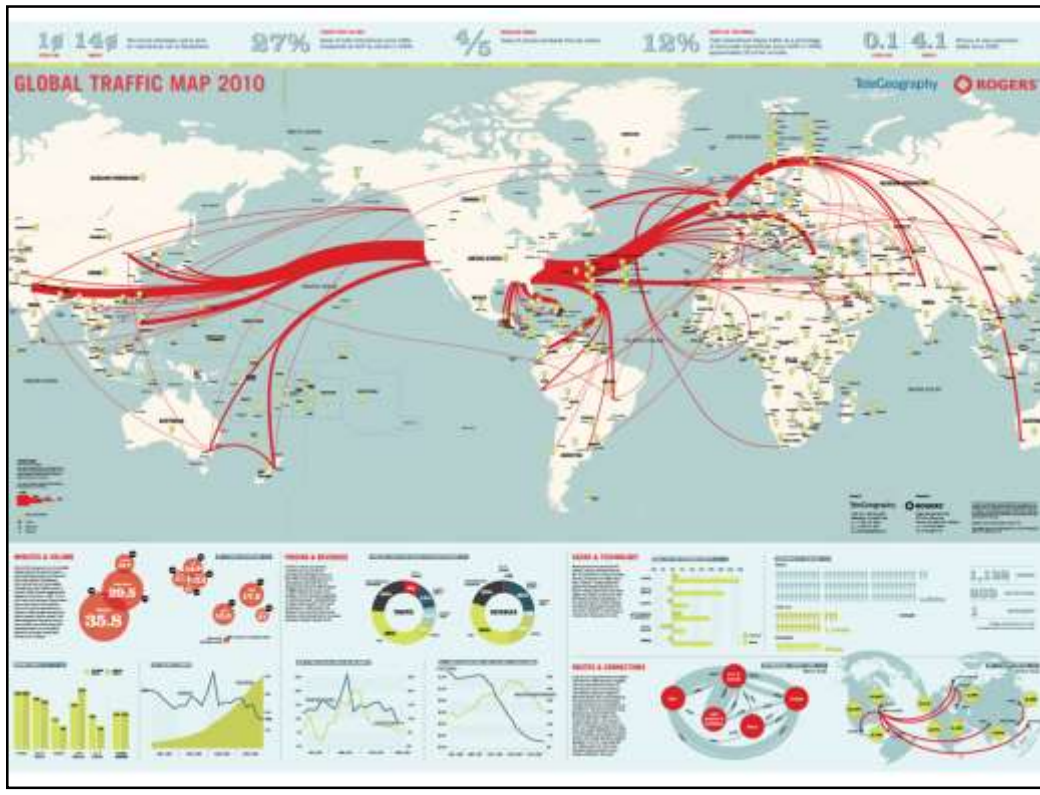
From life to cultural representations

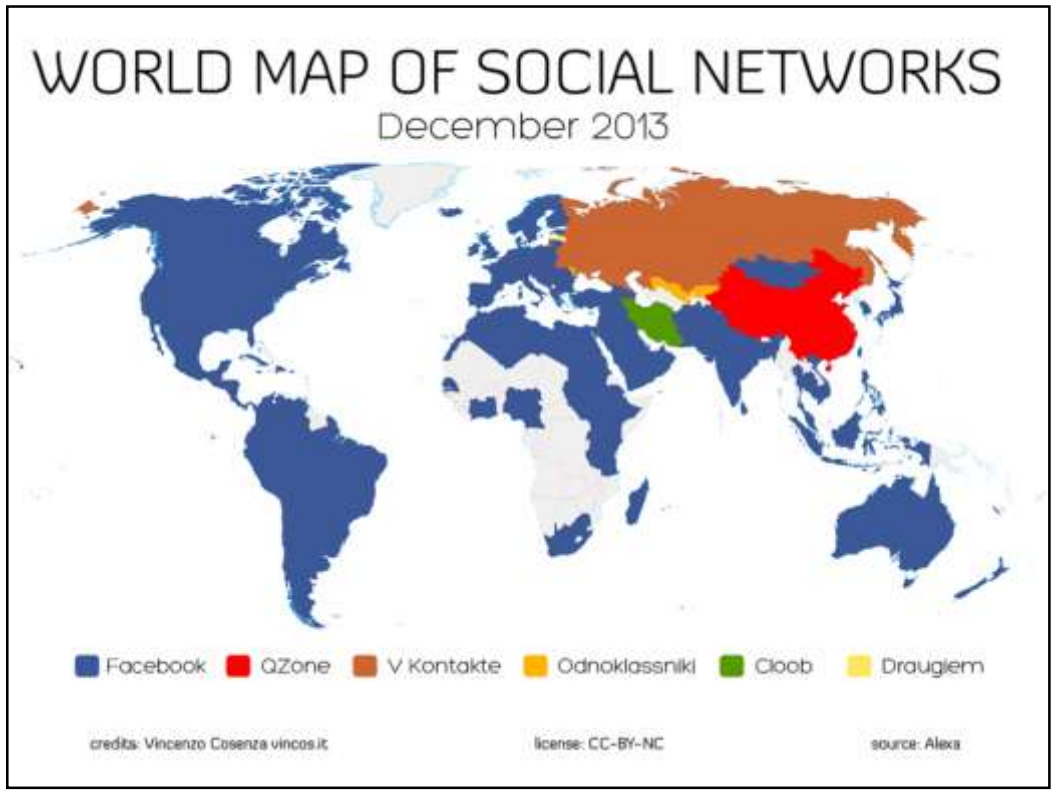
- The paradox of culture: dialectic between stability and variation (Sperber 2005)
- "Culture can be defined as a viral phenomenon, the continual exchange of representations in the everyday communication" (Benozzo 2011)
- Non-western paradigm: "knowledge is not the *accumulation* or *control* of savoirs, but the expansion of consciousness" (Torella 2008)

Maps, graphs, trees…

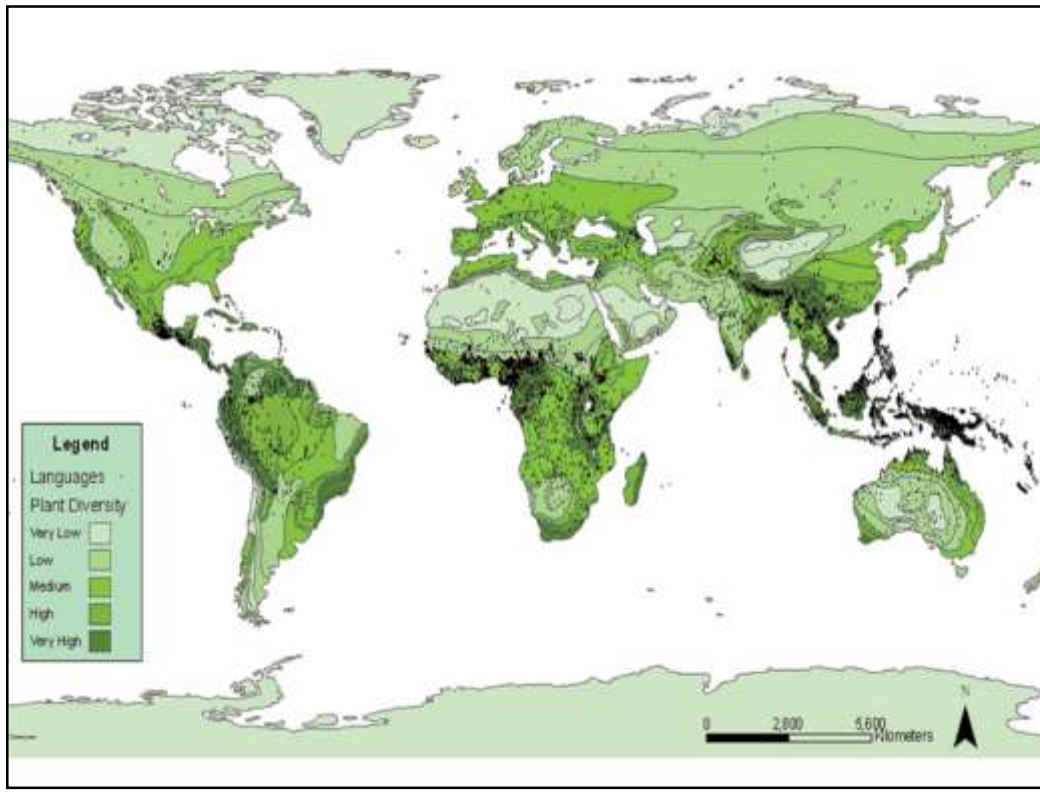
- In the following slides we will observe strange overlapping between maps from different contexts…
- It is a way for introducing us to the world of [digital] cultural representations

Orbis Typus Universalis, world map drawn by Martin Waldseemüller and first published in 1513.

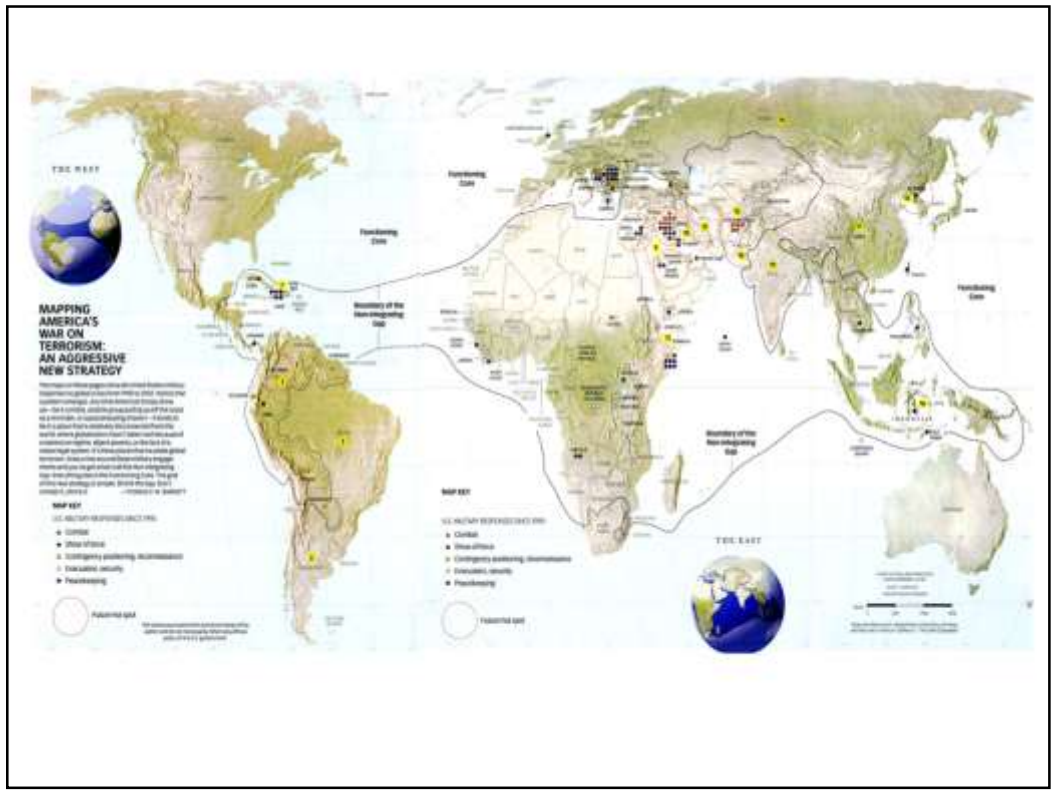

GLOBAL TRAFFIC MAP 2010
TeleGeography
ROGERS



FB is the dominant social network in in 127 out of 137 countries analyzed: in the latest months FB lost Kyrgyzstan, but stole Syria from Maktoob. FB has now 1,189 billion monthly active users, but it is growing less rapidly than before (*it has added just 34 million active users in 6 months*). 351 million users in Asia, 276 million in Europe, 199 million in US & Canada, 362 million in remaining countries.



Over the past decade began to emerge the idea that linguistic diversity is inextricably linked to biodiversity (Maffi 2001). The Index of Biocultural Diversity adopted in 2004 by a group of interdisciplinary scholars that later will found the terralingua.org international NGO, shows that three areas of the world "emerge as 'core areas' of exceptionally high biocultural diversity: the Amazon Basin, Central Africa and Indomalaysia/Melanesia" (Maffi 2010: 6).


MAPPING AMERICA'S WAR ON TERRORISM: AN AGGRESSIVE NEW STRATEGY
MAP KEY
Combat
Show of force
Contingency positioning, reconfiguration
Evacuation / security
Peacekeeping
Future hot spot
MAP KEY
Combat
Show of force
Contingency positioning, reconfiguration
Evacuation / security
Peacekeeping
Future hot spot

The reality show

- "This disturbing map is from the Pentagon's new war on terror; part of their new, more aggressive strategy under the Bush presidency. (…) what is interesting about it is that it contains these regions marked out with dotted lines, called the boundary of the *non-integrated gap*, which end up encompassing the center of the world. The heartland and origin of civilization is now defined as the non-integrated gap, **that's where we all came from**." (Bowker 2010)

The "cultural law" of the artifact

- Material and cognitive artifacts are subject to the influence of their environment and culture, and depend on the social habits of individuals and groups that devise and make use of them
- The artifact influences and at the same time is influenced by its context
- Technology is always a part of culture, not a cause or an effect of it

The bias of communication

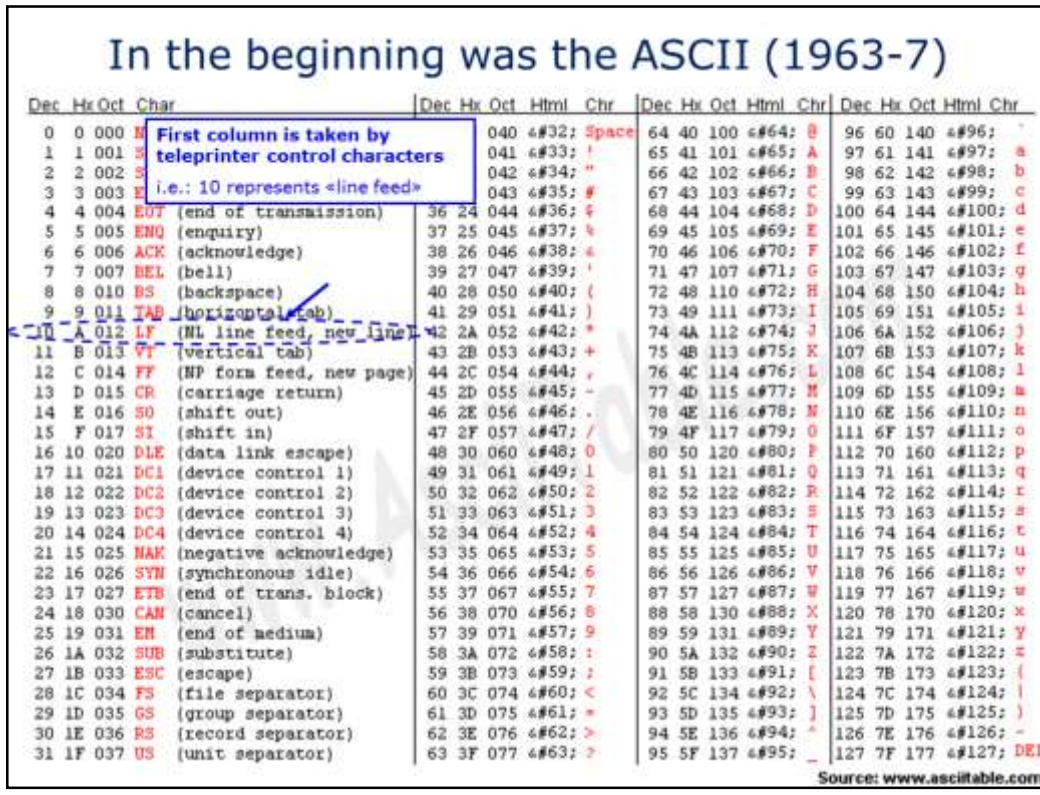
- "All forms of communication imply a *bias*" (Innis 1951)
  - "Innis argued that any form of communication possessed a bias; by its nature it was most adept at reducing signaling time and controlling space or strengthening collective memory and consciousness and controlling time. This bias hardened into a monopoly when groups came to control the form of communication and to identify their interests, priestly or political, with its capacity." (Carey 1992: 167)
- The main difference between cultural evolution and natural evolution lies in the active role played by self-descriptions, i.e. in the influence that representations [models] have on the object (Lotman/Uspenskij 1977)

La prima frase in realtà già riassume benissimo la mia visione delle DH e dell'edizione.

Le questioni teoriche che discuto qui originano dalla cosiddetta Scuola di Toronto (Innis, McLuhan, Havelock, Ong) e i communication studies (Carey, Slack). Link a Innis 1951: http://www.media-studies.ca/articles/innis.htm

Riprendi anche le nozioni di Lotman sulla semiotica della cultura, come cornice con la quale arricchire i fondamenti dell'edizione digitale. Lotman parla del «modello»! Una autodescrizione normativa diventa una metalingua nel momento in cui assume un ruolo di modello per lo scambio e la trasmissione culturale (es. la filologia). Rappresentazione culturale è altro termine chiave: esso corrisponde a una serie di fenomeni, strumenti, eventi e discipline che formano l'immagine della cultura.

Ricerca realizzata ad aprile 2008



Dec	Hx	Oct	Char	Dec	Hx	Oct	Html	Chr	Dec	Hx	Oct	Html	Chr	Dec	Hx	Oct	Html	Chr
0	0	000	NUL	32	20	040		Space	64	40	100	@	@	96	60	140	`	`
1	1	001	SOH	33	21	041	!	!	65	41	101	A	A	97	61	141	a	a
2	2	002	STX	34	22	042	"	"	66	42	102	B	B	98	62	142	b	b
3	3	003	ETX	35	23	043	#	#	67	43	103	C	C	99	63	143	c	c
4	4	004	EOT (end of transmission)	36	24	044	$	$	68	44	104	D	D	100	64	144	d	d
5	5	005	ENQ (enquiry)	37	25	045	%	%	69	45	105	E	E	101	65	145	e	e
6	6	006	ACK (acknowledge)	38	26	046	&	&	70	46	106	F	F	102	66	146	f	f
7	7	007	BEL (bell)	39	27	047	'	'	71	47	107	G	G	103	67	147	g	g
8	8	010	BS (backspace)	40	28	050	(	(	72	48	110	H	H	104	68	150	h	h
9	9	011	TAB (horizontal tab)	41	29	051	)	)	73	49	111	I	I	105	69	151	i	i
10	A	012	LF (NL line feed, new line)	42	2A	052	*	*	74	4A	112	J	J	106	6A	152	j	j
11	B	013	VT (vertical tab)	43	2B	053	+	+	75	4B	113	K	K	107	6B	153	k	k
12	C	014	FF (NP form feed, new page)	44	2C	054	,	,	76	4C	114	L	L	108	6C	154	l	l
13	D	015	CR (carriage return)	45	2D	055	-	-	77	4D	115	M	M	109	6D	155	m	m
14	E	016	SO (shift out)	46	2E	056	.	.	78	4E	116	N	N	110	6E	156	n	n
15	F	017	SI (shift in)	47	2F	057	/	/	79	4F	117	O	O	111	6F	157	o	o
16	10	020	DLE (data link escape)	48	30	060	0	0	80	50	120	P	P	112	70	160	p	p
17	11	021	DC1 (device control 1)	49	31	061	1	1	81	51	121	Q	Q	113	71	161	q	q
18	12	022	DC2 (device control 2)	50	32	062	2	2	82	52	122	R	R	114	72	162	r	r
19	13	023	DC3 (device control 3)	51	33	063	3	3	83	53	123	S	S	115	73	163	s	s
20	14	024	DC4 (device control 4)	52	34	064	4	4	84	54	124	T	T	116	74	164	t	t
21	15	025	NAK (negative acknowledge)	53	35	065	5	5	85	55	125	U	U	117	75	165	u	u
22	16	026	SYN (synchronous idle)	54	36	066	6	6	86	56	126	V	V	118	76	166	v	v
23	17	027	ETB (end of trans. block)	55	37	067	7	7	87	57	127	W	W	119	77	167	w	w
24	18	030	CAN (cancel)	56	38	070	8	8	88	58	130	X	X	120	78	170	x	x
25	19	031	EM (end of medium)	57	39	071	9	9	89	59	131	Y	Y	121	79	171	y	y
26	1A	032	SUB (substitute)	58	3A	072	:	:	90	5A	132	Z	Z	122	7A	172	z	z
27	1B	033	ESC (escape)	59	3B	073	;	;	91	5B	133	[	[	123	7B	173	{	{
28	1C	034	FS (file separator)	60	3C	074	<	<	92	5C	134	\	\	124	7C	174			|
29	1D	035	GS (group separator)	61	3D	075	=	=	93	5D	135	]	]	125	7D	175	}	}
30	1E	036	RS (record separator)	62	3E	076	>	>	94	5E	136	^	^	126	7E	176	~	~
31	1F	037	US (unit separator)	63	3F	077	?	?	95	5F	137	_	_	127	7F	177	&#127;	DEL

95 caratteri ASCII stampabili, dal 32 al 126. “ASCII developed from telegraphic codes. Its first commercial use was as a 7-bit teleprinter code promoted by Bell data services.”

The problem of standards

- "The control of standards is a central feature of economic life… classifications and standards are material, as well as symbolic." (Bowker & Leigh Star 1999)
- "Any attempt to create an obligatory system of classification, rigid and universal, will result in residual categories (…) It is necessary to root the awareness of what happens every time one tries to standardize. In other words, in this creation there is someone who wins and someone who loses. (Leigh Star 2006)

Who does control standards?

- ASCII old style (126 characters) was the encoding standard of the Net until 2007
  - Most commonly used web languages, like HTML e XML, are still based on the 126 symbols of the ASCII table
- But what does ASCII mean?
  - *American Standard Code for Information Interchange*
- Where does it come from?
  - ASCII developed from telegraphic codes. Its first commercial use was as a seven-bit teleprinter code promoted by Bell data services
  - They were adapted to provide a user interface to early mainframe computers and minicomputers, sending typed data to the computer and printing the response
  - Almost 1/3 of symbols encode teletype symbols and *not text*!

Muchos de los caracteres de control ASCII servían para marcar paquetes de datos, o para controlar protocolos de transmisión de datos (por ejemplo **ENQ**uiry, con el significado: ¿hay alguna estación por ahí?, **ACK**nowledge: recibido o "acuse de recibo", Negative AcKnowledge: No recibido, Start Of Header: inicio de cabecera, Start of TeXt: inicio de texto, End of TeXt: final de texto, etc.). **ESC**ape y **SUB**stitute permitían a un protocolo de comunicaciones, por ejemplo, marcar datos binarios para que contuviesen códigos con el mismo código que el carácter de protocolo, y que el receptor pudiese interpretarlos como datos en lugar de como caracteres propios del protocolo. (Fuente: Wikipedia, Abril 2008)

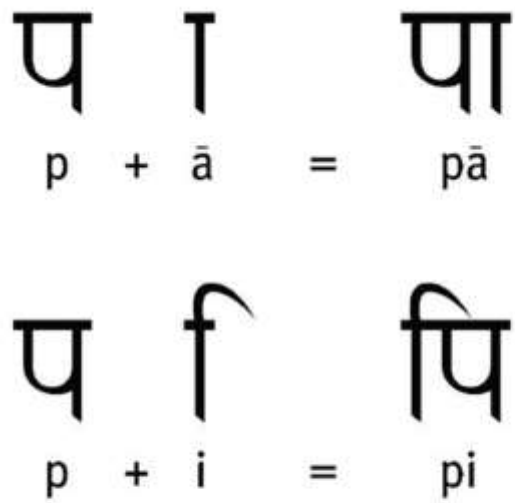


Fig. 2 - Fonte: Lussu, *La forma del linguaggio*, cit., p. 41.

Often in Indian systems aspects of a graphic nature prevail over the reading order of the graphemes. Unicode experts argue that Indic scripts are represented in its system according to a "logical scheme" that ignores "typographic" details (Perri 2009)

The encoding bias

- "But why on earth should the order of characters corresponding to the phonetic segment be considered logical by an Indian literate? Who says that the linearity of Saussure's alphabetic signifier should play a role in his writing practices?
- It is therefore all too evident that the alphabetic filter, the rendering software and the automatic process of normalization of Indic scripts are the result of *a choice that reflects the need for structural uniformity as opposed to the emic cultural practices of the real user*" (Perri 2009)

"The emic approach investigates how local people think" (Kottak, 2006): How they perceive and categorize the world, their rules for behavior, what has meaning for them, and how they imagine and explain things. "The etic (scientist-oriented) approach shifts the focus from local observations, categories, explanations, and interpretations to those of the anthropologist. The etic approach realizes that members of a culture often are too involved in what they are doing to interpret their cultures impartially. When using the etic approach, the ethnographer emphasizes what he or she considers important." (Wikipedia)

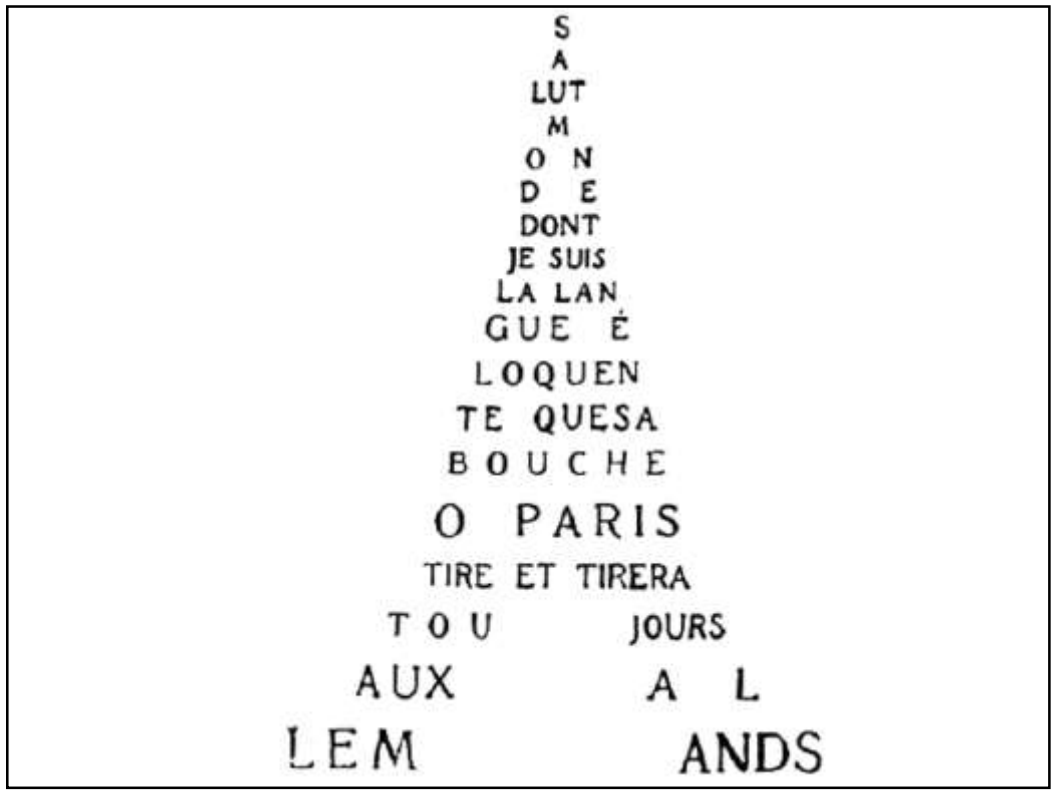
A Platonic standard

- "In Unicode – **an idealist, abstract, Platonic standard** – there is a clear distinction between the abstract character – which is the thing that Unicode provides a code for – and the glyph, or image, of the character."
- "When a symbol is to some degree iconopoeic, it is even less possible, less defensible, to divorce the abstract entity from its visual manifestation."
- "We are in part trying to illustrate, in very clear ways, the material and graphic dimensions of text, and the truth of Johanna Drucker's declaration: 'Design is information'. (Walsh – Hooper 2012)

iconopoeia = words depicting images.

"The liberty of invention: alchemical discourse and information technology standardization."

John A. Walsh, and Wallace Edd Hooper. *Literary and Linguistic Computing*, 27(1):55-79 (2012)


S
A
LUT
M
O N
D E
DONT
JE SUIS
LA LAN
GUE É
LOQUEN
TE QUESA
BOUCHE
O PARIS
TIRE ET TIRERA
T O U JOURS
AUX A L
LEM ANDS

RUMORGRAFIA
(SCRITTURA DEI RUMORI)
8-MAGGIO-1915
MILANO
DEPERO

Private (UNI)CODE

- Board of Directors
  - Google (1 member)
  - Microsoft (2)
  - Apple (1)
  - Intel (1)
  - IBM (1)
  - IMS - Appature (1)
- President of the Executive Office: Mark Davies, Google engineer since 2006
- UC Berkley is the only educational organization (Technical officers)

The "Lock-in" effect

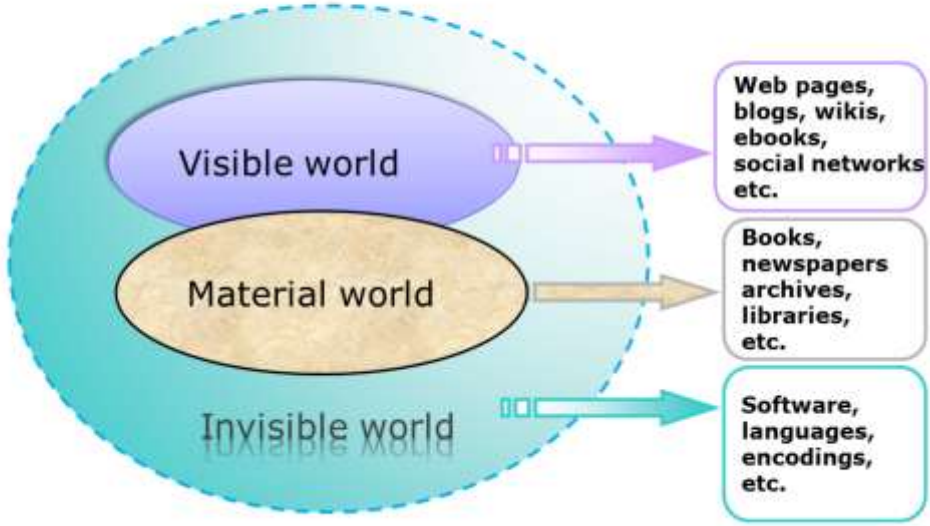
- "The fateful, unnerving aspect of information technology is that a particular design will occasionally happen to fill a niche and, once implemented, **turn out to be unalterable**."
- "Lock-in removes ideas that do not fit into the winning digital representation scheme"
- "Lock-in, however, removes design options based on what is easiest to program, what is politically feasible, what is fashionable, or what is created by chance." (Lanier 2010)

Jaron Lanier, *Your are not a gadget*, Allen Lane, 2010.
Niche: nicho
Unnerving: desconcertante
Feasible: viable, factible
By Chance: casualmente

III. From digital texts to digital editions

(Where I discuss the cultural foundations of our idea of 'representation', 'text' and 'edition')

The three dimensions of the digital edition



Digital encoding

- "Writing is metalinguistic" (Olson 1997)
  - Writing it's not just a transcription, but an abstract representation (i.e. a conceptual model) of speech
- Digital encoding is a representation and a "translation" of a cultural object by means of another semiotic system (a meta-language)
- In both cases (encoding and translation), the hermeneutic circle needs to be periodically renovated
- Encoding is a critical activity that entails to build a conceptual and formal model of our document

Olson, David (1997), *On the relations between speech and writing*, in Pontecorvo, Clotilde (ed.) (1997), *Writing development. An interdisciplinary view*, John Benjamins, Amsterdam-Philadelphia, pp. 3-20.

Ivanov, V. V. – Lotman, J. M. – Piatigorskij A. M. – Toporov, V. N. – Uspenskij, B. A. ([1973] 2006), *Tezisy k semiotičeskomu izučeniju kul'tur (v primenenii k slavjanskim tektstam)*, in *Semiotyka i Struktura Tekstu. Studia święcone VII międz. Kongresowi slawistów*, a c. di M. R. Mayenowa, Warszawa, pp. 9-32 (trad. it. *Tesi per un'analisi semiotica della cultura*, in Lotman, J. M., *Semiotica della cultura*, Roma, Meltemi, 2006, pp. 105-147.

Theoretical foundations of the DSE

- "So digital preservation is not a simple process of preserving physical objects but one of preserving the ability to reproduce the objects. The process of digital preservation, then, is inseparable from accessing the object. You cannot prove that you have preserved the object until you have re-created it in some form that is appropriate for human use or for computer system applications." (Thibodeau 2002)
- "Quello che interessa conservare non è la 'persistenza' dell'oggetto digitale nel formato originale, ma la sua 'operatività' (*performance*), e emerge l'esigenza di distinguere fra gli aspetti 'tecnici' o 'fisici', e gli aspetti 'concettuali' o 'intellettuali' dell'oggetto digitale". (Buzzetti 2003)
- Is that all?

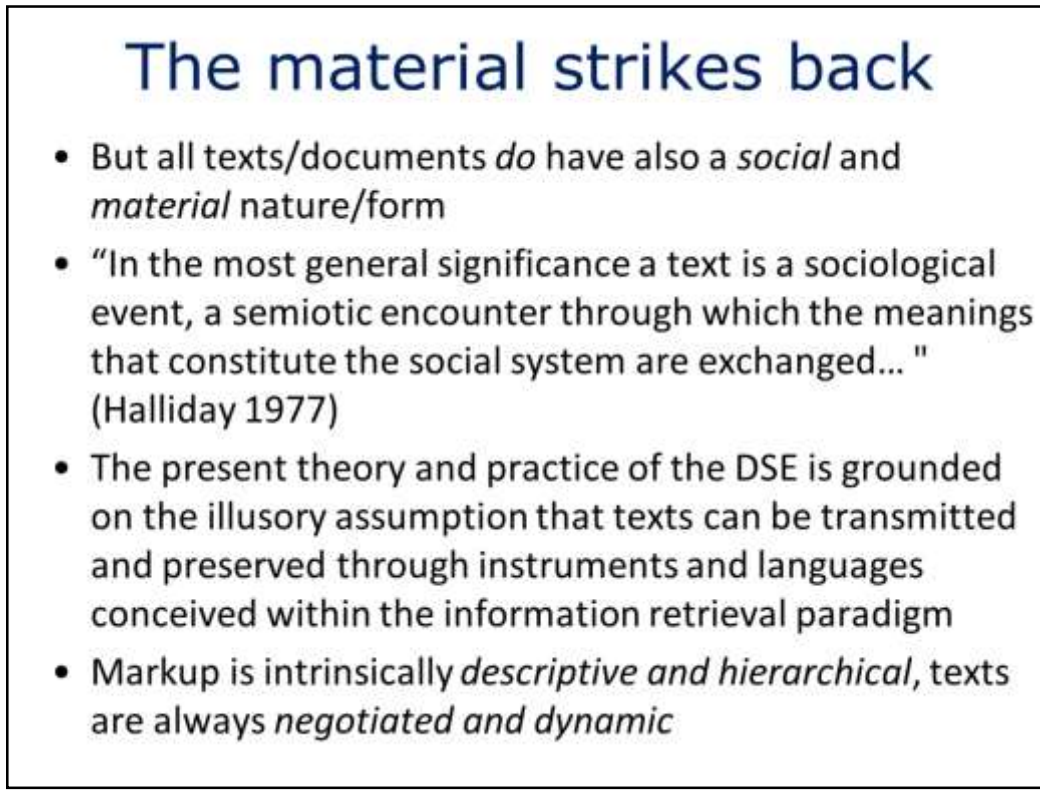
All digital objects are entities with multiple inheritance; that is, the properties of any digital object are inherited from three classes. Every digital object is a physical object, a logical object, and a conceptual object, and its properties at each of those levels can be significantly different. A *physical* object is simply an inscription of signs on some physical medium. A *logical* object is an object that is recognized and processed by software. The *conceptual* object is the object as it is recognized and understood by a person, or in some cases recognized and processed by a computer application capable of executing business transactions. (Thibodeau, http://www.clir.org/pubs/reports/pub107/thibodeau.html)

The immaterial image

- "The text does not have a **material nature**; that the text is only and always an **image**; … any attempt to identify it with a material witness whatsoever, even an autograph original, is an attempt to conceal [its] unavoidable problematic nature." (Buzzetti 2009)
- " Il testo è dunque soltanto un'immagine, virtuale se situata al termine della ricostruzione stemmatica, reale se risulta dalla lettura del suo **simulacro**, l'autografo. Sempre **un'immagine**. Un'immagine di un discorso. Questo discorso è l'algoritmo di una enunciazione linguistica; ma è possibile coglierlo solo come enunciato." (Segre 1985)

A DSE formal recipe

- Present DSE do not exploit "the distinctive features of a digital form of textual representation to obtain better critical and analytical results" (Buzzetti 2009)
- "An edition in digital form, then, is primarily considered here, besides its conventional use, as textual data to be processed. From this point of view, our primary concern becomes that of extracting a processable semantics from character data, by assigning them a functional structure and a suitable formalism." (Ibidem)



Dunque il problema non è (solo) che i linguaggi e strumenti informatici attuali siano inadeguati a rappresentare il testo (nella sua accezione di "corpo vivo"), ma che informatica e critica testuale di stampo strutturalista si sono alleate per perpetuare un'idea di testo come "oggetto" trasmissibile. (Di qui, per altro, l'ostilità verso le interfacce, unico punto di sutura con la realtà... gli utenti!) Esse sono il frutto di uno stesso paradigma epistemologico (qui cit. Benozzo) che considera i dati e non i contesti, concepisce il testo come "ricostruzione" ("immagine", rappresentazione, ecc.) e non come performance (la stabilità infatti è frutto solo delle nostre scelte, non deriva dal testo, come è chiaro in altre culture, per es. quella indiana), ed infine eredita dalla cibernetica l'ossessione del controllo e dalla scienza dell'informazione l'esigenza dell'information retrieval. Il testo così diventa un oggetto da "analizzare", ecc.

Dunque come è possibile pensare a un'edizione digitale senza pensare prima all'oggetto che abbiamo di fronte, alla sua storia e al suo contesto di fruizione? Paradossalmente, questa ossessione per la conservazione si scontra con la fragilità degli strumenti che la dovrebbero garantire. Come dico qui accanto il problema è come intendo la comunicazione, non i modelli che applico, ma qual è il contesto in cui nascono...

A tale of many editions

- "La edición crítica a la europea (que no se confunde, pues, con la *scholarly edition*) es uno de los artefactos más admirables que han producido las artes de humanidad
- En sustancia, se trata de un sistema de síntesis y ordenación de datos que contempla la existencia de diversos avatares del texto en cuerpo mayor y postula, implícita o explícitamente, que éste ofrece la versión auténtica de la obra en cuestión, mientras que las variantes suponen otras tantas desviaciones respecto a ella.
- (...) una edición sólo es hoy concebible *como parte de un proyecto de ediciones*." (Rico 2011)

«[L]a gerarchizzazione tipografica dell'edizione critica dà l'impressione che il testo che occupa il corpo della pagina sia l'unico

testo reale, e che le varianti a pie' di pagina non siano altro che incidenti minori. Così, a dire il vero, sminuisce, invece di rafforzarla, la percezione dell'effettiva pluralità testuale, che non consiste nel leggere tutti i testi contemporaneamente, ma nella consapevolezza che si sta leggendo solo uno dei diversi testi possibili. (...) Ad ogni modo, quando si può disporre facilmente di tutti i rappresentanti di una determinata opera, quando la tecnica consente di presentare una accanto all'altra tutte le sue versioni storiche e ipotetiche, la vecchia edizione critica cambia profondamente senso.» (Rico 2012)

Beyond the conservation paradigm?

- "Ho l'impressione che, finché si resterà nella situazione di una filologia come ripristino e *restitutio* (cioè di una filologia come flottazione e restauro), non verranno per niente regolati i conti con l'alterità di ciò che ci parla attraverso i documenti, con la storia vera e propria, con ciò che sta dietro i nostri manoscritti, in definitiva con ciò che i testi sono e raccontano." (Benozzo 2011)
- Ironically, this obsession with conservation clashes with the fragility of the digital instruments which are supposed to safeguard it
- So the problem is not only what models and methodologies I'm using, but the context in which they have originated

Towards a *living* paradigm of the text

- An Aristotelian prejudice permeates the paradigm of the standard theory of communication that inspires current DSE theories: a theory that ultimately would exchange the "dead body" of communication for its "living body" (Ronchi 2008)
- Textual variation, and writing artifacts in general, are similar to living phenomena: they are subject to change as to increase their chances to propagate and survive
- "The text is a continuous process. There is a constantly shifting relation between a text and its environment, both paradigmatic and syntagmatic." (Halliday 1977)

Beyond the DSE?

- “ L’edizione [digitale] diventa un artefatto problematicamente inserito in un processo di ricerca il cui fine ultimo può cambiare a seconda della prospettiva che si assume e degli interessi che la animano.
- Il che forse rende obsoleta la domanda sulla scomparsa o meno della filologia – perché casomai occorrerà interrogarsi su che cosa sta diventando la ricerca (e non solo quella umanistica). ” (Fiormonte - Schmidt 2014)

«[L]a gerarchizzazione tipografica dell’edizione critica dà l’impressione che il testo che occupa il corpo della pagina sia l’unico

testo reale, e che le varianti a pie’ di pagina non siano altro che incidenti minori. Così, a dire il vero, sminuisce, invece di rafforzarla, la percezione dell’effettiva pluralità testuale, che non consiste nel leggere tutti i testi contemporaneamente, ma nella consapevolezza che si sta leggendo solo uno dei diversi testi possibili. (...) Ad ogni modo, quando si può disporre facilmente di tutti i rappresentanti di una determinata opera, quando la tecnica consente di presentare una accanto all’altra tutte le sue versioni storiche e ipotetiche, la vecchia edizione critica cambia profondamente senso.» (Rico 2012)

From cultural variance to digital variation

(Our experience with digital editing)

From where we started?

In every phase of the process, the variants are not a plus one or worse, a useless set of faulty versions, but the very condition of survival of history. The non-*occurrence* of variants means the end of evolution and we *must* believe that it spells the end of that form of life: whether it is a species of plant, *animal,* or man or his stories. (Sobrero 2009: 72).

«La trasmissione della cultura non è da intendersi come "passaggio" di informazioni: non vi sono entità che passano, piuttosto si stabiliscono *links comunicativi che richiedono*, inevitabilmente, un atto di ri-creazione da parte del ricevente. Ciò che si comunica è un atto distinto che ri-crea e trasforma totalmente lo stimolo originario, in un differente universo mentale e con identità in qualche modo mutata. La cultura di un individuo o di un gruppo, più che collezione di informazioni, è quindi conoscenza plurale che, situata, in azione e inseparabile dalle pratiche, come nelle malattie, si espande, contagia e opera commistioni.»

"In ogni fase del processo **le varianti** non sono un "di più" o peggio un inutile insieme di versioni difettose, ma **sono la condizione stessa della sopravvivenza della storia.** Il non prodursi di varianti significa la fine dell'evoluzione e dobbiamo pensare che alla lunga comporti la fine della forma di vita: che si tratti di una specie vegetale, animale, dell'uomo o dei suoi racconti"

DV theoretical framework

- Filologia e critica del testo: ricostruzione e interpretazione del testo
- Critica genetica e critica delle varianti: il testo mobile
- Bibliografia testuale: aspetti visuali, materiali e culturali
- La psicologia della composizione: il processo
- Etnofilologia: considerazioni di tutti i reperti "culturali"
- Cinque che filoni danno luogo a:
  - Filologia computazionale: codifica come atto ermeneutico, automazione procedure ecdotiche
  - Filologia digitale: rappresentazione della genesi del testo, ruolo del pubblico, dei produttori, dei supporti, ecc.
  - Filologia 2.0: accesso e fruizione diretta ai documenti, coautorialità, piattaforme scientifiche, *knowledge site*

I principali assunti teorici dell'etnofilologia (Cfr. Benozzo, "Verso un'etnofilologia", *Ecdotica*, 2007) possono essere riassunti nei seguenti punti:

i testi antichi possono essere indagati non solo come testi che vivono in relazione con altri testi, ma in primo luogo come esperienze di comunicazione tra esseri umani;

i testi antichi, nella loro forma di manoscritti, sono reperti di codificazione formale che propagano immagini e modi di essere di una civiltà;

il campo di studio dell'etnofilologia comprende tanto i reperti scritti quanto documenti di tipo orale, onomastico, folklorico, gestuale, iconografico, archeologico e materiale;

lo scopo dell'indagine etnofilologica consiste principalmente nel riconoscere le fonti indagate come codificazioni mnemotecniche di immagini, parole e credenze;

come accade per le scienze fisiche, anche il contesto in cui si collocano le ricerche etnofilologiche si espande con l'espandersi delle conoscenze e delle domande che tali ricerche portano con sé;

l'etnofilologo non è un osservatore neutro, ma la sua indagine implica sempre (come in etnografia) una partecipazione attiva e un coinvolgimento di fondo nel proprio modo di approssimarsi a ciò di cui si occupa;

l'etnofilologia considera se stessa, al pari dell'etnolinguistica, anzitutto una

scienza sociale.

Variation and cultural interfaces

- Filologia e critica del testo in quanto metodi di ricostruzione del documento sono "meccanismi metaculturali" attraverso i quali le culture si diffondono e riproducono (Lotman)
- La stabilità testuale si basa su due principi:
  - Capacità di diffusione, penetrazione, inclusione, ecc.
  - Standardizzazione dei supporti
- La *mouvance* testuale è un nodo cruciale (e irrisolto) all'interno di questi meccanismi di trasmissione
- La disponibilità, sotto forma di un sistema multi-strato, di questa *mouvance*, è in grado di cambiare la percezione della cultura?

Uso Variance e mouvance solo per comodità terminologica, senza per questo aderire in toto ai postulati della critica genetica francese – che pure a mio parere ha dei meriti indiscutibili.
Note a margine di Lotman, *Tesi per un'analisi semiotica delle culture*: "… dato che l'equivalenza fra i testi non è identità la traduzione da un sistema testuale in un altro comporta sempre un certo elemento di intraducibilità" (p. 138).
Commento punto 1: ovviamente sono solo alcuni dei principi
Commento 2: la filologia è interfaccia della trasmissione culturale
Commento ultimo punto: in sostanza la tecnologia, per riprendere Lotman, è un device metaculturale. Ciò che appare chiaro, al di là dell'elemento tecnologico, è che questi linguaggi
possono essere considerati, sia dal punto di vista intrinseco che estrinseco, vere e proprie 'metalingue' capaci di rappresentare e tradurre la conoscenza.
Se le culture comunicano fra loro con l'ausilio di meccanismi metaculturali (Lotman – Uspenskij 2006: 149), fra cui le traduzioni, si può proporre l'ipotesi di considerare la codifica (e a maggior ragione una codifica attraverso metalinguaggi di markup) un «sistema modellizzante secondario» , cioè un tipo di (meta)traduzione culturale.
D'altra parte l'approccio semiotico-strutturale di Lotman era esplicitamente richiamato da Gigliozzi. (Lotman, Jurij M. (2006), *Tesi per una semiotica delle culture*, Meltemi, Roma.)

DV solutions (1996-2012)

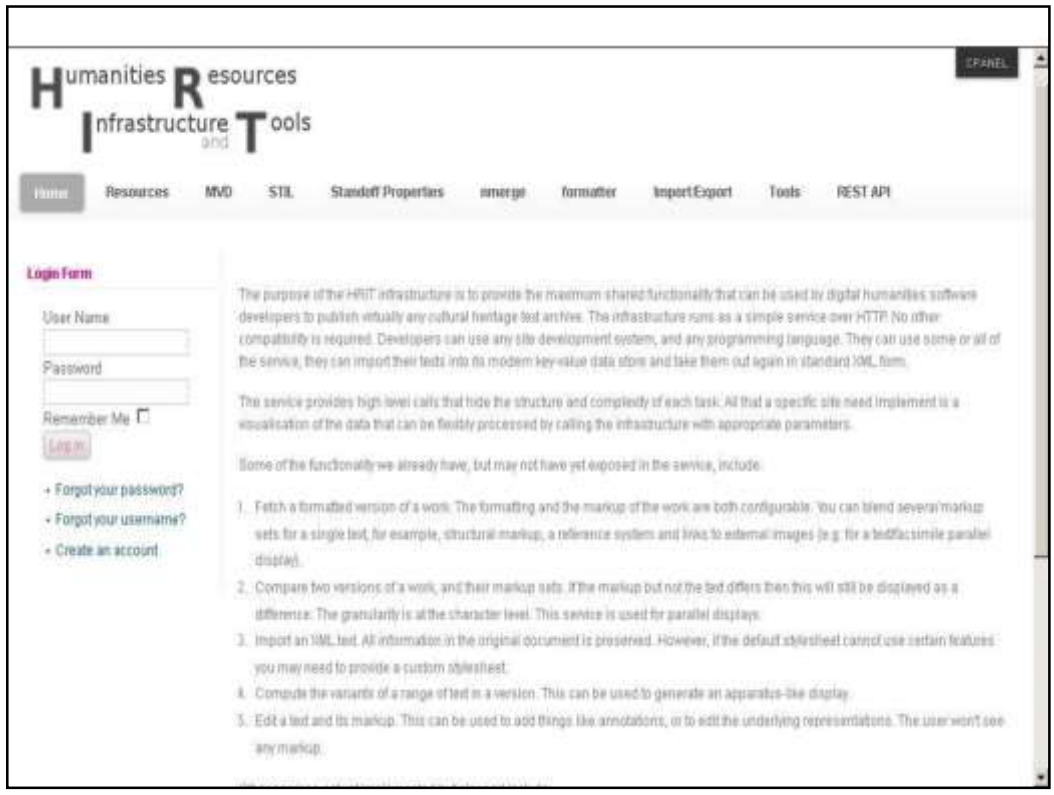
- Word:
  - Sanvitale
- HTML / Java Script
  - *all* ...
- SGML-TEI
  - Sanvitale
- Flip-Zooming
  - Sanvitale
- XML-TEI
  - Magrelli, Cerami
- Flash
  - Magrelli
- MVD
  - Cerami, Magrelli, Galiano, Petrolini, Vangelo della Sibilla, etc.
- Humanities Resources Infrastructure and Tools, AustESE

DV 1996-2001

DIGITAL
variants
[Versión Española] | [Versione Italiana]
Project Overview
Editorial Board
Authors
Texts
DV as a Teaching Resource
Information, links and credits
Computers, Literature and Philology (Rome 1999, Alicante 2000, Duisburg 2001)
Written and Spoken Italian: A one day workshop
Project overview

DV 2002-2007

Home
Events
News
Contact us
University of Edinburgh
variants
Authors
Texts
Teaching
forum
Forum
Search
Versión Española
Versione Italiana
Project overview
Digital Variants (=DV) is a contemporary authors digital archive founded in 1996 by the Department of Italian of the University of Edinburgh. The aim of the project is to make available on the Internet texts of living authors at different stages of writing.
Well-known writers of different literatures, at the height of their activity, agreed to open the "kitchen" of the text, showing us the complex writing phenomena lying under the final version of a work.
(read more...)
or choose a link on the left...

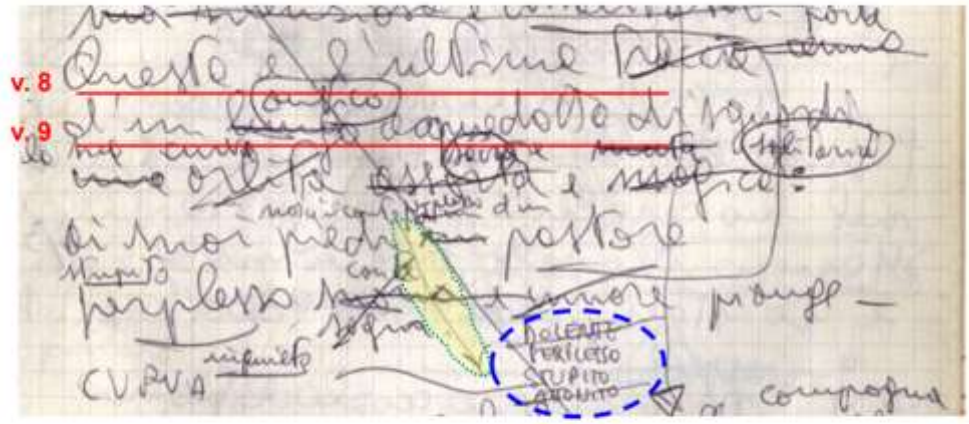

Humanities Resources
Infrastructure and Tools
CPANEL
Home
Resources
MVD
STIL
Standoff Properties
xmerge
formatter
Import/Export
Tools
REST API
Login Form
User Name
Password
Remember Me
Log in
• Forgot your password?
• Forgot your username?
• Create an account
The purpose of the HRIT infrastructure is to provide the maximum shared functionality that can be used by digital humanities software developers to publish virtually any cultural heritage text archive. The infrastructure runs as a simple service over HTTP. No other compatibility is required. Developers can use any site development system, and any programming language. They can use some or all of the service, they can import their texts into its modern key-value data store and take them out again in standard XML form.
The service provides high level calls that hide the structure and complexity of each task. All that a specific site need implement is a visualisation of the data that can be flexibly processed by calling the infrastructure with appropriate parameters.
Some of the functionality we already have, but may not have yet exposed in the service, include:
1. Fetch a formatted version of a work. The formatting and the markup of the work are both configurable. You can blend several markup sets for a single text, for example, structural markup, a reference system and links to external images (e.g. for a text/facsimile parallel display).
2. Compare two versions of a work, and their markup sets. If the markup but not the text differs then this will still be displayed as a difference. The granularity is at the character level. This service is used for parallel displays.
3. Import an XML text. All information in the original document is preserved. However, if the default stylesheet cannot use certain features you may need to provide a custom stylesheet.
4. Compute the variants of a range of text in a version. This can be used to generate an apparatus-like display.
5. Edit a text and its markup. This can be used to add things like annotations, or to edit the underlying representations. The user won't see any markup.

About Services People Events Publications Related Contact
The University Of Queensland
nectar
ESE
AustESE
Australian Electronic Scholarly Editing
eResearch tools to support the collaborative authoring and management of electronic scholarly editions
The AustESE project is a collaboration between The University of Queensland, University of NSW, Curtin University, University of Sydney, Queensland University of Technology, Loyola University, Chicago and the University of Saskatchewan. It is being project managed out of the ITEE eResearch Group at The University of Queensland.
Funding has been approved by NeCTAR (National eResearch Collaboration Tools and Resources) via the eResearch Tools program from June 2012 to Dec 2013. Contract negotiations completed and executed on 5 June 2012.
The aim of the AustESE project is to develop a set of interoperable services to support the production of electronic scholarly editions by distributed collaborators in a Web 2.0 environment.
The more specific objectives are to provide the Australian scholarly editing community with an online integrated 'workbench' that provides:
collation tools for automatically detecting, identifying and highlighting variations between different versions of a work and that allow the relationship between texts to be visualised, authored or edited e.g. building on Multi-Version-Document nmerge, CATT, CorTex and CollateX;
annotation tools that:
enable scholars to create and reply to scholarly commentary attached to texts, variants and images;

Worst-case scenario is now

What sort of model can we develop for representing the 'textual' structure of this object?

V. Magrelli, fragment of the manuscript of *Ecco la lunga palpebra*



Ecco la lunga palpebra della donna,
il sopracciglio vasto che attraversa
il pensiero dopo la pioggia
e lo illumina. Il suo arco
misura nel silenzio la sera
percorrendo assorto
la chiarità curva del cielo.
Questa è l'ultima porta
d'un antico acquedotto di sguardi.
(ver. def. Magrelli 1980)

Gli aggettivi *stupito*, *perplesso* e *inquieto* dell'ultimo verso potrebbero essere stati inseriti tutti nello stesso momento, oppure *stupito* e *inquieto* potrebbero essere stati inseriti in seguito ed essere considerati le due varianti di *perplesso*. Per questo li abbiamo considerati come una lista di varianti all'interno del verso, senza ricostruirne una successione cronologica. Immediatamente dopo troviamo altre due varianti (*canta* e *sogna*) che vanno a sostituire la parola cancellata (*suona*)

The 'old' solution: XML-TEI

```
<seg type="l">Questa è l'ultima
 <app>
  <rdg varSeq="1">traccia</rdg>
  <rdg varSeq="2"><del hand="M"
   type="overstrike">traccia</del><add hand="M"
   place="right">cenno</add></rdg>
   <rdg varSeq="3"><del hand="M"
   type="overstrike">cenno</del><add hand="M"
   place="supralinear">porta</add></rdg>
 </app>
</seg><lb/>
<seg type="l">d'un
 <app>
  <rdg varSeq="1">lungo</rdg
  <rdg varSeq="2"><del hand="M"
   type="overstrike">lungo</del><add hand="M"
   place="supralinear"><emph rend="circle"
   id="adj_2">antico</emph></add></rdg>
 </app> acquedotto di sguardi,
</seg><lb/>
```

Codifica del 1 rigo dell'autografo. in questo caso abbiamo utilizzato le indicazioni della TEI per la trascrizione dell'apparato critico di un manoscritto, racchiudendo la variazione testuale all'interno del tag <app> e dotando ogni lettura del verso <rdg> di un numero progressivo. Dunque la sequenza processuale è racchiusa all'interno di un apparato 'fittizio'.

```
<seg type="l">ai suoi piedi
 <app>
  <rdg varSeq="1">un pastore</rdg>
  <rdg varSeq="2"><add hand="M" place="supralinear">
   <note type="arrow" resp="M" place="foot">
    <list type="simple" id="adj_list2">
     <item><del hand="M" type="overstrike">DOLENTE</del></item>
     <item>PERPLESSO</item>
     <item>STUPITO</item>
     <item><del hand="M" type="overstrike">ATTONITO</del></item>
    </list>
   </note>
  d'un pastore</add></rdg>
  <rdg varSeq="3">nasce il canto
   <note type="arrow" resp="M" place="foot">
    <list type="simple" id="adj_list2">
     <item><del hand="M" type="overstrike">DOLENTE</del></item>
     <item>PERPLESSO</item>
     <item>STUPITO</item>
     <item><del hand="M" type="overstrike">ATTONITO</del></item>
    </list>
   </note>
   <del hand="M" type="overstrike">d'un</del>
   <add hand="M" place="supralinear">perplesso d'un</add>pastore</rdg>
 </app>
```

Codifica del 4 rigo dell'autogr. In seguito prob. all'inserimento di *nasce il canto* l'autore inserisce una freccia che rimanda a una lista dei quattro aggettivi in stampatello a piè di pagina. Questo tipo di segnali sono interpretabili come 'azioni' dell'autore vòlte a produrre certi effetti – ma non dei 'risultati', perché l'autore è ancora in fase di esplorazione. Per rendere conto di questo movimento abbiamo scelto di rappresentare la freccia attraverso una nota, utilizzando l'elemento paratestuale TEI <note> e dotandolo dell'attributo type="arrow", al cui interno si presenta la lista (<list>) degli aggettivi in alternativa (ciascuno contraddistinto da <item>). Poiché riteniamo che la lista/nota dei quattro aggettivi sia stata inserita in un secondo momento, la inseriamo all'interno del secondo passaggio (<rdg varSeq="2">); nell'immagine infatti la freccia si collega alla "o" di *[nasce il] canto* (frase aggiunta sopra la quarta riga).

Ricapitolando dunque, nella sequenza che codifica la quarta riga dell'autografo abbiamo inserito la 'descrizione' della lista di aggettivi a piè di pagina sotto forma di componenti (<item>) di una lista (<list>). Questa soluzione, che inverte la distribuzione spaziale degli elementi così come si presenta nell'a. (la lista si trova alla fine del **quinto** rigo del manoscritto, e non all'interno del quarto come nella codifica), costituisce un ulteriore esempio di come codificare il processo di scrittura possa implicare, attraverso il rimodellamento/ricostruzione operato dai linguaggi di markup, il rovesciamento della dimensione (della localizzazione) nello spazio. Ciò che nella realtà materiale del documento si presenta in una regione determinata (e dominata) dallo sguardo, nella dimensione pragmatica del processo si localizza su più livelli temporali. Poiché codificare vuol dire scegliere un punto di vista, nel nostro caso ci siamo sforzati di rappresentare, attraverso strumenti lineari quali il markup, qualcosa che lineare per sua natura non è. Una dimensione che l'autore ci indica "performativamente" attraverso segnali non esclusivamente testuali, quasi a mostrare che la dimensione del processo è cognitivamente irriducibile (nella durata creativa) alla sequenza gerarchico-lineare. (*Accenna anche al problema della ridondanza della ripetizione della lista 2 volte)*

What did we encode, really?

- Encoding of the writing process can imply, through the remodelling/reconstruction brought about by markup languages, the overturning of the spatial dimension.
- In other words, what appears in the *material* reality of the document as a particular region, determined visually, in the pragmatic dimension of the process may be located on several *temporal* levels.
- Since to encode means to select one point of view, in our case we are forced to represent, through the linear instruments of markup, something *which is not linear by its very nature.*
- This is a dimension which the author indicates "performatively" through signals that are not exclusively textual, as if to show that *the dimension of the process cannot be cognitively reduced to a hierachical sequence*. (Fiormonte – Martiradonna – Schmidt 2010)

«Ciò che nella realtà materiale del documento si presenta in una regione determinata (e dominata) dallo sguardo, nella dimensione pragmatica del processo si localizza su più livelli temporali. Poiché codificare vuol dire scegliere un punto di vista, nel nostro caso ci siamo sforzati di rappresentare, attraverso strumenti lineari quali il markup, qualcosa che lineare per sua natura non è. Una dimensione che l'autore ci indica "performativamente" attraverso segnali non esclusivamente testuali, quasi a mostrare che la dimensione del processo è cognitivamente irriducibile (nella durata creativa).»

alla sequenza gerarchico-lineare.

Conclusions

- Digital philology should not be limited to the preservation and reconstruction of cultural heritage documents, but it is required to act as an interface to knowledge
- There is no pre-existing language, model or structure that can be conceived and isolated from its cultural function and material context
- There should be no separation between the scope of the representation and the needs of the user
- "Do not ask questions… because the answer is a process." (Bowker 2010)