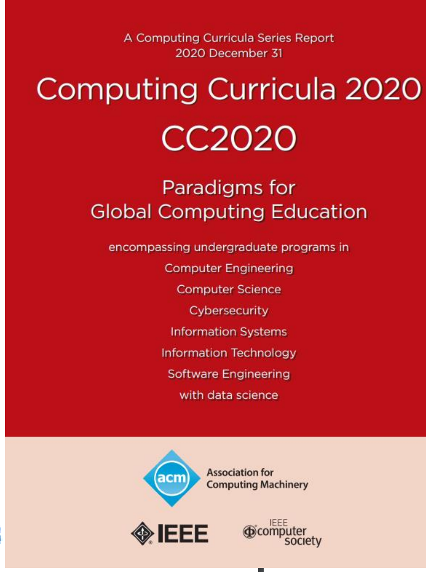
Visual Data Science @ CU

ACM CC, Skiena, Moller... Datova veda@matfyz

ACM CC 2020



1.4.1: Current Discipline Structure

The baccalaureate disciplines for which computing curricula exist or are in the development process at the time of this writing are as follows.

- Computer engineering (CE)
- Computer science (CS)
- Cybersecurity (CSEC)
- Information systems (IS)
- Information technology (IT)
- Software engineering (SE)
- Data science (DS) [1]

Each of these disciplines has a recent volume (or will soon complete a volume) sponsored by ACM and IEEE-CS for undergraduate curriculum guidelines that one or more international professional and scientific societies have endorsed and published. These disciplines have affected a large majority of undergraduate students worldwide who are majoring in computing.

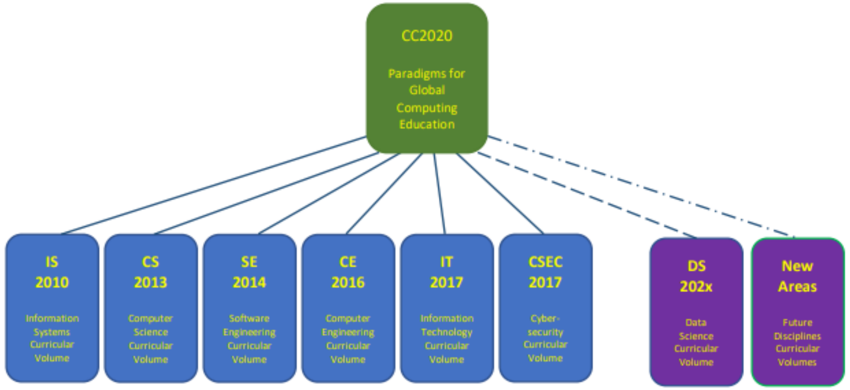


Figure 1.2 Structure of the Computing Curricula Series

JOURNAL OF STATISTICS AND DATA SCIENCE EDUCATION
2021, VOL. 29, NO. S1, S40–S50
https://doi.org/10.1080/10691898.2020.1851159

Taylor & Francis
Taylor & Francis Group

OPEN ACCESS
Check for updates

Data Science in 2020: Computing, Curricula, and Challenges for the Next 10 Years

Aimee Schwab-McCoy[a], Catherine M. Baker[b], and Rebecca E. Gasper[a]

[a]Department of Mathematics, Creighton University, Omaha, NE; [b]Department of Computer Science, Design and Journalism, Creighton University, Omaha, NE

Table 2. Topics most often taught in introductory data science courses for all disciplines.

Category	Description	Frequency of inclusion
Data visualization	Exploratory data analysis: multivariate plots and graphs, summary statistics.	56
Data cleaning	Data cleaning and wrangling (i.e., restructuring data tables, creating new variables).	51
Professional practices	Data ethics and responsible data use.	43
Data management	Data curation and data quality.	36
Statistical methods	Regression models: simple linear regression, least squares regression, logistic regression.	36
Professional practices	Reproducible research: computational reproducibility, empirical reproducibility, statistical reproducibility.	35
Professional practices	Data lifecycle and data collection, data quality evaluation.	34
Statistical methods	Research methods: research cycle, hypothesis definition and testing.	28
Data architecture	Data architecture, data types, and data formats.	27
Machine learning	Text mining: natural language processing, topic modeling, text visualization (i.e., word clouds, frequency plots).	27
Visualization	Customizing data visualizations: color and composition, accessibility, "grammar of graphics."	27
Machine learning	Supervised learning: decision trees, Naive Bayes classification, neural networks, support vector machines (SVM), ensemble methods.	26

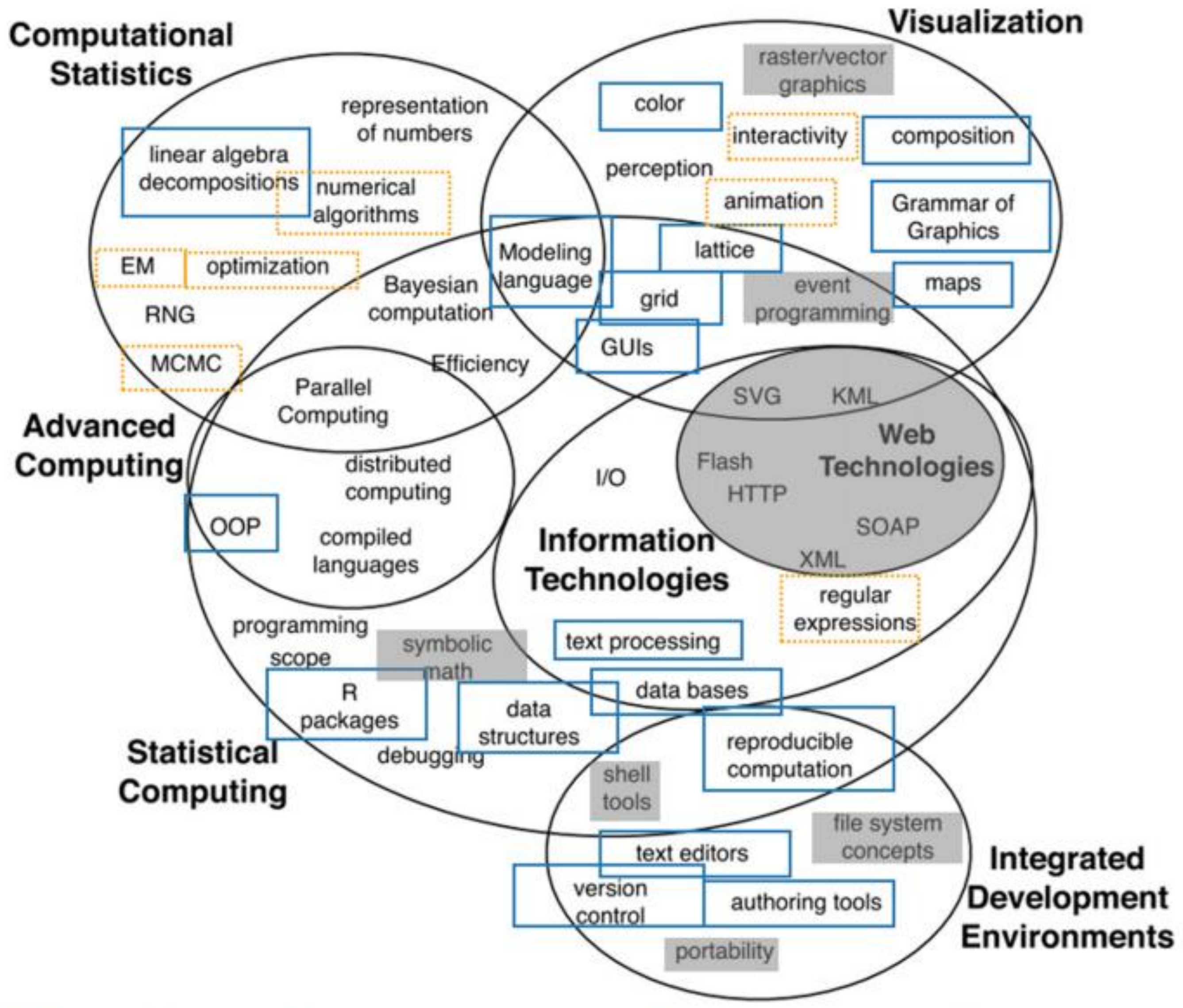


Figure 6. Computational topics relevant to statistics as indicated by Nolan and Temple Lang (2010). Topics covered in more than 75% of data science curricula in our survey are boxed in blue (solid line). Topics included in more than 50% of data science curricula are boxed in orange (dotted line). Topics not addressed in our survey are greyed out.

Knowledge

Appendix A The Body of Knowledge: Computing Competencies for Data Science......41

2.3.7: Data Science (Under Development)

Data science (DS) is a new area of computing that is closely related to the fields of data analytics and data engineering. One definition of data science is "a set of fundamental principles that guide the extraction of knowledge from data … [and] involves principles, processes, and techniques for understanding phenomena via the (automated) analysis of data." [Pro1]

Several DS projects have emerged in recent years. These include the *EDISON Data Science Framework (2017)* project [Edi1], the *National Academies Report on Data Science for Undergraduates (2018)* [Nas1], the *Park City Report (2017)* [Par1], the *Business Higher Education Framework (BHEF) Data Science and Analytics (DSA) Competency Map (2016)* [Bhe1], and the *Business Analytics Curriculum for Undergraduate Majors (2015)* [Ban1]. ACM conducted initial DS workshops in 2015; a report described the discussions, reflected the diversity of opinions, and proposed a list of knowledge areas useful for the field [Cas1]. In August 2017, the ACM Education Council created a task force to articulate the role of computing in the DS field [Dat1]. The task force produced an initial draft report tentatively tagged as (DS202x) in February of 2019 [Dat2] followed by a second draft report in December of 2019 [Dat3].

The second draft describes a "competency framework" that addresses knowledge areas representing a body of material for data science degree programs that capture high-level competencies, skills, and dispositions. The knowledge areas include (a) computing fundamentals, (b) data acquirement and governance, (c) data management, storage, and retrieval, (d) data privacy, security, and integrity, (e) machine learning, (f) data mining, (g) big data, (h) analysis and presentation, and (i) professionalism. For a full curriculum, these areas need augmentation with courses covering calculus, discrete structures, probability theory, elementary statistics, advanced topics in statistics, and linear algebra.

https://dstf.acm.org/DSReportDraft2Full.pdf

- **DSTF page https://dstf.acm.org/, Draft 1, slides...**

Data Acquisition, Management, and Governance (DG)

As the base of data science, data should be acquired, integrated and preprocessed. This is an important step to ensure both quantity and quality of data and improve the effectiveness of the following steps of data processing. Thus, a data scientist must understand concepts and approaches of data acquisition and governance including data shaping, information extraction, information integration, data reduction and compression, data transformation as well as data cleaning. In our ever increasing reliance on the quantity and quality of data in all forms of decision making, the data scientist has an ethical responsibility of protecting the integrity of data and proper use of data.

Scope	Competencies
• Shaping data and their relationships • Acquiring data from physical world and extracting data to a form suitable for analysis • Traditional Data Integration Methods: Pattern Mapping, Data Matching, Entity Recognition • Integrating heterogeneous data sources • Preprocessing and cleaning data for applications • Improving data quality • Ensuring data integrity including privacy and security	• Construct and tune the governance process according to the requirements of applications, including data preparation algorithms and steps. (Process Construction and Tuning) • Define and write semantics rules for data governance, including information extraction, data integration and data cleaning (Rules Definition) • Develop scalable and efficient algorithms for data governance according to the property of data and the requirements of applications, including information extraction, data integration, data sampling, data reduction, data compression, data transformation and data cleaning algorithm (Algorithm Development) • Describe and discover the static and dynamic properties of data, changing mechanisms of data and similarity between data. (Property Description and Discovery) • Develop policies and processes to ensure the privacy and security of data.
Sub-domains	
DG-Data Acquisition – T1, T2 DG-Information Extraction – T1, T2 DG-Working with Various Types of Data – T2 DG-Data Integration – T1	DG-Data Reduction and Compression – T1, T2 DG-Data Transformation – T1 DG-Data Cleaning – T1 DG-Data Privacy and Security – T1

DG-Data Acquisition – T1

As the initial step in data governance policies, data acquisition is the process of obtaining raw data from real-world objects. The process of data acquisition should fully consider the physical properties of the subject, and at the same time `consider the characteristics of the data application. Due to the limited resources available during data acquisition (such as network bandwidth, sensor node energy, website tokens, etc.), it is necessary to effectively design data collection techniques to maximize valuable data within limited resources and minimize valueless data. Also due to resource constraints, the data acquisition process is unlikely to obtain all the information of the data description object, so the data acquisition technology needs to be carefully designed to minimize the deviation between the collected data and the real objects.

Knowledge

- The sources of data
- Pull-based and push-based approaches
- Various data acquisition with the features of acquired data
- Data acquisition acceleration techniques
- Data discretization method
- Security and Privacy standards and best practices

Skills

T1:

- Select data source for the applications
- Design techniques for data acquisition according to the features of data sources and applications.
- Plan following steps including data discretization, transmission as well as storage.

T2:

- Design the acceleration and parallelization strategies for data acquisition according to the applications

Dispositions

- An ability to assess the trade-off between accuracy and efficiency in data acquisition.

DG-Information Extraction – T2

Information extraction (IE) is the task of automatically extracting structured information from unstructured and/or semi-structured machine-readable documents. It is an important technique to acquire data from documents, web pages, and even multimedia.

Information Extraction is relevant to the requirements of data acquisition and governance, but is described elsewhere in this report. See Information Extraction in the DataMining KA.

DG-Working with Various Types of Data – T2

Data comes in many forms. Some projects will rely completely on numeric data. Others will require processing of text or image or other media data. The data scientist must have an overview understanding of all types of data representation and processing, and must be competent to interact with some types of data as an expert.

Knowledge

- Data representation: numbers, text, images, data precision
- Text data processing: bag-of-words, word-count, TF-IDF, n-grams, Lexical analysis, syntax analysis, semantic analysis, stop word filtering, stemming, basic applications
- Image processing: data representation: multi-dimensional matrices of integers, features, image operators, video operators. Object recognition. Higher order feature extraction

Skills

- Write programs to perform basic operations on data of each type. Compute summary statistics, extract n-grams, do modifications to an image.

Dispositions

- Recognize the importance of choice of data type for encoding information.

DG-Data Integration – T1

In the data acquisition process, since the data may come from an autonomous data source, it is difficult to ensure the consistency of the data mode, modality, semantics, etc.. However, in many applications, these data from multiple autonomous data sources need to be summarized and used together to generate new value, this is the task of data integration, which is a crucial step for data acquisition and governance.

Knowledge

- The concepts and application scenarios of government database, data warehouse and mediator-based information integration
- The concepts and approaches of schema mapping
- The concepts and approaches of data mapping
- The concepts and approaches of data semantic transformation
- The techniques of cross-domain data integration

Skills

- Choose the scheme of data integration i.e. traditional data integration VS. cross-domain data integration
- Choose the architecture of data integration according to the features of applications
- Select or develop appropriate algorithms for schema mapping, data mapping and data semantic transformation
- Develop proper algorithms for cross-domain data integration

Dispositions

- Understand the challenges brought by heterogeneous data sources
- Know the roles of AI in data integration

DG-Data Reduction and Compression

The goal of data reduction and compression is to eliminate the redundancy of data and decrease the size of data involved in the next data processing steps. This involves data sampling, filtering and compression.

Knowledge

T1:

- The role of reduction and compression in data process
- Various data sampling approaches
- Data filter techniques
- Data compression techniques

Skills

T1:

- Determine whether data reduction and compression steps are required
- Perform data sampling and filtering

T2:

- Analyse the properties of data sampling
- Select data compression techniques according to the computation, communication and storage requirements
- Develop query-friendly data compression approach

Dispositions

- Understand the trade-off between data computation effectiveness and efficiency.

DG-Data Transformation – T1

Data collected from data sources often have different dimensions and ranges. These data may be correct, but they cannot be directly used. It is often necessary to transform the collected data and

https://dstf.acm.org/DSReportDraft2Full.pdf, p. 68-69 convert the data into "appropriate" form to understand the data or visualize the data to achieve effective application of the data.

Knowledge
- Data Transformation pipeline
- Simple function transformation methods and their applications
- Data standardization and its applications
- Data normalization and its applications
- Data encoding approaches and their applications
- Data smooth approaches and their applications

Skills
- Evaluate and compare the dimension and range of data and those of the requirements in the applications.
- Determine the process of data transformation
- Choose proper data algorithms for the task
- Evaluate the effectiveness of data transformation

Dispositions
- Understand the importance of data transformation to data usage
- Have an understanding of the links between data transformation and data quality

DG-Data Cleaning – T1

Data quality is an important aspect of data usability. There is a perception that if data is "suitable for its intended use in operations, decision making, and planning," it is generally considered to be of high quality. There are also views that if the data correctly represents the real-world entities that it refers to, then it is also considered to be of high quality. Data quality issues and the resulting knowledge and decision-making mistakes have had terrible consequences on a global scale. Data cleaning is an important solution for data quality problems.

Knowledge
- The dimensions of data quality
- The approaches to improve data quality
- Data cleaning algorithms including entity resolution, truth discovery, rule-based data cleaning.
- Various forms for data quality rules such as functional dependencies (FD), conditional functional dependencies (CFD), conditional inclusion dependencies (CIND), and matching dependencies (MD)

Skills
- Evaluate data quality
- Write rules for data cleaning according to the requirement of applications and data semantics
- Develop data cleaning pipeline according to the data quality requirements.
- Develop algorithms for efficient and effective data cleaning

Dispositions
- Hold an awareness of the harm of data quality problems
- Appreciation of and ability to handle the role of data cleaning in data usage.

DG-Data Privacy and Security – T1

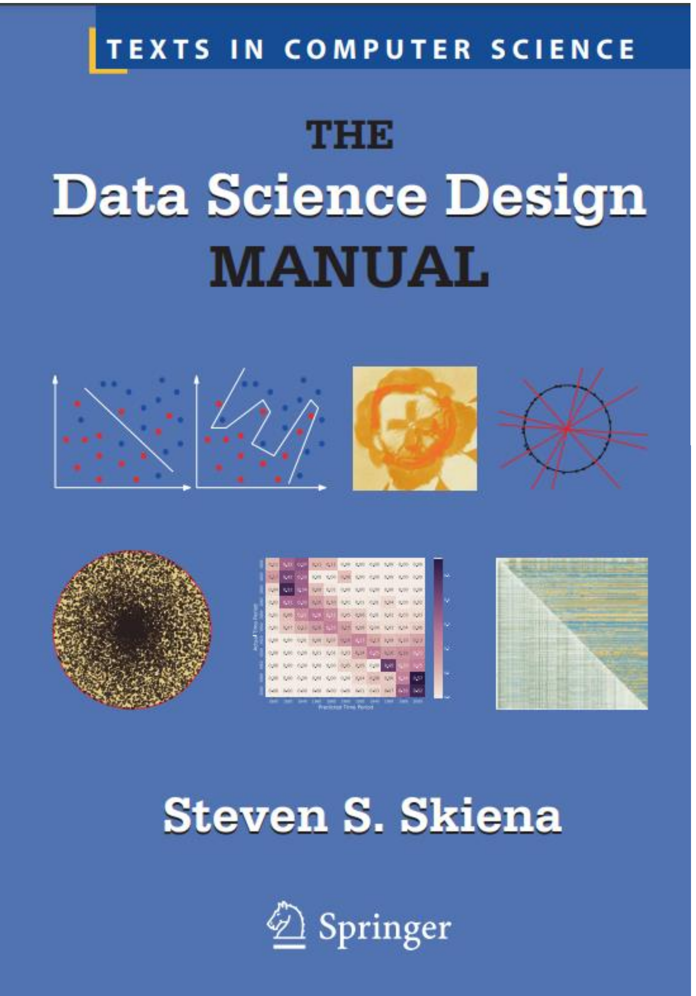
Knowledge
- The relationships between individuals, organizations, or governmental privacy requirements
- The cross-border privacy and data security laws and responsibilities
- A comprehension of how organizations with international engagement must consider variances in privacy laws, regulations, and standards across the jurisdictions in which they operate.

Skills
- Explain how laws and technology intersect in the context of the judicial structures that are present – international, national and local – as organizations safeguard information systems from cyberattacks.
- Explain requirements of the General Data Protection Regulation (GDPR). And Privacy Shield agreement between countries, such as the United States and the United Kingdom, allowing the transfer of personal data.
- Describe how [Section 5 of the U.S. Federal Trade Commission, State data security laws, State data-breach notification laws, Health Insurance Portability Accountability Act (HIPAA), Gramm Leach Bliley Act (GLBA), and Information sharing through US-CERT, Cybersecurity Act of 2015] and other laws impact data security

Dispositions
- Have an appreciation for the ethical implications of data governance policies and actions
- Hold an awareness of the harm of data loss due to security and privacy failures
- Maintain the upmost ethical standards regarding legal and social responsibility for data

Data Science Design Manual by Steven Skiena

- https://www.webpages.uidaho.edu/~stevel/517/The%20Data%20Science%20Design%20Manual.pdf



Chapter 1

What is Data Science?

The purpose of computing is insight, not numbers.

– Richard W. Hamming

- Visual Computer, T.L.Kunii
  - Visual Computing

Questions >> Images .. Abela 2013

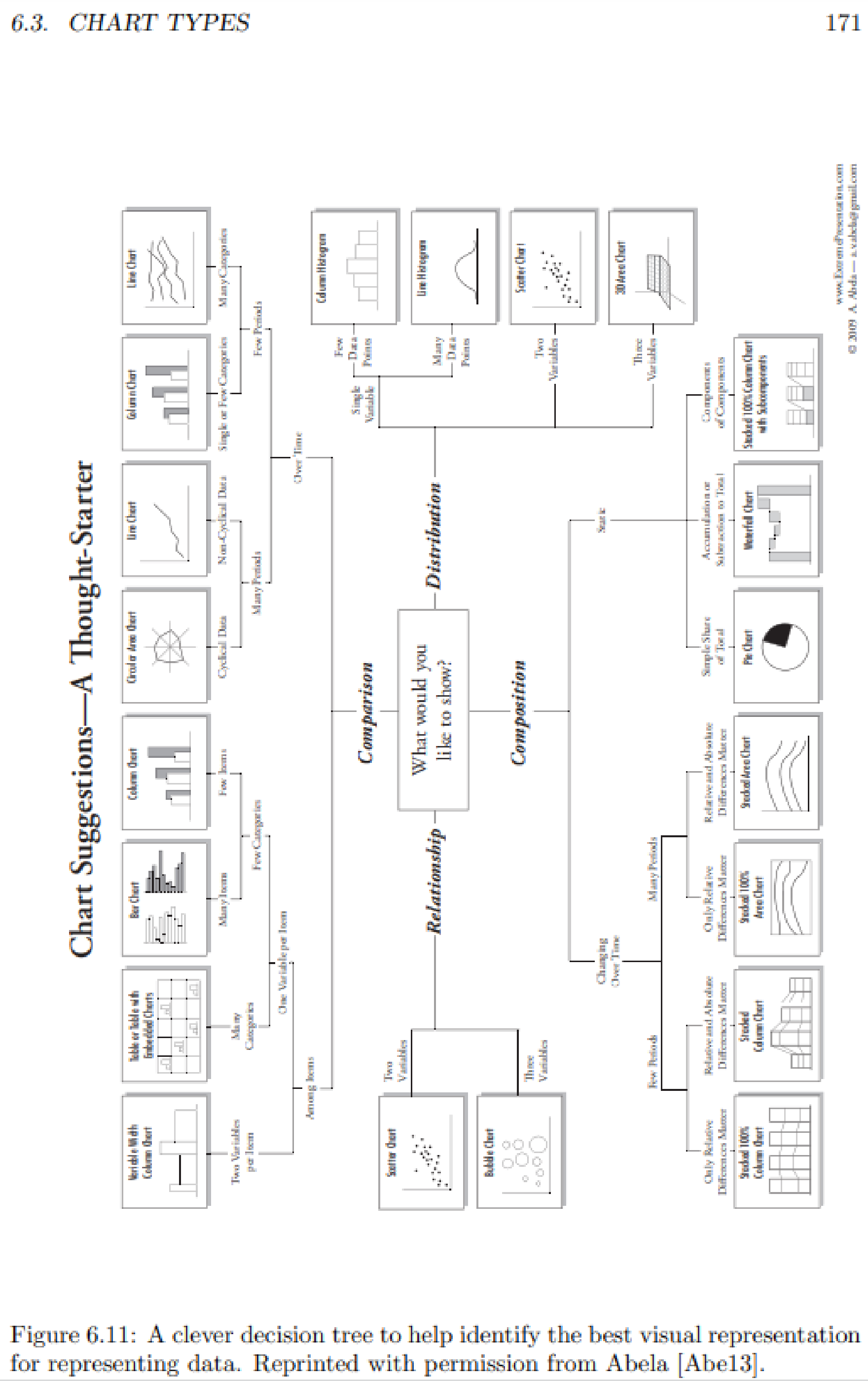
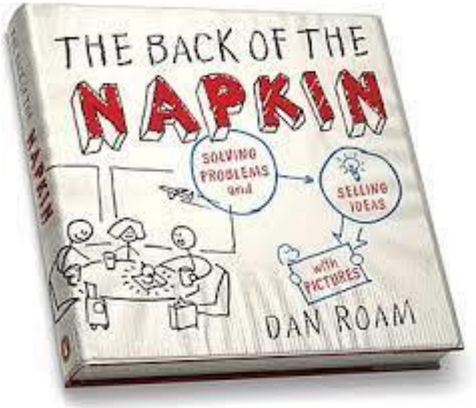

6.3. CHART TYPES 171

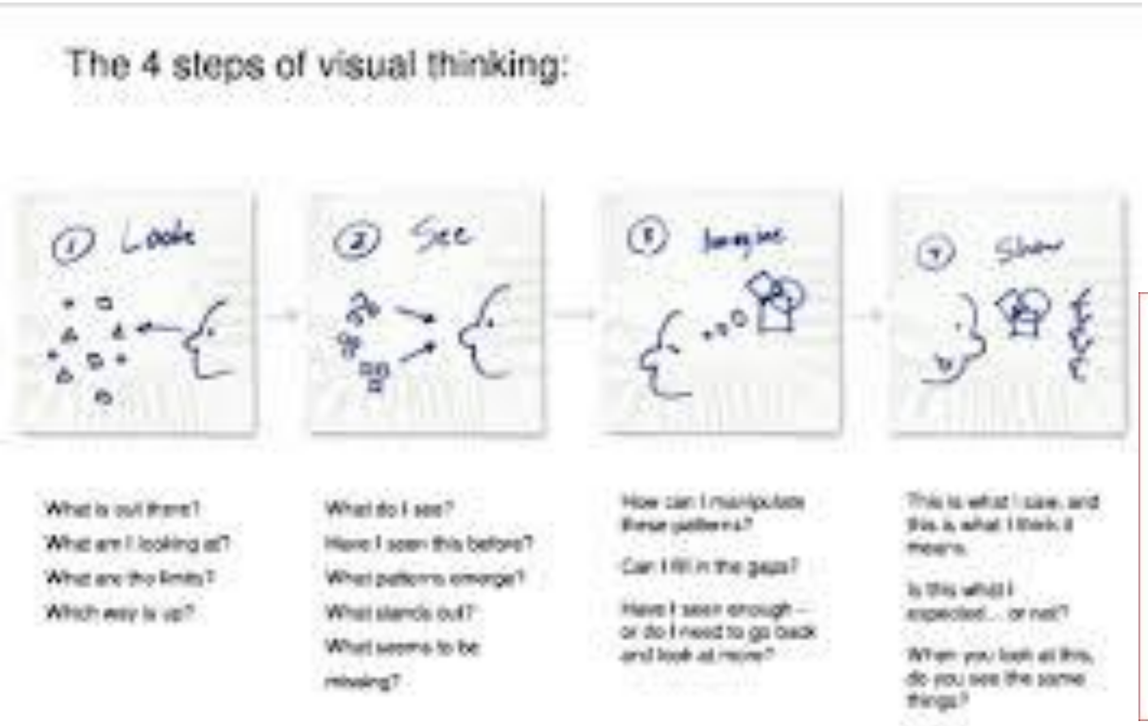
Figure 6.11: A clever decision tree to help identify the best visual representation for representing data. Reprinted with permission from Abela [Abe13].

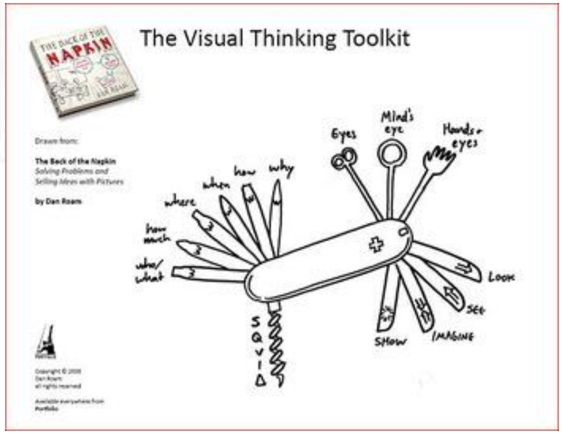
- 100 * “interesting”, not defined

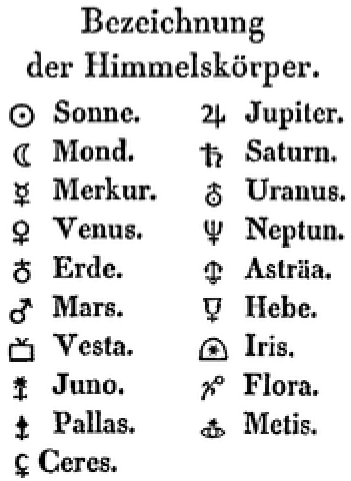
Rosetta Stone, Visual Thinking Codex



- Symbolic >> visual, e.g. 6W, ascii, svg, vrml, x3d... nD data
- Non-meaning >> meaning, e.g. Rosetta: Greece >> hieroglyph

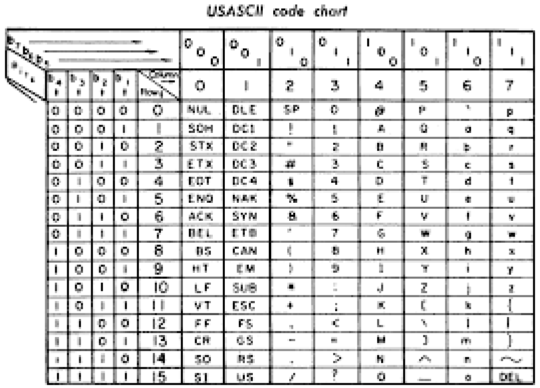






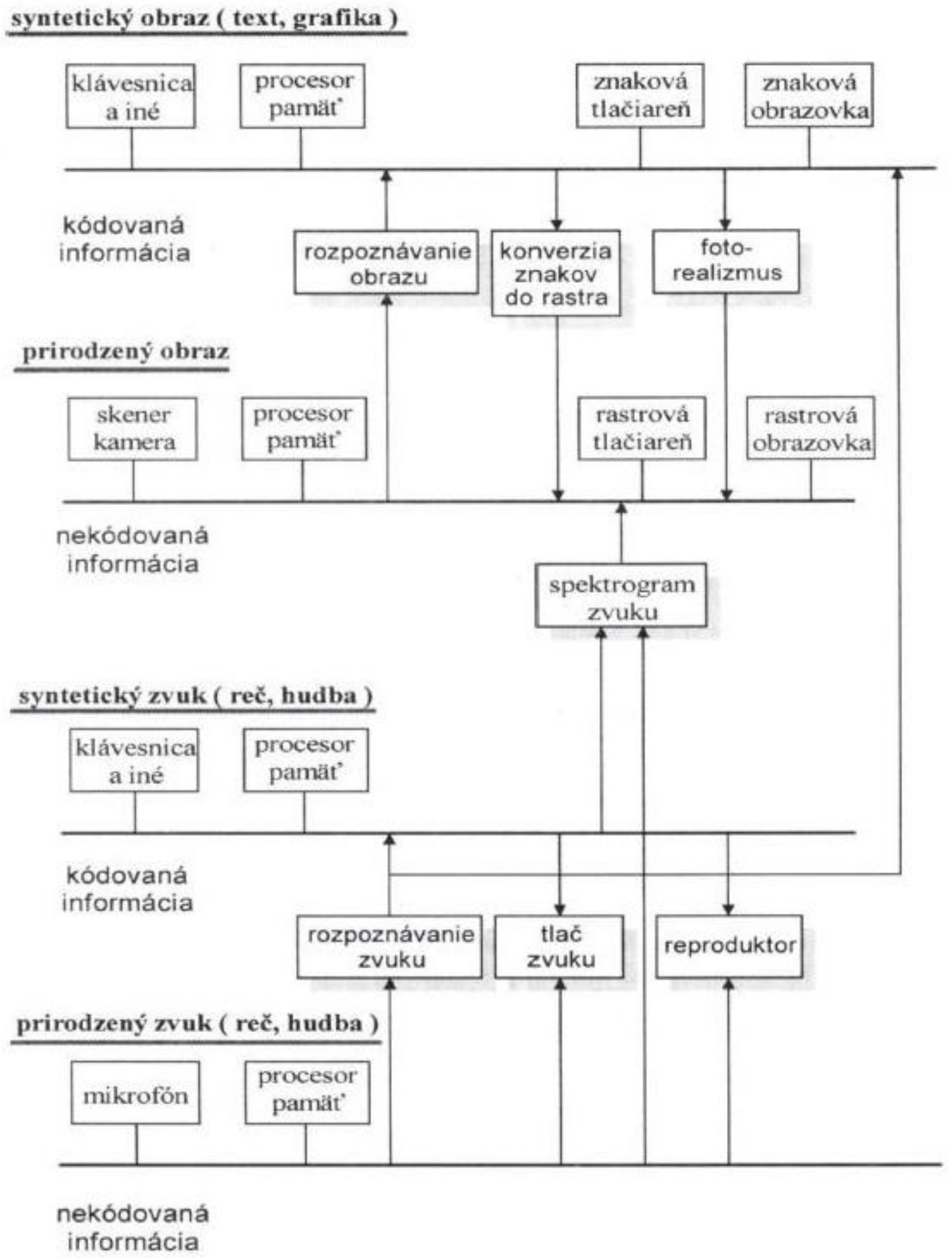
Bezeichnung
der Himmelskörper.

☉ Sonne.	♃ Jupiter.
☾ Mond.	♄ Saturn.
☿ Merkur.	⛢ Uranus.
♀ Venus.	♆ Neptun.
♁ Erde.	Asträa.
♂ Mars.	Hebe.
Vesta.	Iris.
Juno.	Flora.
Pallas.	Metis.
Ceres.	

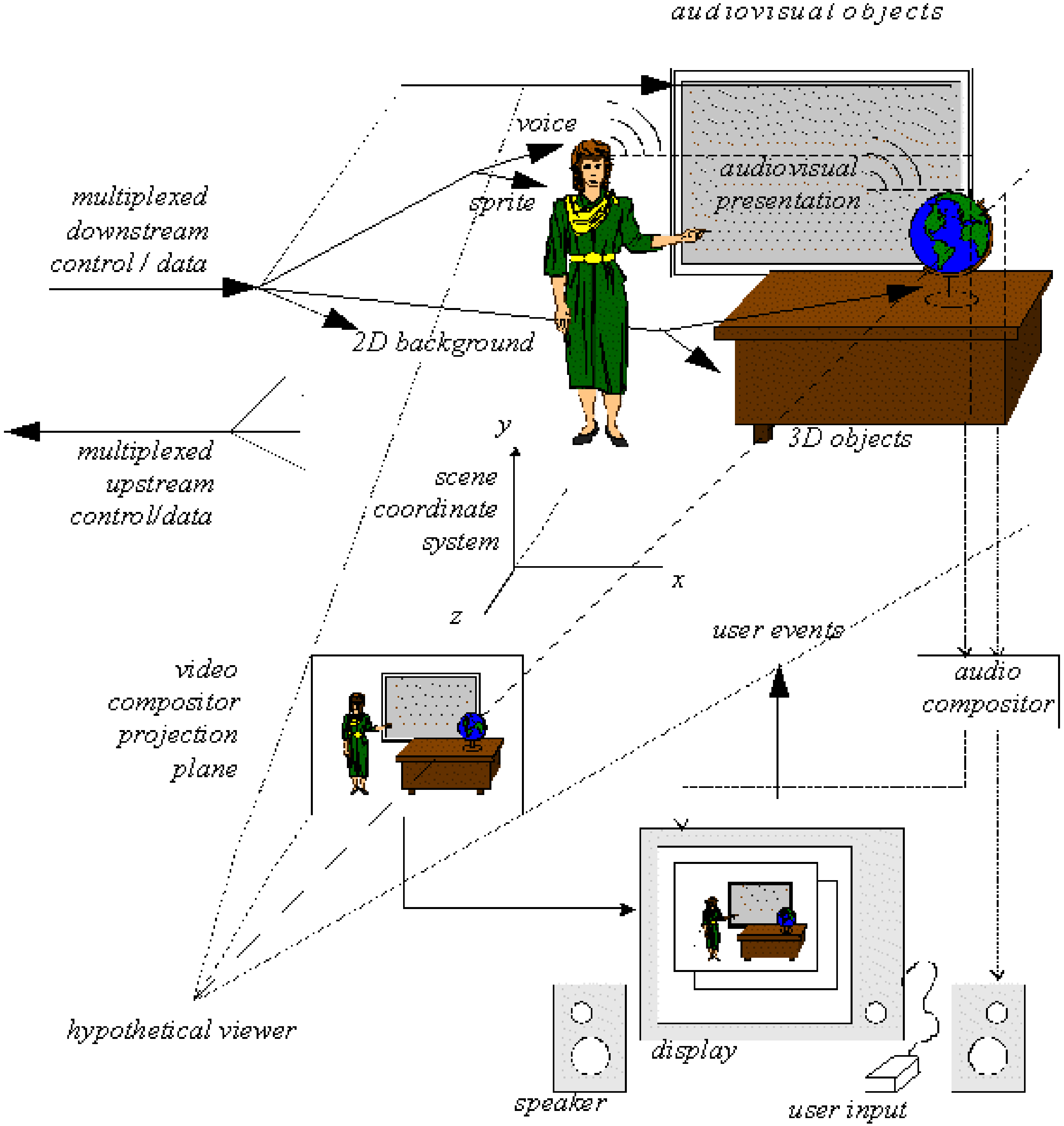


USASCII code chart

b7 b6 b5					000	001	010	011	100	101	110	111
b4	b3	b2	b1	Row \ Column	0	1	2	3	4	5	6	7
0	0	0	0	0	NUL	DLE	SP	0	@	P	`	p
0	0	0	1	1	SOH	DC1	!	1	A	Q	a	q
0	0	1	0	2	STX	DC2	"	2	B	R	b	r
0	0	1	1	3	ETX	DC3	#	3	C	S	c	s
0	1	0	0	4	EOT	DC4	$	4	D	T	d	t
0	1	0	1	5	ENQ	NAK	%	5	E	U	e	u
0	1	1	0	6	ACK	SYN	&	6	F	V	f	v
0	1	1	1	7	BEL	ETB	'	7	G	W	g	w
1	0	0	0	8	BS	CAN	(	8	H	X	h	x
1	0	0	1	9	HT	EM	)	9	I	Y	i	y
1	0	1	0	10	LF	SUB	*	:	J	Z	j	z
1	0	1	1	11	VT	ESC	+	;	K	[	k	{
1	1	0	0	12	FF	FS	,	<	L	\	l	|
1	1	0	1	13	CR	GS	-	=	M	]	m	}
1	1	1	0	14	SO	RS	.	>	N	^	n	~
1	1	1	1	15	SI	US	/	?	O	_	o	DEL



Obr. 20.1 Architektúra multimediálneho systému



Ruzicky, PREMO, CGRM

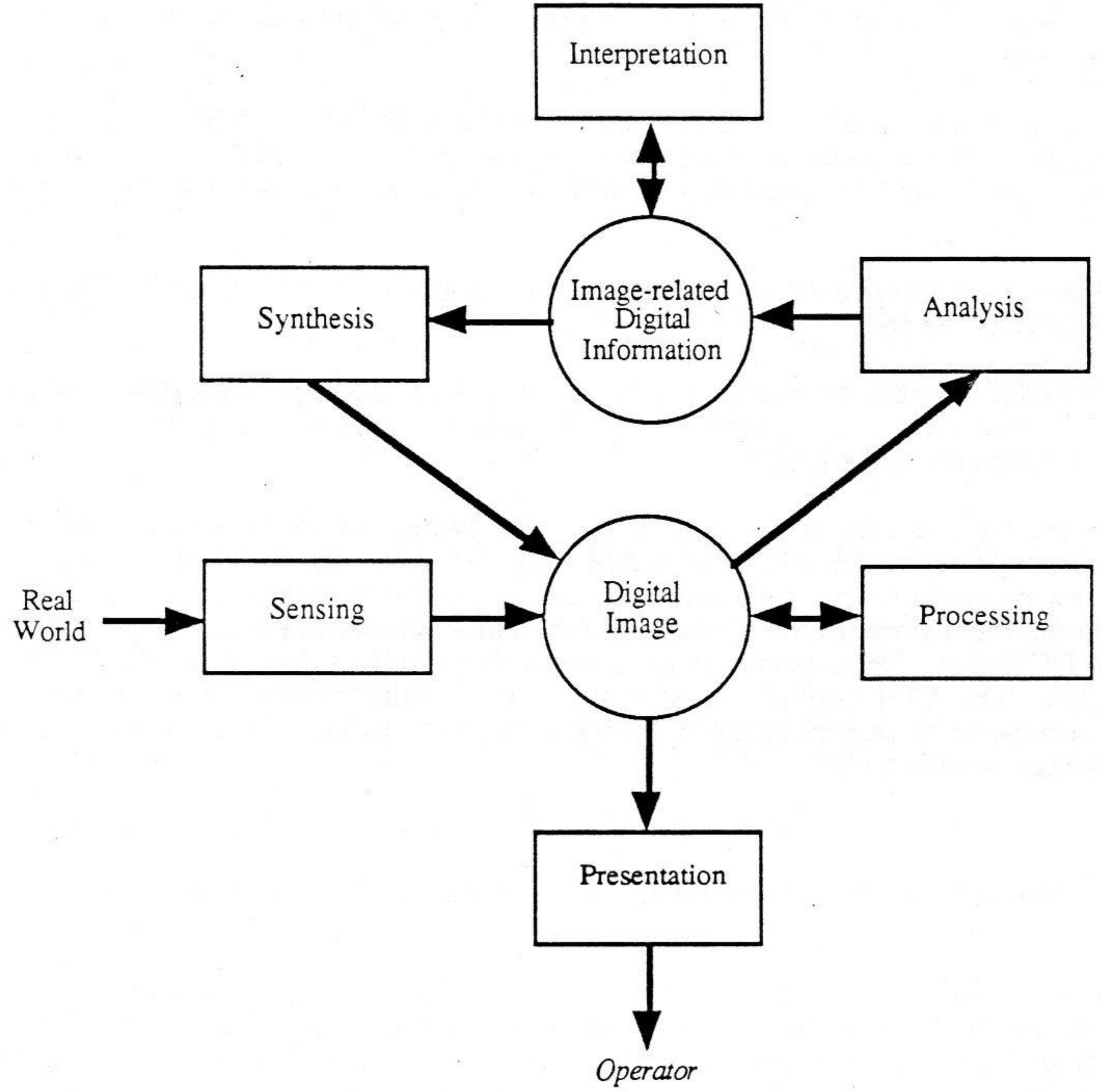
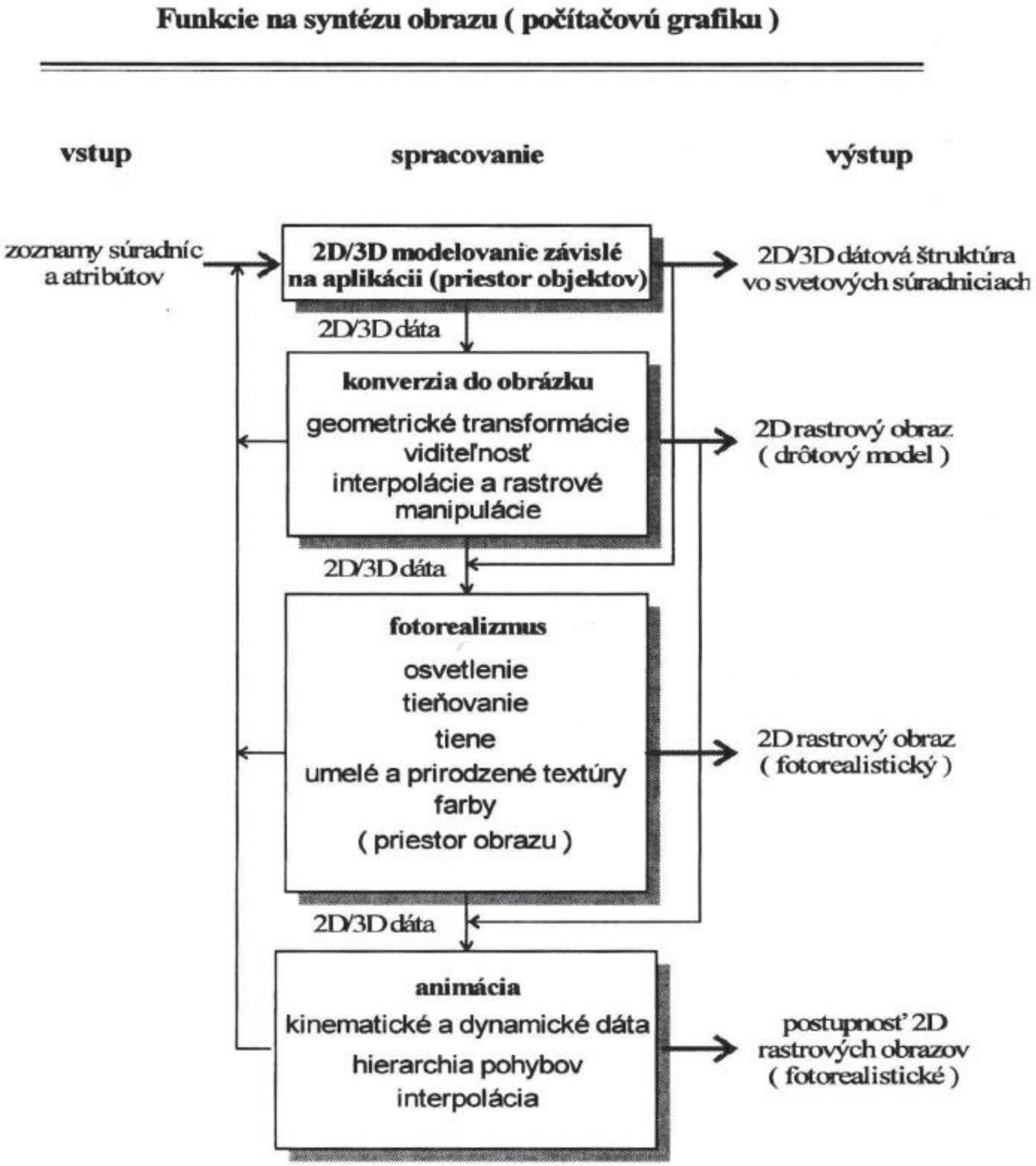


Figure B.1 – Computer imaging model



Obr. 1.7 Funkcie na vytváranie obrazu (pasívna počítačová grafika)

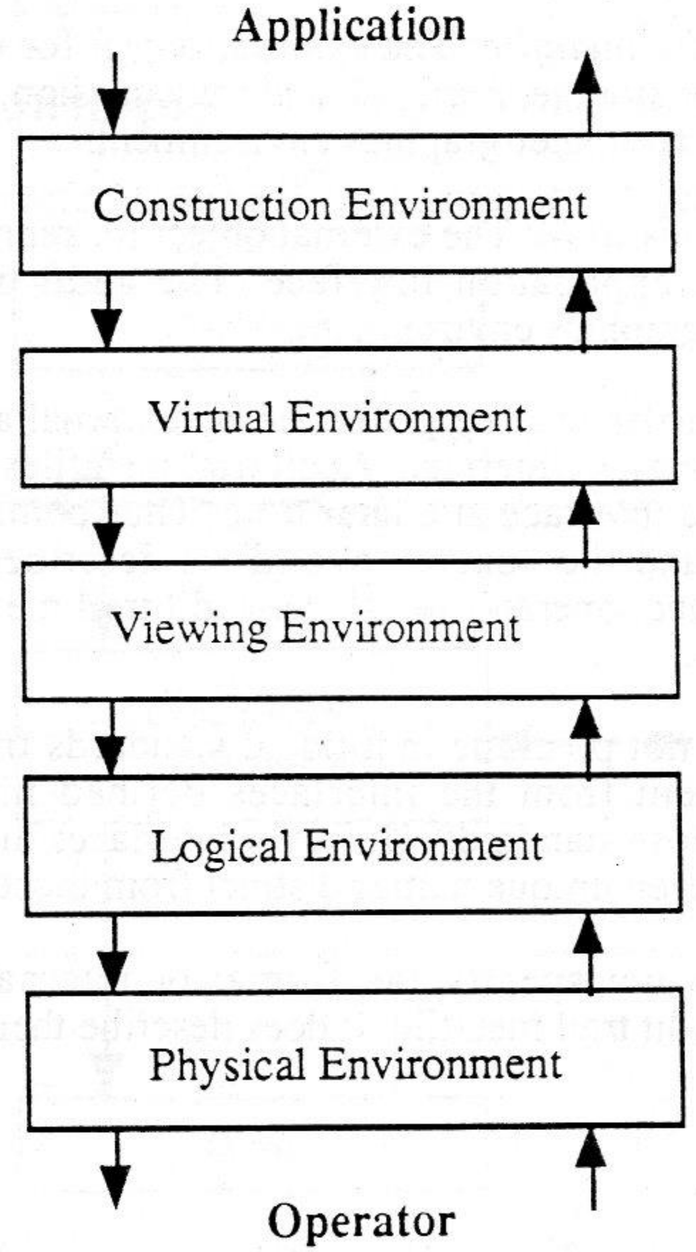


Figure 2 – Computer graphics environments

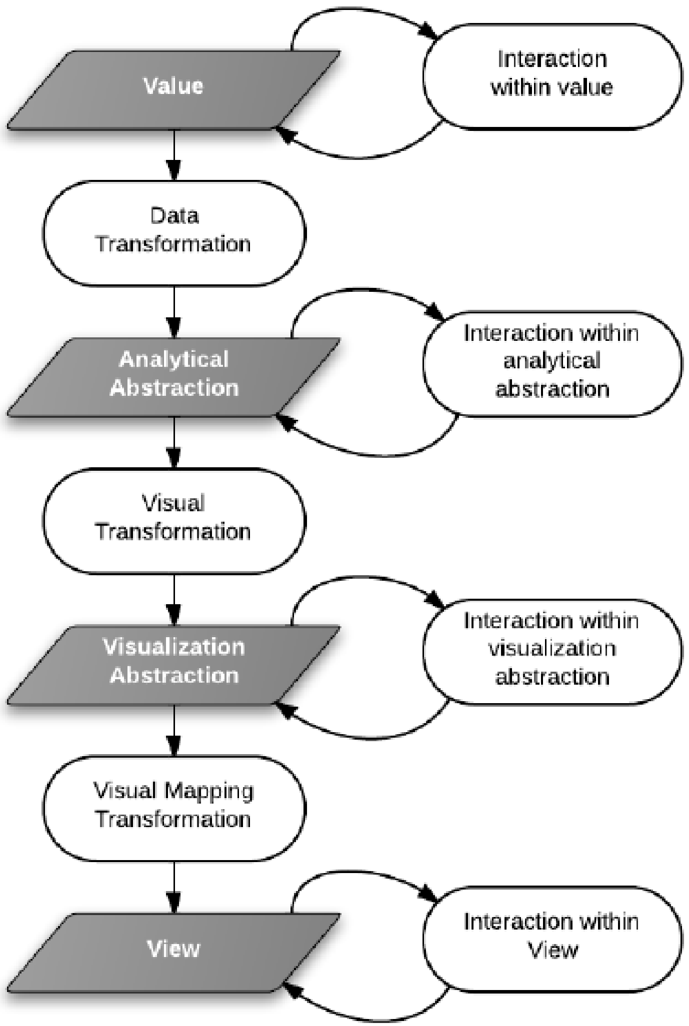
Four data buses metaphor

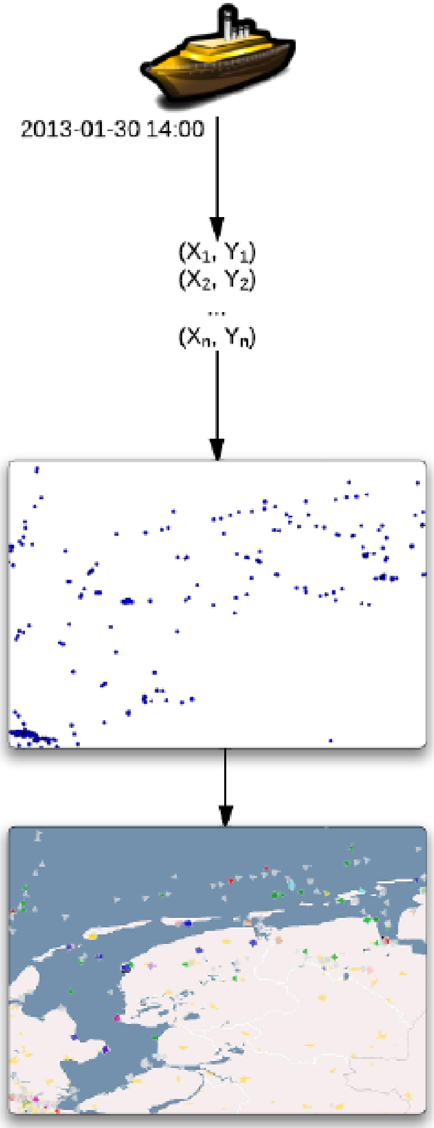
- **A. Data bus for structured pictures,** TXT, SVG, CGM, VRML…
- **B. Data bus for unstructured images,** JPG, BMP, PNG, TIFF…
- **C. Data bus for structured sounds**
- **D. Data bus for unstructured, natural sounds**

- **Input for A & C: model, data and functionality**
- **Input for B & D: scanner and microphone**

- **[Stuc91] STUCKI, P.: Graphics and Multimedia, tutorial at Eurographics Conference, Vienna 1991, cited in Ruzicky**

Data State Model, InfoVis Taxonomy [Chi2000]





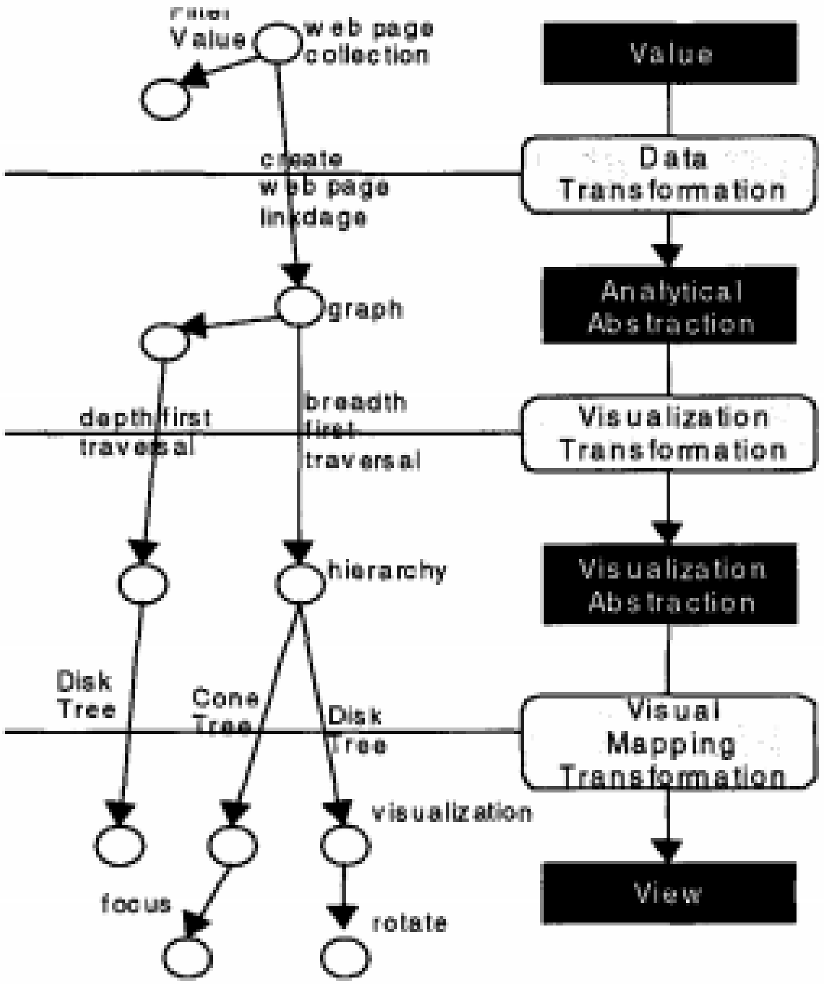


Figure 2: Data State Model applied to Web sites

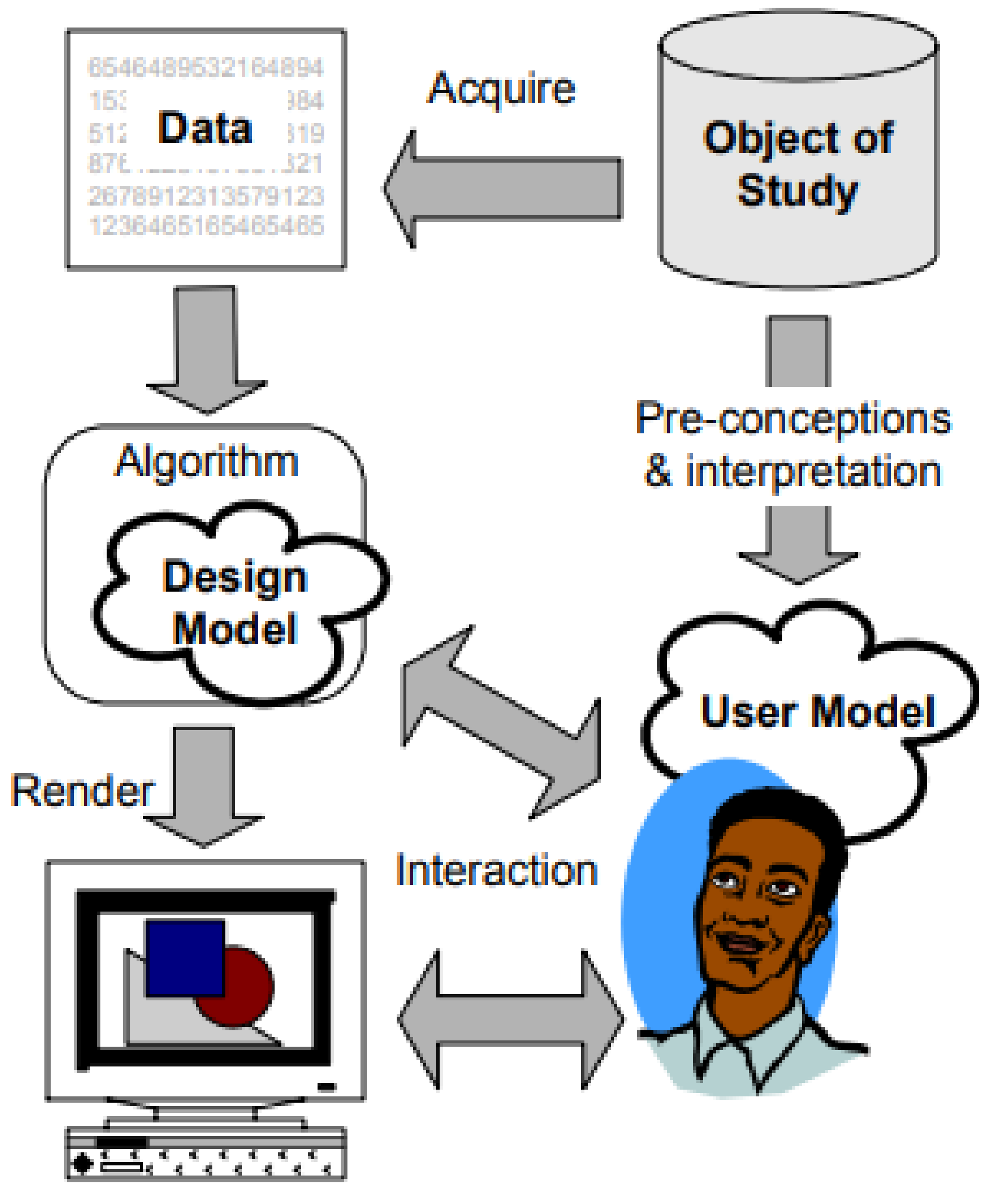


Figure 1: Relationships between object of study, data, visualization algorithm, design model, and user model.

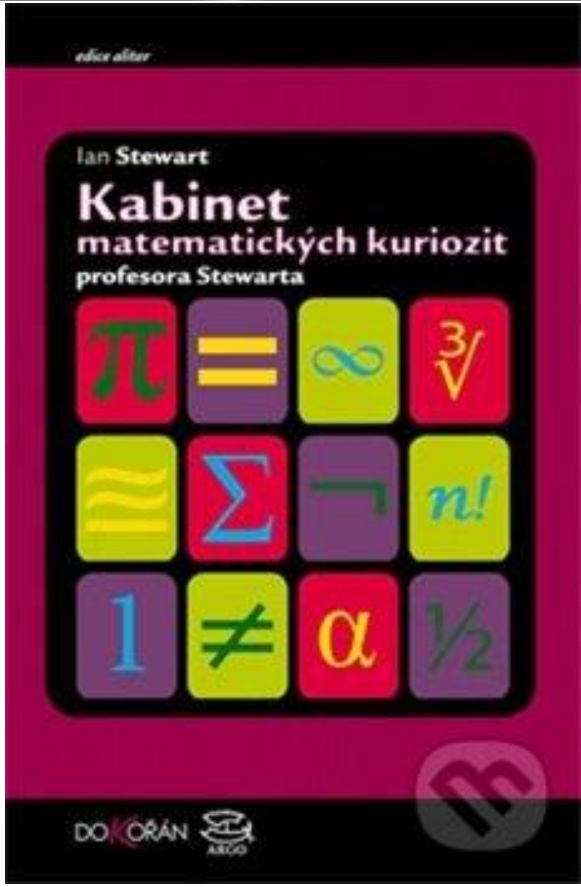
Table 1: High-level visualization taxonomy, illustrated by examples. Design models are classified based on whether they are discrete or continuous and by how much the algorithm designer chooses display attributes (spatialization, timing, colour, and transparency). Examples show different constraints on spatialization.

	Display Attributes		
	Given	***Constrained***	***Chosen***
Continuous	Images (e.g., medical) Fluid / gas flow, pressure distributions Molecular structures (distributions of mass, charge, etc.) Globe – distribution data (e.g., elevation levels)	Distortions of given / continuous ideas (e.g., flattened medical structures, 2D geographic maps, fish-eye lens views) Arrangement of numeric variable values	Continuous (high-dimensional) mathematical functions Continuous time-varying data, when time is mapped to a spatial dimension Regression analyses
Discrete	Classified data / images (e.g., segmented medical images) Air traffic positions Molecular structures (exact positions of components) Globe – discrete entity data (e.g., city locations)	Distortions of given / discrete ideas (e.g., 2D geographic maps, fish-eye lens views) Arrangement of ordinal or numeric variable values	Discrete time-varying data, when time is mapped to a spatial dimension Arbitrary entity-relationship data (e.g., file structures) Arbitrary multi-dimensional data (e.g., employment statistics)

- Rethinking visualization: A high-level taxonomy
    - Melanie Tory, Torsten Möller, 2004

Museum, prehistory

- Collection, cabinet, museum, encyclopaedia
- Codex Gigas, al-Mukaddima, Orbis pictus
- https://cs.wikipedia.org/wiki/Kabinet_kuriozit
- Professor Stewart's Cabinet of Mathematical Curiosities
- ICOM, 55,000 museums in 202 countries



the oldest label

Visual Data Science: Advancing Science through visual reasoning

Torsten Möller
Visualization and Data Analysis
University of Vienna

Vasant Dhar, "Data Science and Prediction", (2013)

What is data science?

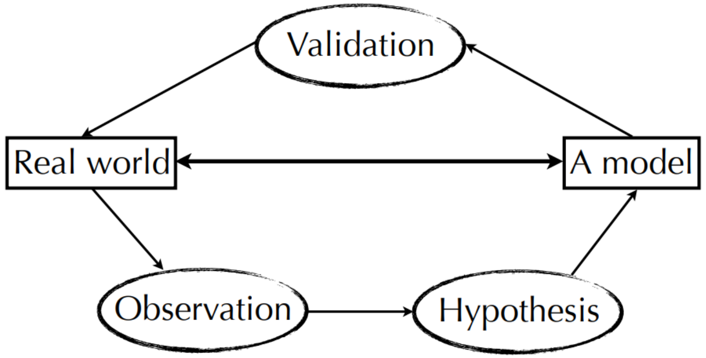
- Dhar 2013: "Data Science is the study of the generalizable extraction of knowledge from data."

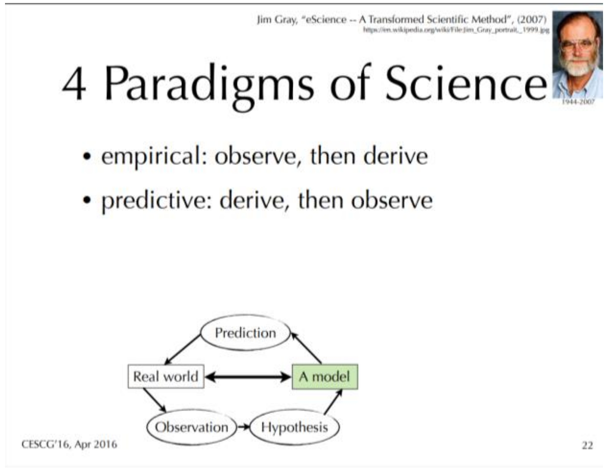
Overview

- Data Science is all about modelling
- The three types of modelling
  - Computational modelling
  - Statistical modelling
  - Empirical modelling
- Challenges of Visual Data Science
- Conclusions

after Hans Christian Ørsted, "First Introduction to General Physics" (1811

Scientific Method



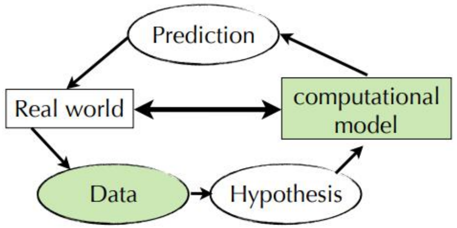


Jim Gray, "eScience -- A Transformed Scientific Method", (2007)
https://en.wikipedia.org/wiki/File:Jim_Gray_portrait,_1999.jpg

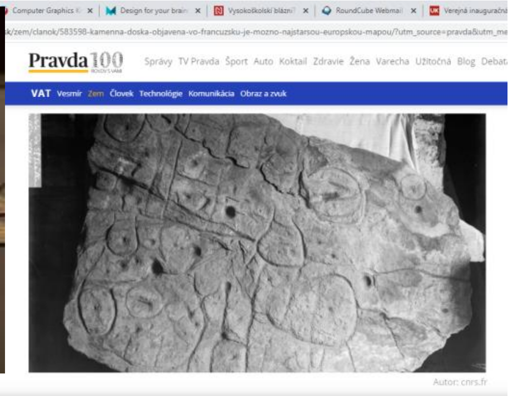
1944-2007

4 Paradigms of Science

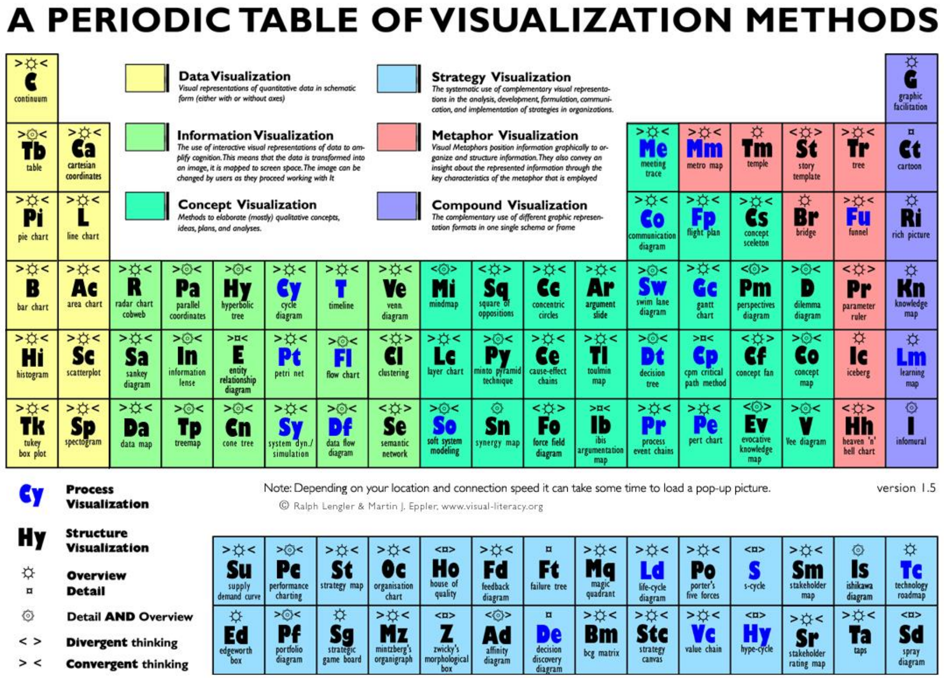
- empirical: observe, then derive
- predictive: derive, then observe
- computational: simulate
- data-driven: measure

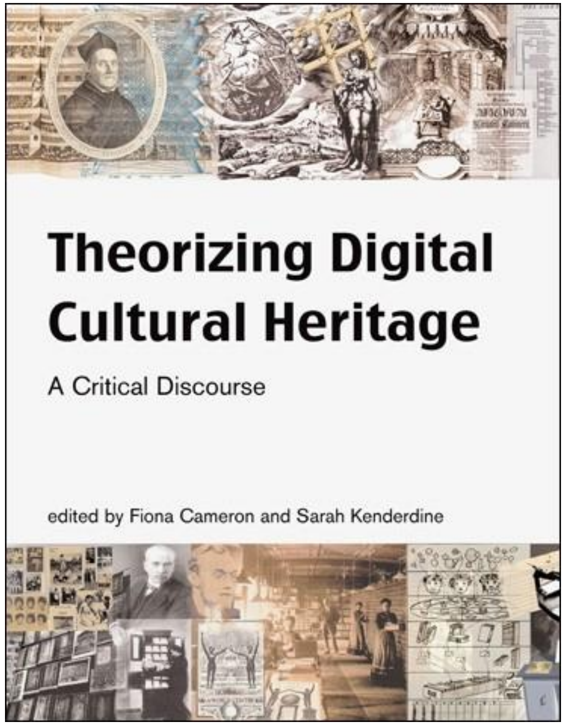


Virtual Heritage



- Virtual Heritage (VH) can be defined as “the use of computer-based interactive technologies to record, preserve, or recreate artefacts, sites and actors of historic, artistic and cultural significance, and to deliver the results openly to a global audience in such a way as to provide formative educational experiences through electronic manipulations of time and space” [Stone, R.J. “Virtual Heritage: “The willing suspension of disbelief for the moment…””; UNESCO World Heritage Review; October, 1999; pp.18-27.]

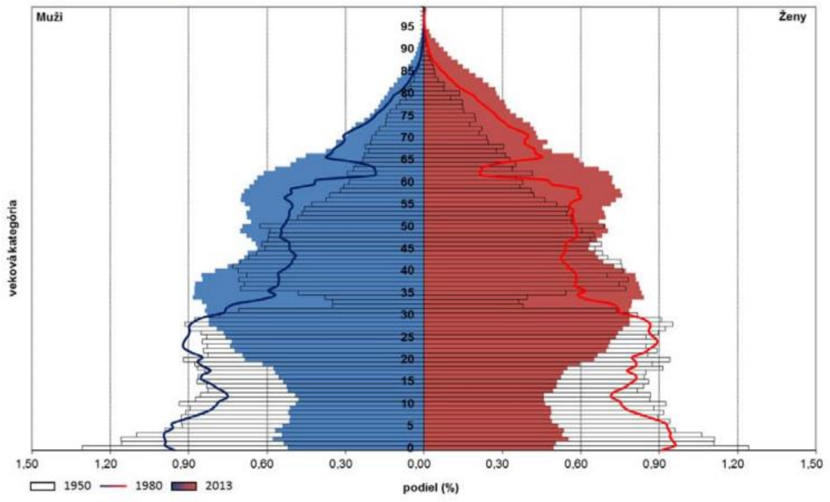




2. STARNUTIE OBYVATEĽSTVA SLOVENSKA

Predlžujúca sa stredná dĺžka života a znižujúci sa počet živonarodených detí sú dva súčasne sa uplatňujúce demografické princípy vo väčšine populácií, ktoré smerujú k procesu populačného starnutia. Vývoj vekovej štruktúry obyvateľstva Slovenska v rokoch 1950 – 2013 (graf č. 1) tento trend potvrdzuje.

Graf č. 1: Veková pyramída obyvateľstva SR (1950, 1980, 2013)



Zdroj údajov: [14]

Virtual museum: VMC, Smithsonian 9/11

- A **virtual museum** is a digital entity that draws on the characteristics of a museum, in order to complement, enhance, or augment the museum experience through personalization, interactivity and richness of content. Virtual museums can perform as the digital footprint of a physical museum, or can act independently, while maintaining the authoritative status as bestowed by the International Council of Museums (ICOM) in its definition of a museum. In tandem with the ICOM mission of a physical museum, the virtual museum is also committed to public access; to both the knowledge systems imbedded in the collections and the systematic, and coherent organization of their display, as well as to their long-term preservation.
- Definition of a Virtual Museum by Forte in Qvortrup et al. - adding **"telematic collection of multimedia..."** // DECLARATIVE
- Things, people, environments * Visualization, activising, hermeneutic sites // 9 project options
- Real time – one past, virtual time – 2 pasts (author, user)

Prehľad národných projektov

	Názov projektu	Žiadateľ	Rozpočet v €
1.	Digitálna knižnica a digitálny archív	Slovenská národná knižnica	49 572 033, 60
2	Digitálna galéria	Slovenská národná galéria	15 457 026, 36
3.	Digitálne múzeum	Múzeum Slovenského národného povstania	27 576 538, 32
4.	Digitálny pamiatkový fond	Pamiatkový úrad SR	10 273 680, 00
5.	Digitálna audiovízia	Slovenský filmový ústav	24 089 940, 37
6.	Centrálna aplikačná infraštruktúra a registratúra	Národné osvetové centrum	23 072 112, 16
7.	Centrálny dátový archív	Univerzitná knižnica Bratislava	33 184 470, 48
8.	Harmonizácia informačných systémov	Národné osvetové centrum	7 061 410,40
9.	Dokumentačno-informačné centrum rómskej kultúry	Štátna vedecká knižnica Prešov	1 913 406,46
10.	Digitalizácia kultúrneho dedičstva rezortných a mimorezortných pamäťových a fondových inštitúcií (dopytovo orientovaná výzva)	Múzeá, galérie, vedecké, akademické inštitúcie	20 222 346

Špecifické ciele digitalizácie

Digitalizácia kultúrneho dedičstva znamená pre Slovenskú republiku jedinečnú a historickú príležitosť na sprostredkovanie jej kultúrneho bohatstva pre budúce generácie. Cieľom digitalizácie je zároveň zjednodušiť a zefektívniť prístup ku kultúrnym objektom a sprostredkovať ich laickej a odbornej verejnosti. Napriek tomu, že má dosť neprajníkov (aj v radoch pracovníkov kultúry, múzeí atď.), nie je neja-

Žabková, S. 2013. Digitalizácia kultúrneho dedičstva. Projekt: Digitálne múzeum https://www.muzeologia.sk/index_htm_files/MuzeologiaKD_1_2013_Zabkova.pdf


Digitalizačné centrum v Banskej Bystrici

Konkrétne ciele projektu

a) Zabezpečiť vizualizáciu najvýznamnejších hmotných a nehmotných dokladov vývoja prírody a spoločnosti na základe obsahového rámca a v zmysle lokálneho a časového harmonogramu

b) Podporiť jednoznačnú identifikáciu digitálnych objektov, ich bezpečné uloženie a integráciu s vedomostným systémom

c) Spracovať digitálne objekty do podoby vizuálnych zástupcov hmotných dokladov a vytvoriť tak podmienky na ich plnohodnotné on-line využitie v oblasti vedy, prezentácie a reprezentácie kultúry

d) Zdokonaľovať technológie a metódy digitálnej vizualizácie kultúrnych objektov a uplatňovať ich v prebiehajúcich procesoch vedomostného zhodnocovania múzejných zbierok a uplatňovania v oblasti výchovy verejnosti a prezentácie krajiny

e) Umožniť rýchly a selektívny prístup ku kultúrnym objektom prostredníctvom ich vizuálnych zástupcov za účelom:
- správy a kontroly kultúrneho dedičstva
- ďalšieho vedomostného zhodnocovania zbierok pamäťových a fondových inštitúcií
- využitia kultúrneho dedičstva na vzdelávacie a výchovné účely
- uplatnenia kultúrneho dedičstva v oblasti prezentácie, reprezentácie a propagácie kultúry
- využitia vizuálnych zástupcov kultúrnych objektov v kultúrnom priemysle

f) pokračovať vo vizualizácii kultúrneho dedičstva – digitalizáciu ostatných zbierok a dokumentov vzťahujúcich sa k vizualizovaným objektom

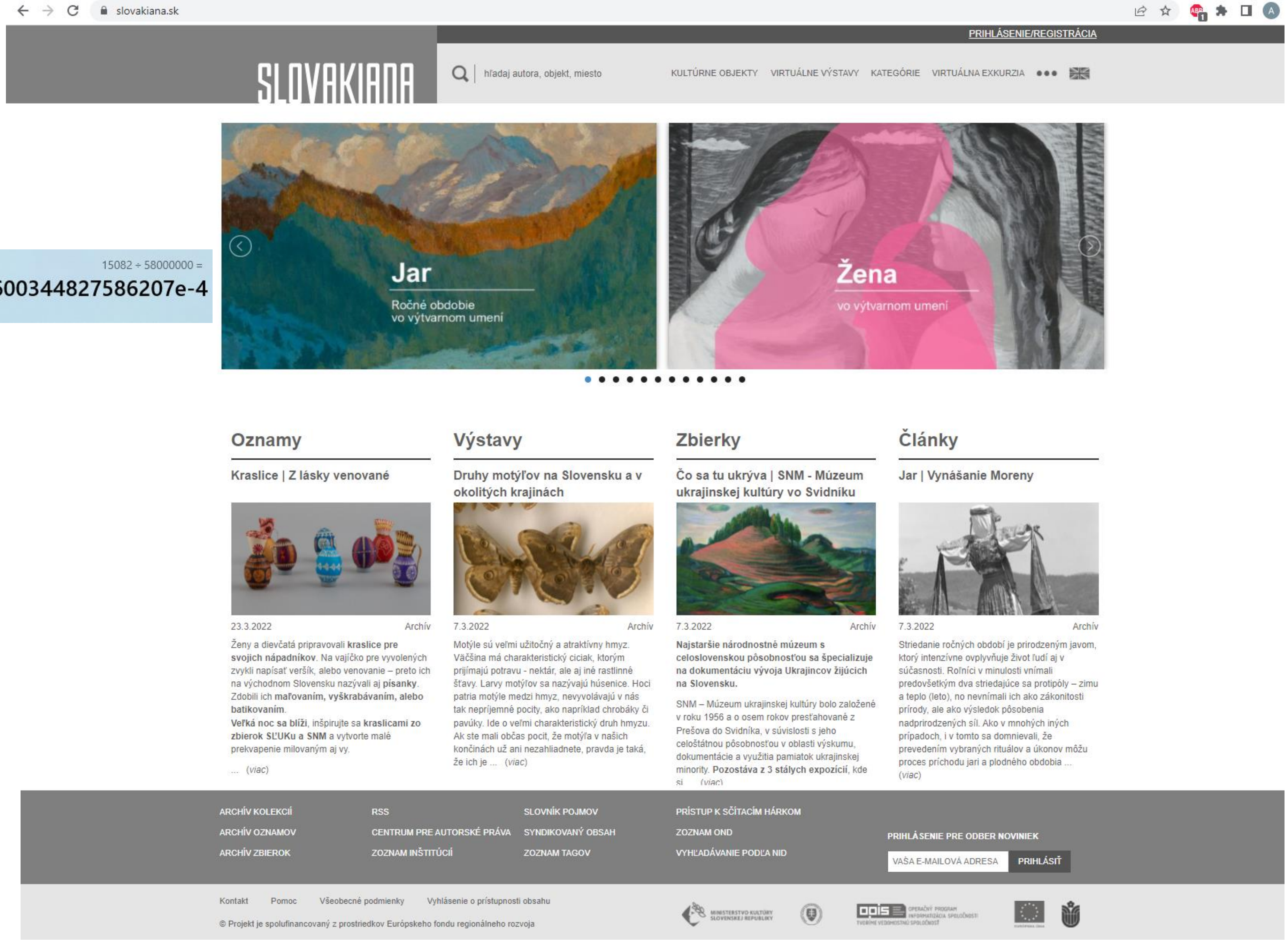
g) zabezpečiť trvalé uchovanie digitálnych objektov

Slovakiana

- **SK/EU = 15082/58 mil. =**

15082 ÷ 58000000 =
2,600344827586207e-4

- **Slovník pojmov ~ CIDOC CRM**
- **https://www.slovakiana.sk/slovnik-pojmov**



Slovakia's culture showcased in Europeana Collections

Europeana Collections features over 58 million objects of which **15,082** are provided by Slovakian institutions.

Aggregation

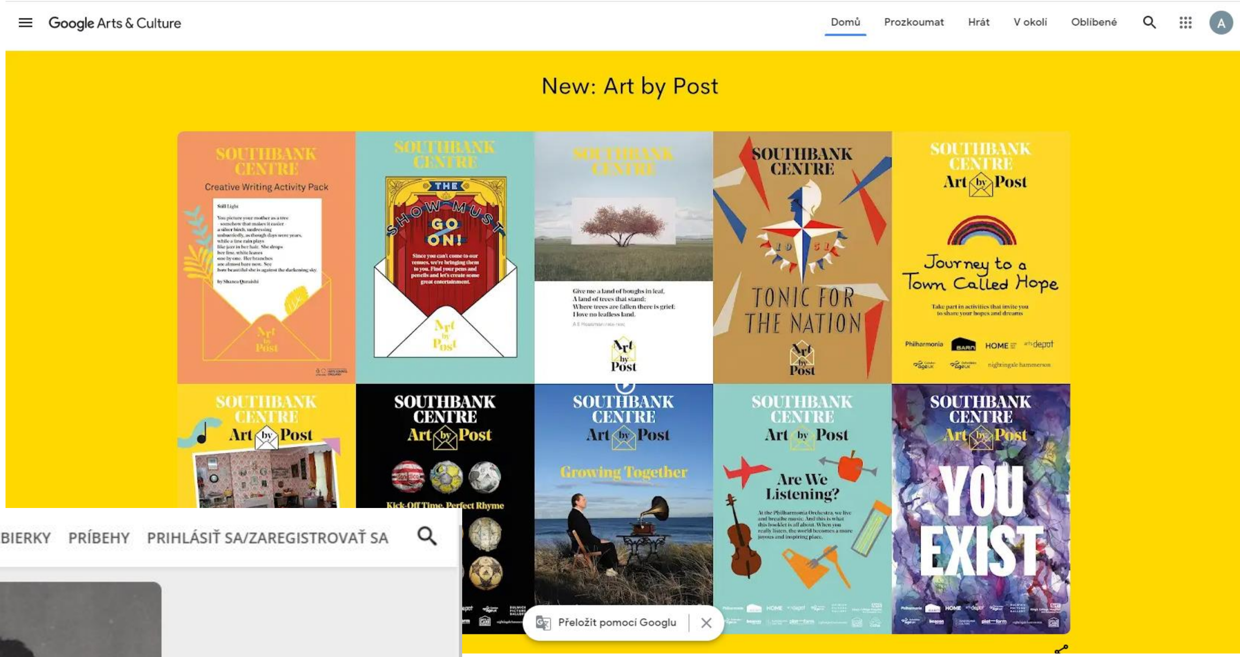
All objects aggregated by Slovakia can be found here[2]. The main point of contact for Slovakia is Slovakiana.

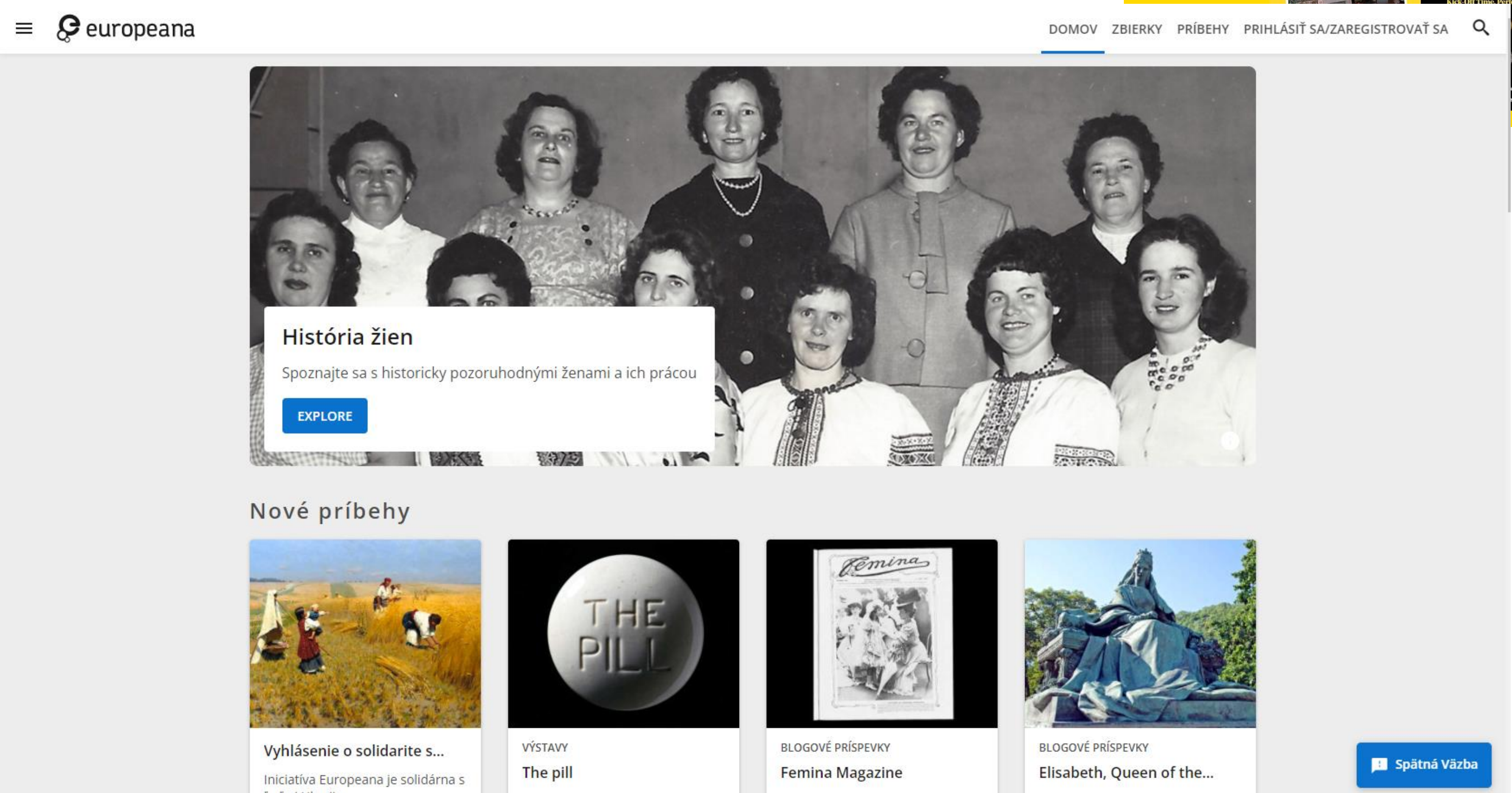
Cultural heritage institutions

4 cultural heritage institutions in Slovakia contribute collections to Europeana. The largest partners are listed in the following table.

Europeana, GAC, **Holbein**

- www.europeana.eu/sk, GAC: SNG+2 2014





Similar initiatives: You@GLAM

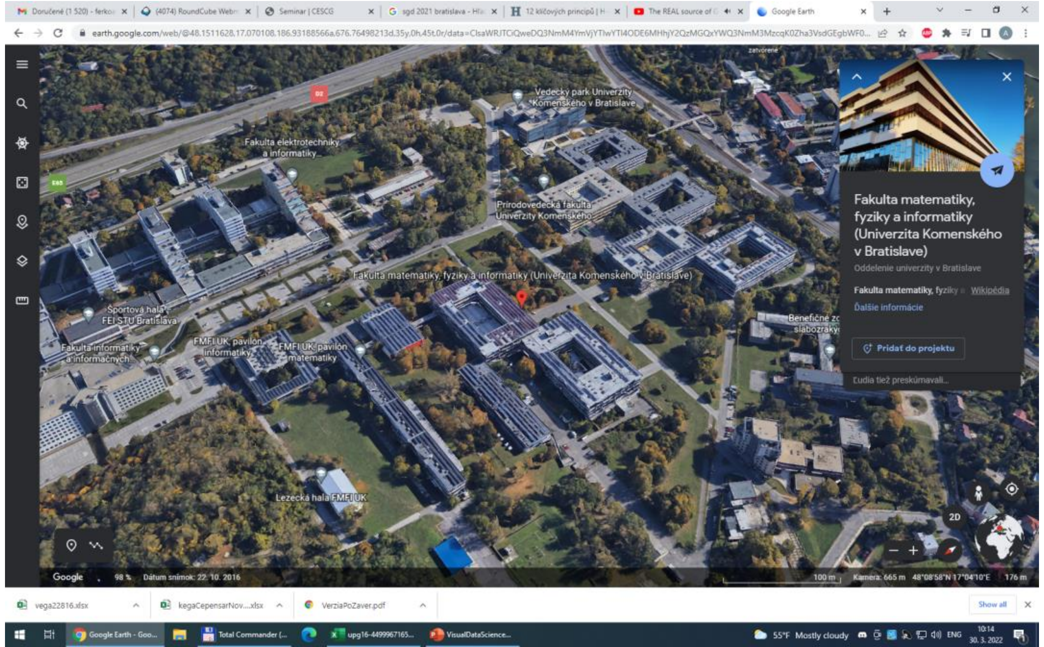
- Europeana is a virtual repository of artworks, literature, cultural objects, relics, and musical recordings/writings from over 2000 European institutions
- The Prado launched a virtual collection, in collaboration with Google Earth, in January 2009. The website contained photos of 14 Prado paintings, each with up to 14 gigapixels.
- The Virtual Museum of Canada is a virtual collection containing exhibits from thousand of Canadian local, provincial and national museums.
- Wikipedia GLAM ("galleries, libraries, archives, and museums' zoological gardens) helps cultural institutions share their resou collaborative projects with experienced Wikipedia editors.

ECO, pollution, post

- Emotional and cognitive overload
- Information pollution generally applies to digital communication, such as e-mail, instant messaging (IM) and social media. The term acquired particular relevance in 2003 when web usability expert Jakob Nielsen published articles discussing the topic.
- Post truth
- False flag operations
- INTERESTINGNESS ~ authenticity, originality, creativity

Technologies 1. ibr

- Image-based rendering:
- Object panorama https://demo.scientica.sk/Objektove_panoramy_3Dmodely/Objetkove_panoramy/Spec7_html/index.html
- Jan Jurkovic, 2021. Objektova panorama kroja https://www.youtube.com/watch?v=xrxx1cw_628
- Street View
- Autostitch
  - http://matthewalunbrown.com/autostitch/autostitch.html

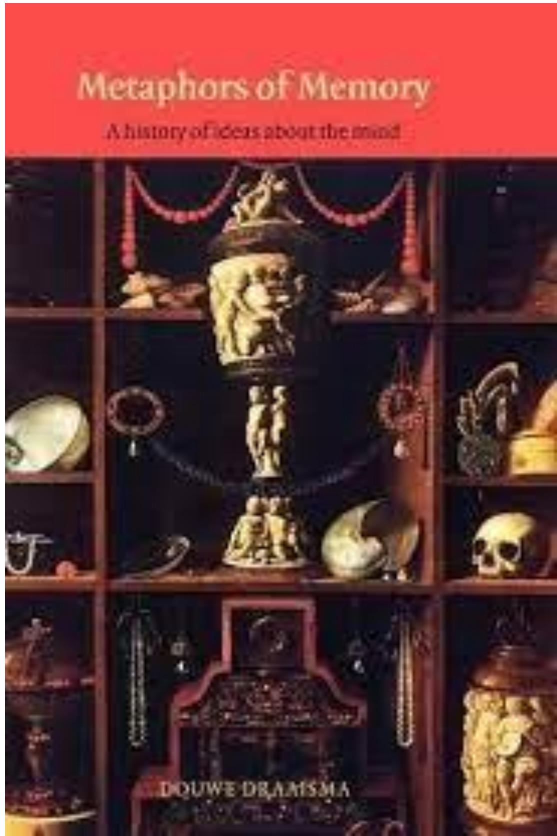


Take home message

- Idea, model, data >> picture (structured, synthetic), objects
- 5 planes JJG, 4 blocks of CG functions, 4 data states ~ 3 workflows

- Sensor >> image (non-structured, natural), views
- Object panorama, Street View, panorama ~ 3 technologies

- World ~ museum (people, things, places)
- Metaphors of memory: sandbox, waxtable… museum



Visual Data Science @ CU

ACM CC, Skiena, Moller... Datova veda@matfyz... vdak