PANEL ON PATTERNS/CONTENT

Patterns and Big Data: Finding a Needle in a Haystack.

ATHENS, GREECE, 23RD OF FEBRUARY 2017

Herwig Mannaert, University of Antwerp, Belgium.

Alexander Mirnig, University of Salzburg, Austria.

Charles Perez, Paris School of Business, France.

Hans-Werner Sehring, Namics, Germany.

Agenda.

FIRST **Namics in a Nutshell.**

SECOND **Proposition.**

FIRST

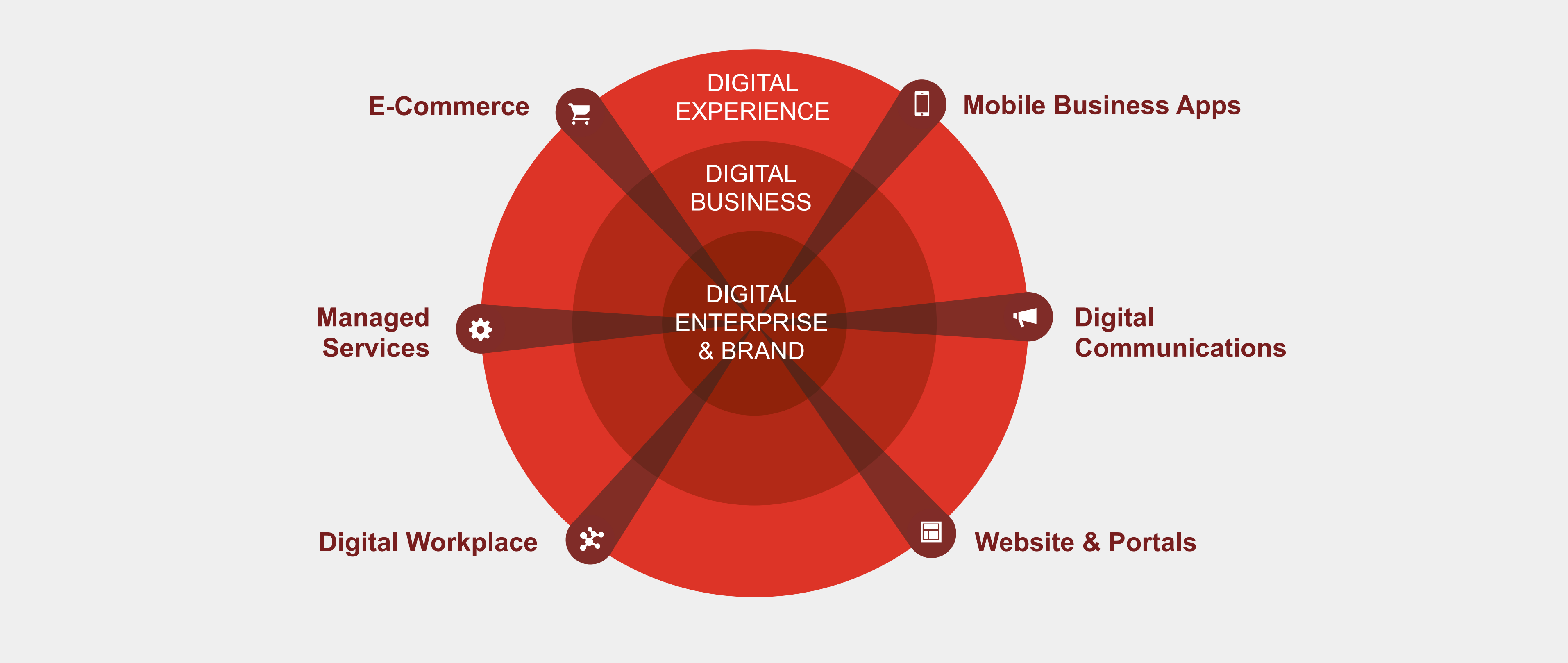
Namics in a Nutshell.

Jürg Stuker

CEO NAMICS

“We are pioneers and experts in the field of digital transformation – and we have been since 1995. As an independent, interdisciplinary full-service partner, we work with you to digitize your business models and critical processes. Your long-term success is the focus of everything we do.”

Our services.



Since 1995. Close proximity. Namics.

6

LOCATIONS

SECOND

Proposition.

Proposition.

Data in databases is typically created according to a schema.

Proposition: To analyze big data collections, a schema might be used "in reverse" in order to match patterns in an existing data collection.

Are there patterns that are similar to a schema/type system?
Is pattern matching in these case a type checking task?

Then, pattern matching could be performed as "duck typing".

Bottom line: We are speaking about a rich "schema"/"type system" here, including higher-level concepts, and constraints, rules, laws.

Some Background from Semiotics and from Philosophy.

Peirce's triadic Semiotics.

Building up Firstness, Secondness, and Thirdness concepts.

Cassirer's Epistemology.

Co-evolution of instances and their classification.

Plus, from a computer science perspective, comp. Joseph Goguen.

Two kinds of classifications: formal structure and meaning.

If I had to solve it: M3L as a "Pattern Recognition" Language.

The Minimalistic Meta Modeling Language (M3L) provides such rich modeling capabilities.

It may be worth the attempt to map it to (existing) Big Data collections.

M3L as a "Pattern Recognition" Language. First Try.

Trivial example: Detect recurring visitors in tracking data.

Q: Do we have recurring visitors? (And how many?)

```
Visitor { Visit is a Time from …; Identifier is a Cookie from …; }
RecurringVisitor is a Visitor {
  Visitor1 is a Visitor; Visitor2 is a Visitor;
  Identifier from Visitor1 is the Identifier from Visitor2;
  Visits; }
|= RecurringVisitor is the Visitor1, the Visitor2 {
  Visit from Visitor1 is a Visits;
  Visit from Visitor2 is a Visits; }
```

hans-werner.sehring@namics.com. Senior Solution Architect.

Thank you. Namics.

Center for
Human-Computer Interaction
Department of Computer Sciences

UNIVERSITÄT
SALZBURG

Patterns: Problems, Solutions, and Everything In-between

About finding the right problem descriptions and the difficulty of proving state-of-the-art solutions.

Content

- Background and Automotive Design
- Design Patterns – a brief overview
- Finding the right problem for your solution
- Proven solutions in a rapidly evolving environment
- Conclusion

contextual.interaction.design.research

Background

Center for Human-Computer-Interaction,
University of Salzburg, Austria

Background:

General Philosophy of Science and Science of Consciousness

Interdisciplinary Workgroup *Neurosignaling*, Department of Zoology, University of Salzburg

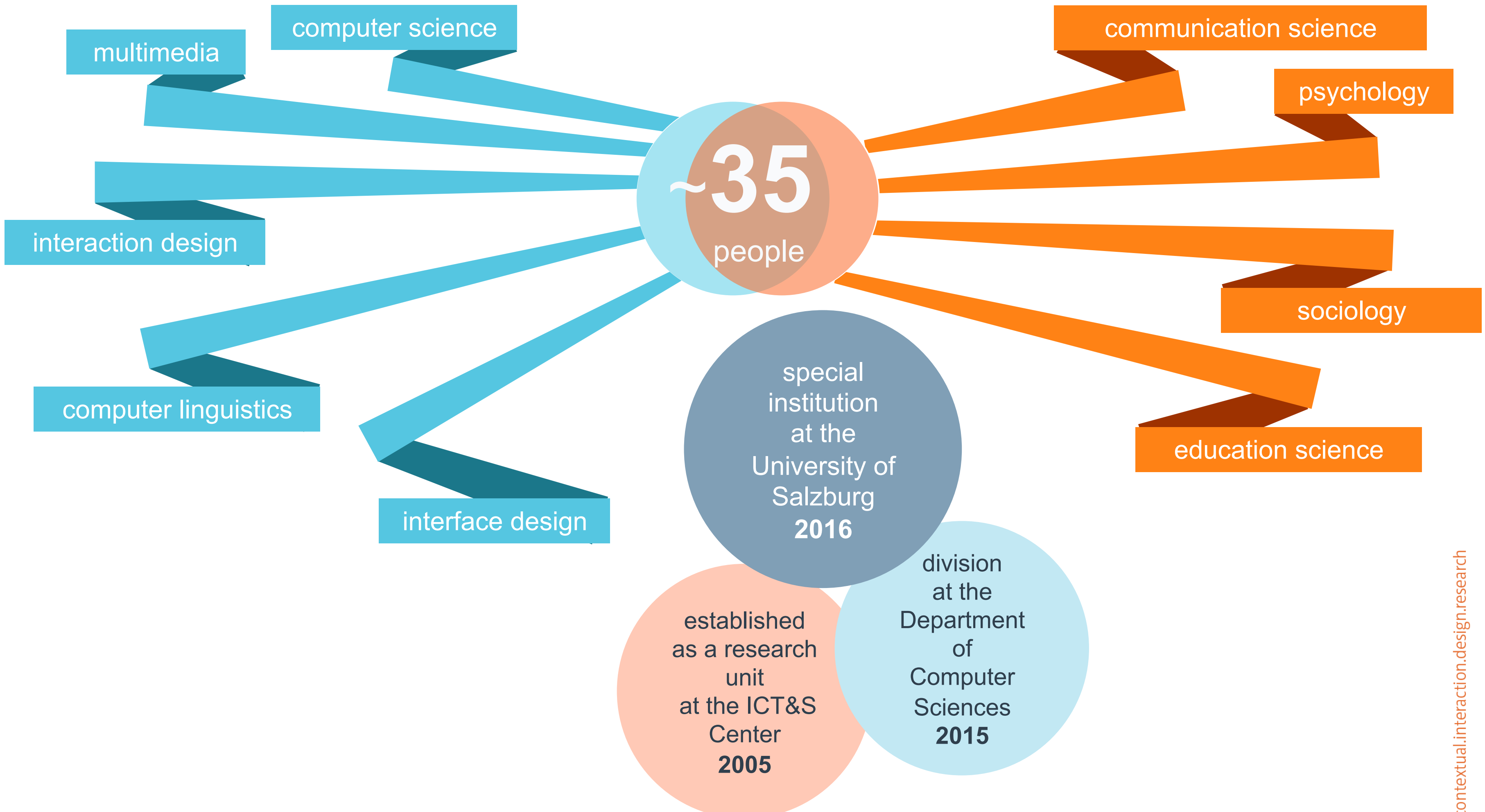
Since 2012: Center for HCI

Main topics:

(Semi-)autonomous vehicles and persuasive interfaces, interface evaluation (Usability and User Experience), definitions and formal approaches in HCI, in-vehicle UIs, theories of consciousness

contextual.interaction.design.research

Background – Center for HCI in Salzburg



contextual.interaction.design.research

Automotive Design

* © Arno Laminger 2015

contextual.interaction.design.research

Design Patterns - Overview

Short Definition:

- A (design) pattern is a structured documentation to a proven solution to a reoccurring problem, embedded in at least one of the contexts it occurs in.

Some advantages:

- they capture expertise and make it accessible to non-experts
- their names collectively form a vocabulary that helps developers communicate better.

Finding the right Problem

Evolution of the problem description

1) Where should what be displayed?

contextual.interaction.design.research

Finding the right Problem

Evolution of the problem description

1) Where should what be displayed?
2) Where should displays be positioned in the cockpit?

contextual.interaction.design.research

Finding the right Problem

Evolution of the problem description

1) Where should what be displayed?
2) Where should displays be positioned in the cockpit?
3) Where does a driver look first? Which areas are quicker to access for the eyes than others?

Finding the right Problem

Final Problem Statement:

- Information sources are spread throughout the cockpits of cars (instrument cluster, center console, in and around the steering wheel). Differently sized displays in different positions make it difficult for the driver to locate the right information at the right time, thus being a potential source of distraction for the driver.

contextual.interaction.design.research

Finding the right Problem

Final Problem Statement:

- Information sources are spread throughout the **cockpits** of cars (instrument cluster, center console, in and around the steering wheel). Differently **size**d displays in different **position**s make it difficult for the driver to locate the right **information** at the right **time**, thus being a potential source of **distraction** for the driver.

contextual.interaction.design.research

Finding the right Problem

What happened?

- As we figure out regarding what we expect from a good solutions, the problem statement gets more and more refined.

contextual.interaction.design.research

Finding the right Problem

What happened?

- As we figure out regarding what we expect from a good solutions, the problem statement gets more and more refined.
- Ref. Philosophy of Science: A question is a good question if we can state what an answer needs to fulfill in order to answer the question either positively or negatively

contextual.interaction.design.research

Finding the right Problem

What happened?

- As we figure out regarding what we expect from a good solutions, the problem statement gets more and more refined.
- Ref. Philosophy of Science: A question is a good question if we can state what an answer needs to fulfill in order to answer the question either positively or negatively
- E.g. What is (the) truth? Is it an object, a linguistic construct, a mathematical function on a set of syntax?

Proving the solution

What does it mean for a novel technology to be proven?

- The Rule of Three!

Proving the solution

What does it mean for a novel technology to be proven?

- The Rule of Three!

- But:

contextual.interaction.design.research

Proving the solution

What does it mean for a novel technology to be proven?

- The Rule of Three!
- But:
  - Do prototypical implementations count?

contextual.interaction.design.research

Proving the solution

What does it mean for a novel technology to be proven?

- The Rule of Three!
- But:
  - Do prototypical implementations count?
  - What if the solution is based on one implementation and lab data?

contextual.interaction.design.research

Proving the solution - Proposal

A novel technology solution is proven if:

- It satisfies the Rule of Three

or:

contextual.interaction.design.research

Proving the solution - Proposal

A novel technology solution is proven if:

- It satisfies the Rule of Three

or:

- It has been successfully implemented at least once in a setting close to its real application context

and

contextual.interaction.design.research

Proving the solution - Proposal

A novel technology solution is proven if:

- It satisfies the Rule of Three

or:

- It has been successfully implemented at least once in a setting close to its real application context

and

- There is compelling evidence to expect the solution to work in other application instances.

contextual.interaction.design.research

Summing Up

- Good answers require good questions being asked

contextual.interaction.design.research

Summing Up

- Good answers require good questions being asked
- It should be possible to provide a provenness-criterion more suitable for rapidly changing domains and environments.

contextual.interaction.design.research

Summing Up

- Good answers require good questions being asked
- It should be possible to provide a provenness-criterion more suitable for rapidly changing domains and environments.

Let me know what you think!

contextual.interaction.design.research

Center for
Human-Computer Interaction
Department of Computer Sciences

UNIVERSITÄT SALZBURG

Contact

Mag. Alexander G. Mirnig
Center for Human-Computer Interaction
University of Salzburg

Jakob-Haringer-Straße 8 / Techno 5
5020 Salzburg, Austria

+43 662 8044/4834
alexander.mirnig@sbg.ac.at
http://hci.sbg.ac.at/mirniga

Patterns and Big Data: Finding a Needle in a Haystack
Detecting Patterns in Big Social Network Data

PATTERNS 2017, February 23, 2017 - Athens, Greece

Charles Perez
PSB Paris School of Business,
Chair D3 Digital, Data, Design,
Paris, France
email:c.perez@psbedu.paris

PSB
PARIS SCHOOL OF BUSINESS

Social network analysis before Big Data

- Euler, 1741: Concept of graph (7 Bridges of Konigsberg)
- Moreno, 1930's: modelling of social relationships as a graph (sociogram, sociometry)
- Simmel, 1950's: importance of nature of relationships more than a group itself
- Barnes, 1954: Social network (Parish in western Norway)
- Milgram, 1967: Small worlds (Familiar Stranger)
- White, 1976: Matrix algebra
- Granovetter, 1973: Strength of weak ties (micro vs macro in sociology)

Sociology, Psychology, Mathematics

The main issue was to grab/find social data

Social network analysis & patterns today
A big data issue

- Social networks (e.g. society) are digitalised (OSN)
  - **Permanent connectivity**: online social networks, smartphones
  - **Permanent memory**: activities are stored as computer data (with context: content, space, time, etc.)

- OSN are composed of billions of users with billions of messages exchanged everyday (Big Data)

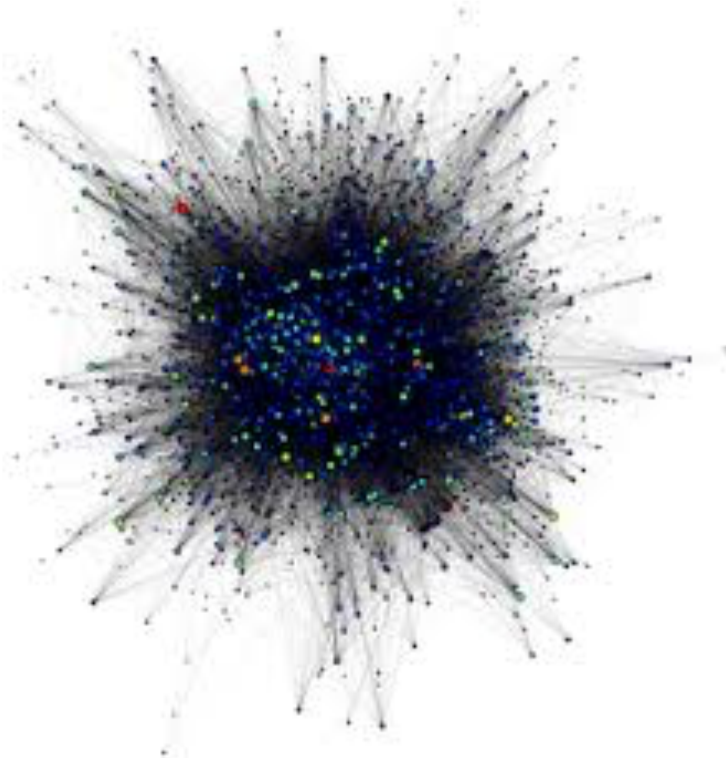
- New challenges & opportunities for patterns discovery

- Implications to business, management, sociology, computer science, mathematics, etc.

The issue is to get value/patterns from Big Data

Network science: a key to patterns discovery?

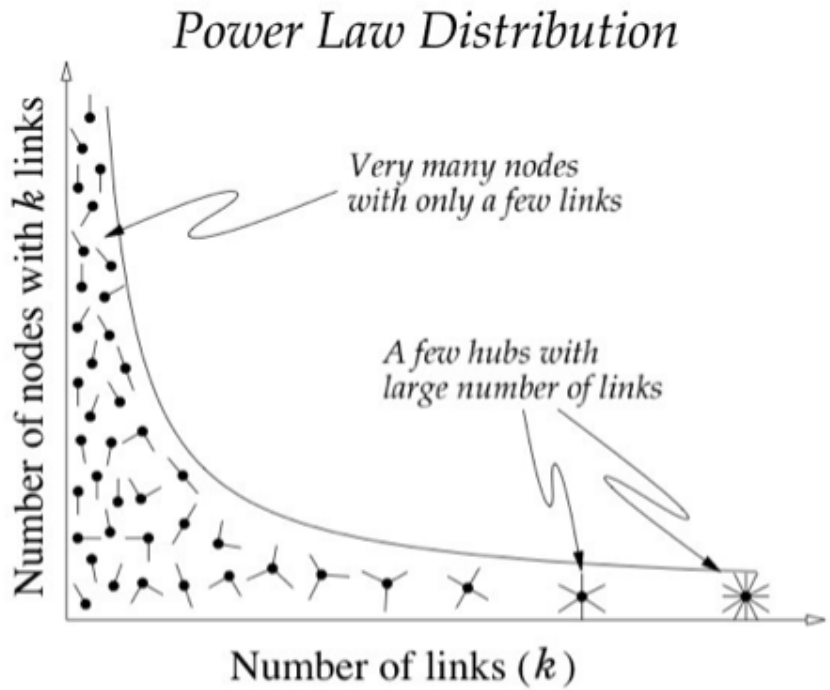
- Modelling complex systems with a basic graph model denoted G(N,E)
  - N = Nodes representing social actors
  - E = Edges representing interactions



OSN patterns discovered in Big Data graphs

- Many OSN have been proved to be:
  - **Small worlds:** small average distance and high clustering coefficient
  - **Scale-free:** Power law degree distribution
- Applications to contagion, influence, buzz, etc.

Ingredients : Time & preferential attachment



From : Network Science ALBERT-LÁSZLÓ BARABÁSI

Questions to be answered and considered for Big data approach to social network analysis

- Generally
  - What data is valuable for patterns search?
  - How to handle Big Data approach with privacy of users?
- Social Network Analysis
  - Identification of the right connections / interactions?
  - Identification of the right analysis scale (space/time)?
  - Automation of patterns detection?
  - What is the source of preferential attachment?
  - What are the limits to human patterns modelling?
- Open questions
  - Are there patterns everywhere in OSN?
  - What is the validity, lifetime of patterns?
  - The scaling of patterns (e.g. migration models)?
  - What is the place of randomness in patterns?

Patterns and Big Data
Finding a Needle in a Haystack

Some Skeptical Thoughts

Herwig Mannaert
University of Antwerp
Normalized Systems Institute

PATTERNS 2017
February 23

Universiteit Antwerpen

My Research

- Study modular structures under change:
  - using systems theoretic stability, thermodynamic entropy, and plain combinatorics
  - to avoid “combinatorial effects” that impede reuse:
    - duplications through lack of separation of concerns
    - ripple effects through various types of coupling
- Examples:
  - Software: law of increasing complexity / lack of reuse
  - Education: duplications in content and descriptions
  - Financial: creation of duplicate ledgers due to reporting
  - Legislation: increasing amount of impacts of new laws
  - …

~ Patterns and Big Data

- Importance of patterns:
  - In order to avoid combinatorial effects, we can derive *very stringent constraints*
  - Designing evolvable modular structures without combinatorial effects requires *domain patterns*
    - E.g. software, education, financial reporting
- Challenges of big data:
  - In order to master complexity, engineers need *hierarchical structure*
    - E.g microprocessors, rockets, software systems
  - Structure is in general both *multi-dimensional* and has *many layers of abstraction*

Thought 1: On Causality

Once upon a time:

Not every correlation is causal. For instance, the correlation between people falling down the stairs and having gray hair, is not causal.

— *Professor Probability Theory*

Thought 1: On Causality

- The continued quest for causal relations is based on the discovery of correlations:
  - Student major in high school as a prerequisite
  - Work absence as a measure of poor health
  - …
- *Could it be that the sheer massive amount of correlations that are being probed today, could jeopardize the validity of traditional measures for statistical relevance ?*

Thought 2: On Relevance

Once upon a time:

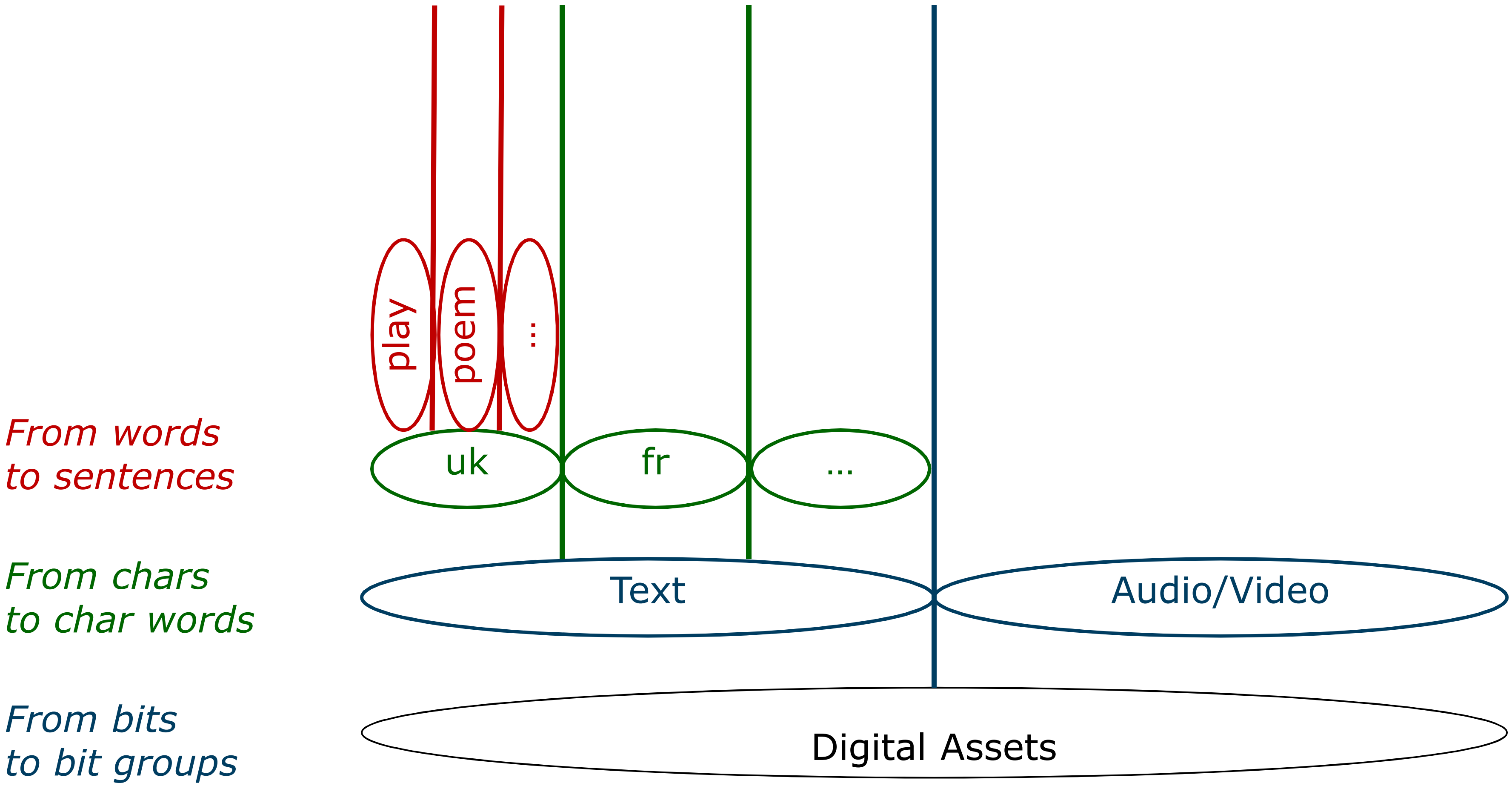
The human genome is complex and often startling. For instance, the correlation seems higher between a human genome and that of a banana, than with the one of a frog.

— *Professor Human Genetics*

Thought 2: On Relevance

- The continued quest for causal relations and/or patterns is based on low-level data:
  - Individuals clicking on individual links
  - Individuals registering an interest
  - …
- *Could it be that on a bit-per-bit basis, the correlation between a Shakespeare play and an adult movie is higher, than the correlation between two Shakespeare plays ?*

Thought 2: On Relevance



Thought 3: On Accordance

Once upon a time:

The purpose of information systems is to create a mirror of the world. In this way, we establish a digital reproduction of the world, and make it available for scientific research.

— *Professor Information Systems*

Thought 3: On Accordance

- The continued quest for data patterns and/or predictions is based on existing models:
  - Entries of existing databases
  - Traces of existing transaction systems
  - …
- *Could it be that the mirrors of information systems today exhibit severe scattering and diffraction, due to increasing complexity and structure degradation ?*

Thought 4: On Suitability

Once upon a time:

Why do we need a live transaction system to tell us the state of a specific IoT device ? The big data lake is able to tell us everything we want to know.

— *Software Solutions Architect*

Thought 4: On Suitability

- The continued quest for data patterns and/or predictions is based on history data:
  - Entries of existing databases
  - Traces of existing transaction systems
  - …
- *Could it be that sometimes, some people turn to statistical estimations and predictions, when the required information is simply available or can easily be computed ?*