



dwhtechnical solutions
simulation services

Integrated Processes for Modelling & Simulation Health Care Systems

Dr. Nikolas Popper

Director DEXHELPP – Decision Support for Health Policy and Planning, Vienna Coordinator Centre for Computational Complex Systems, TU Wien CSO dwh GmbH – Simulation Services and Technical Solutions, Vienna

deMhelpp

Decision Support for Health Policy and Planning















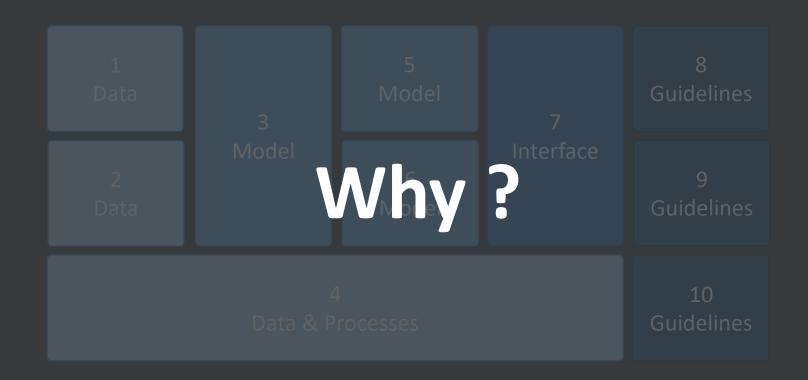
10 Concepts to Integrate

Why?

- -> Dynamics & Complexity
- -> Economics
- -> Big Data

How?

- -> 4th Paradigm
- -> Modelling Dynamics
- -> Data vs. Models
- -> Processes





Man schützt das Klima am besten, indem man kinderlos bleibt, 58,6 Tonnen CO₂ pro Jahr lassen sich so einsparen, heißt es. So ein Unsinn! by Gregor Walter-Drop, DIE ZEIT, 21. März 2019, Nr. 13/2019

Strange Decision Support



"We recommend four widely applicable high-impact (i.e. low emissions) actions with the potential to contribute to systemic change and substantially reduce annual personal emissions: having one fewer child (an average for developed countries of 58.6 tonnes CO₂-equivalent (tCO₂e) emission reductions per year), living car-free (2.4 tCO₂e saved per year), avoiding airplane travel (1.6 tCO₂e saved per roundtrip transatlantic flight) and eating a plant-based diet (0.8 tCO₂e saved per year). "

Seth Wynes and Kimberly A Nicholas 2017 Environ. Res. Lett. 12 074024

Kritik: "...Modellannahmen gehen von falschen Annahmen aus, stellen Fakten absurd zusammen um auf die Ergebnisse zu kommen"



The Seven Tools of Causal Inference, with Reflections on Machine Learning by Judea Pearl, Communications of the ACM, March 2019, Vol. 62 No. 3, Pages 54-60, 10.1145/3241036

Causal Inference refl. ML



"...Machine learning researchers have noted current systems lack the ability to recognize or react to new circumstances they have not been specifically programmed or trained for.... "

"...Another obstacle is "explainability," or that "machine learning models remain mostly black boxes" unable to explain the reasons behind their predictions or recommendations, thus eroding users' trust..."

Classification of causal information in terms of the kind of questions each class is capable of answering:

Level (Symbol)	Typical Activity	Typical Questions	Examples
1. Association $P(y x)$	Seeing	What is? How would seeing X change my belief inY?	What does a symptom tell me about a disease? What does a survey tell us about the election results?
2. Intervention $P(y do(x), z)$	Doing, Intervening	What if? What if I do X?	What if I take aspirin, will my headache be cured? What if we ban cigarettes?
3. Counterfactuals $P(y_x x',y')$	Imagining, Retrospection	Why? Was it X that caused Y? What if I had acted differently?	Was it the aspirin that stopped my headache? Would Kennedy be alive had Oswald not shot him? What if I had not been smoking the past two years?

Economics Example Health Systems



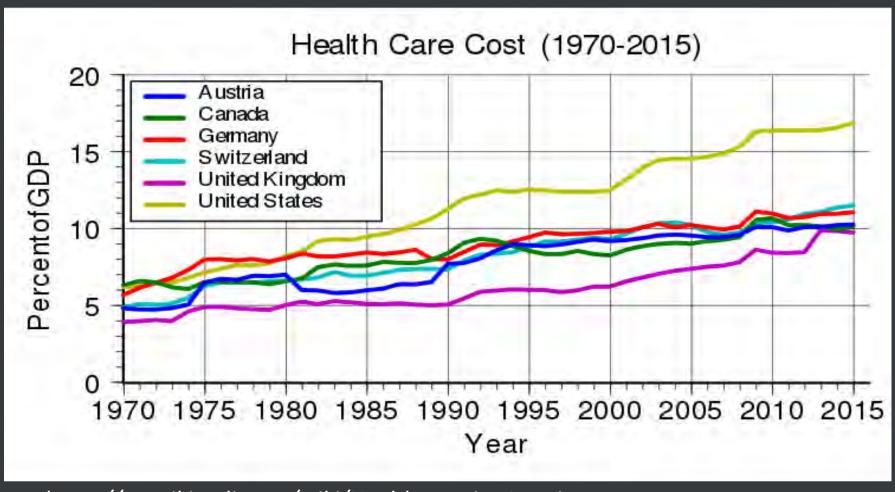
Dissociation of Prooved Benefit of Costs for Medical Care

Pharma- ceutical	Application	Discovery	Costs € (Patient/Yea r)	Effects (Survival)
Insulin	Typ 1 Diabetes	1920	500	Decades
Statins	Cardiology	1990	5.000	Years
Monoclonal Antibodies	Onkology	2000	50.000	Month/Weeks
Encyme Alternation	Metabolism	2010	500.000	???

Source: C. Wild, Ludwig Boltzmann Institute for Health Technology Assessment, Vienna





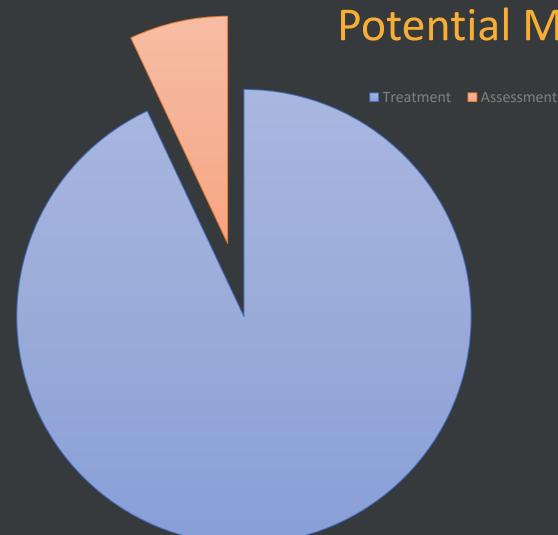


https://en.wikipedia.org/wiki/Healthcare_in_Austria



Health Care Assessment of Interventions

Potential Market



Austria – 2-3 Billion €
UK – 13-20 Billion €
Germany – 17-27 Billion €



Development vs. Measuring

Advanced Treatment

New Medication

Cutting Edge Technology

Innovative Diagnosis

Fair Distribution of Health Services

Disease Prevention

Dealing with Increasing Costs

Estimating Reachable Outcomes

Measuring of Adherence

Managing # of Patients

Developing Registers

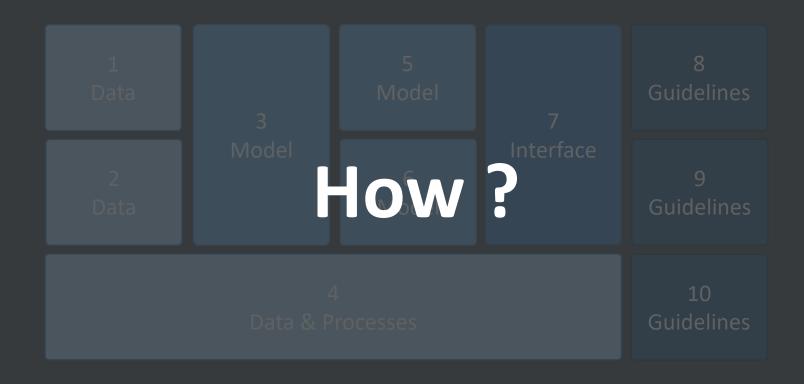
Benchmarking Interventions



Economical Need

Combination of Methods Complex & Dynamic Processes Increasing & Complex Data





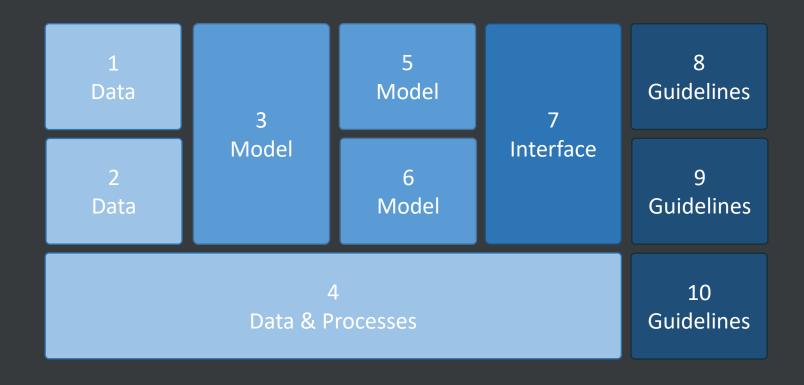


Data Driven System Simulation

10 Concepts to Integrate: Implementing Future Simulation Models, including Data Processes

https://www.eurosim.info/tcs/tc-ddss/





The Fourth Paradigm by Jim Gray (Microsoft)



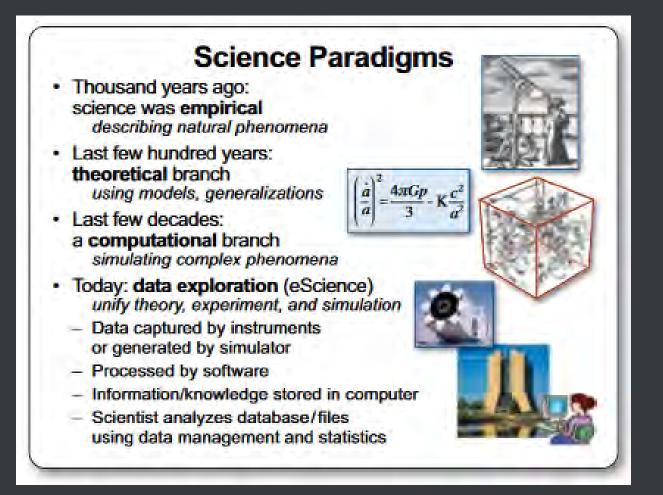
10 Concepts to Integrate

Why?

- -> Dynamics & Complexity
- -> Economics
- -> Big Data

How?

- -> 4th Paradigm
- -> Modelling Dynamics
- -> Data vs. Models
- -> Processes



Describe Model Simulate **Explore Data**

Hey, T., Tansley, S. & Tolle, K. (eds.) (2009). The Fourth Paradigm: Data-Intensive Scientific Discovery. Redmond, Washington: Microsoft Research.

Modelling Dynamics



10 Concepts to Integrate

Why?

- -> Dynamics & Complexity
- -> Economics
- -> Big Data

How?

- -> 4th Paradigm
- -> Modelling Dynamics
- -> Data vs. Models
- -> Processes



Which Questions are to be answered?



Which Data
Ressources are
available?



Which Systems & Processes are described?





N. Popper:"Comparative Modelling and Simulation, A Concept for Modular Modelling and Hybrid Simulation of Complex Systems", PhD Thesis

Modelling Dynamics



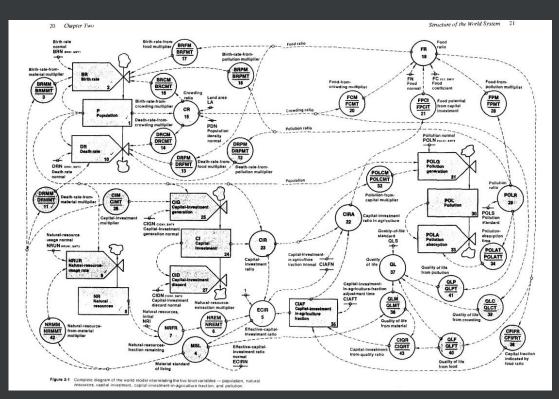
10 Concepts to Integrate

Why?

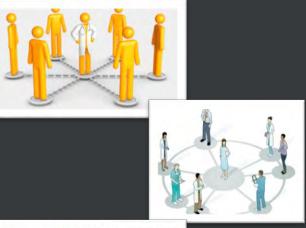
- -> Dynamics & Complexity
- -> Economics
- -> Big Data

How?

- -> 4th Paradigm
- -> Modelling Dynamics
- -> Data vs. Models
- -> Processes



Structures & Knowledge





Forrester, Jay W., World Dynamics. 1973 second ed. 1971, Waltham, MA: Pegasus Communications. P. 144, Reproduced by permission of Jay W. Forrester.

Data vs. Models in Health System Research



10 Concepts to Integrate

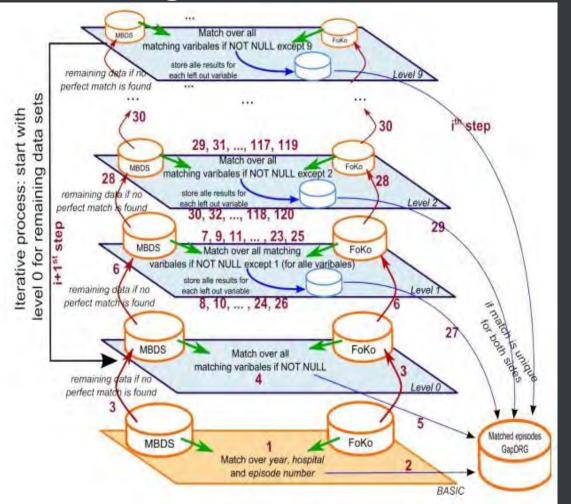
Why?

- -> Dynamics & Complexity
- -> Economics
- -> Big Data

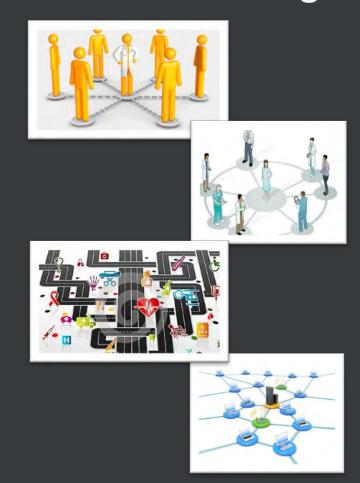
How?

- -> 4th Paradigm
- -> Modelling Dynamics
- -> Data vs. Models
- -> Processes

Big Data Sets



Structures & Knowledge



Modelling and Simulation Processes



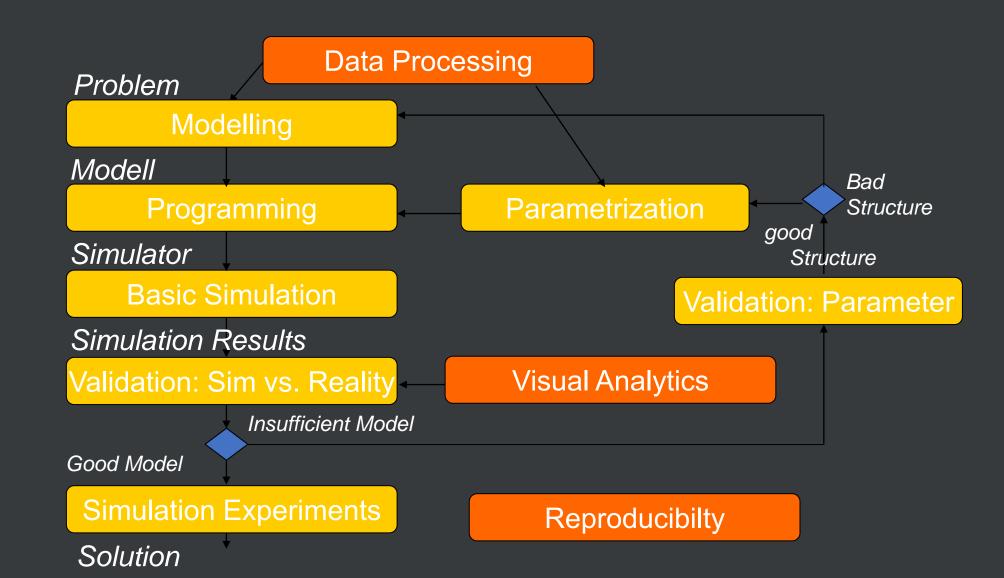
10 Concepts to Integrate

Why?

- -> Dynamics & Complexity
- -> Economics
- -> Big Data

How?

- -> 4th Paradigm
- -> Modelling Dynamics
- -> Data vs. Models
- -> Processes



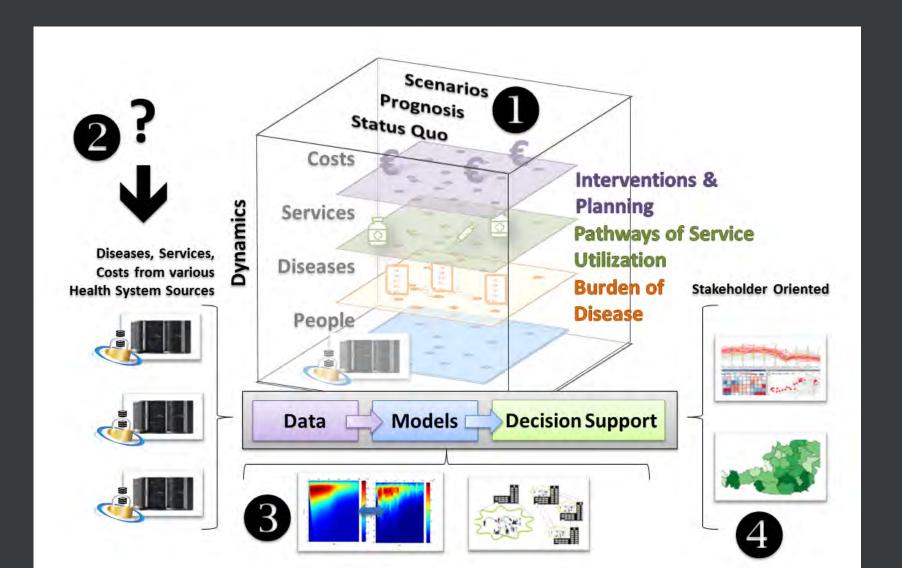
Example Health System: DEXHELPP



DEXHELPP

Decision Support for Health Policy and Planning: Methods, Models and Technologies based on existing health care data

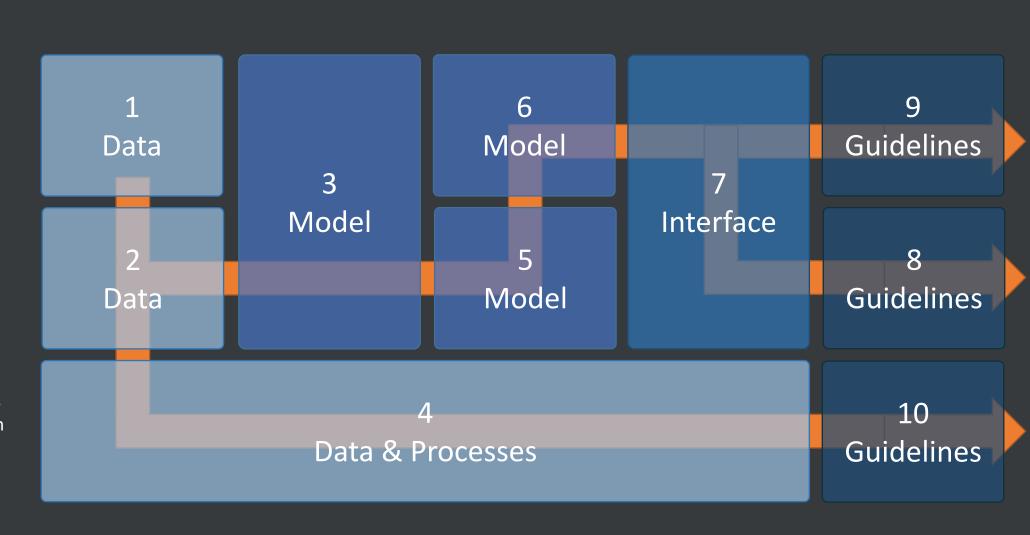




10 Concepts to Integrate



- Methods to Assess and Improve Quality of Data
- Potential to Integrate Missing Data
- 3. Modular & Efficient Solutions
- 4. Reproducible Processes
- Different Methods for Different Questions (Complexity)
- 6. Comparability of Results
- 7. Make it Understandable
- 8. Open and Independent Solutions
- 9. Priority for Data Security and Stake Holder Interests
- 10. Broad Applications (Health System, Energy, Industry, Energy, Mobility, Infrastructure)



Concept 1



10 Concepts to Integrate

 Methods to Assess and Improve Quality of Data

Problems of Collected Data

- Biased Collection of Data e.g. Sensor
 Data or Reimbursement Data is available
- Pre-processing at Various Stakeholders
- Privacy Demands

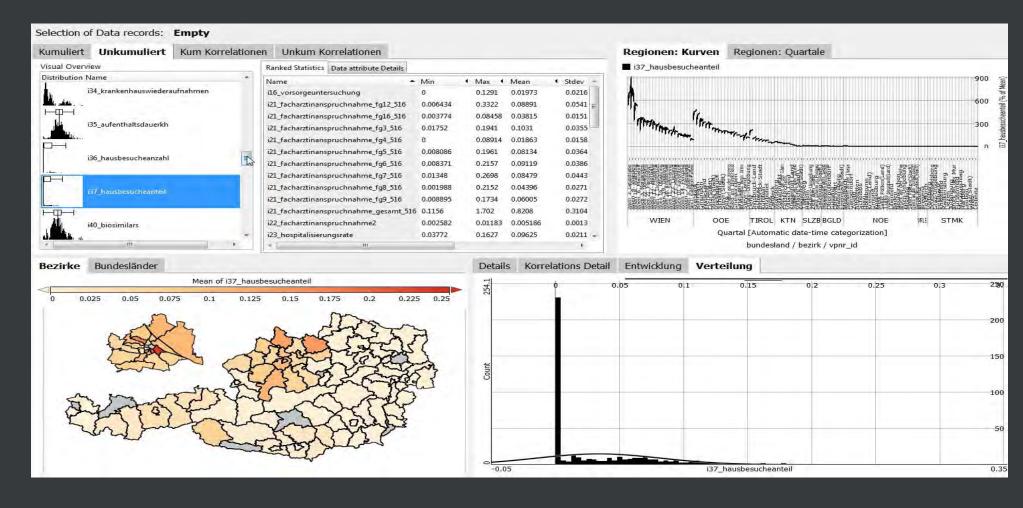
METHOD: Explorative Visual Computing – Visual Analytics and Statistics

Interactive Dashboards



10 Concepts to Integrate

1. Methods to Assess and Improve Quality of Data



H. Piringer, M. Gyimesi et al: Stroke Analyse und Modellierung, DEXHELPP

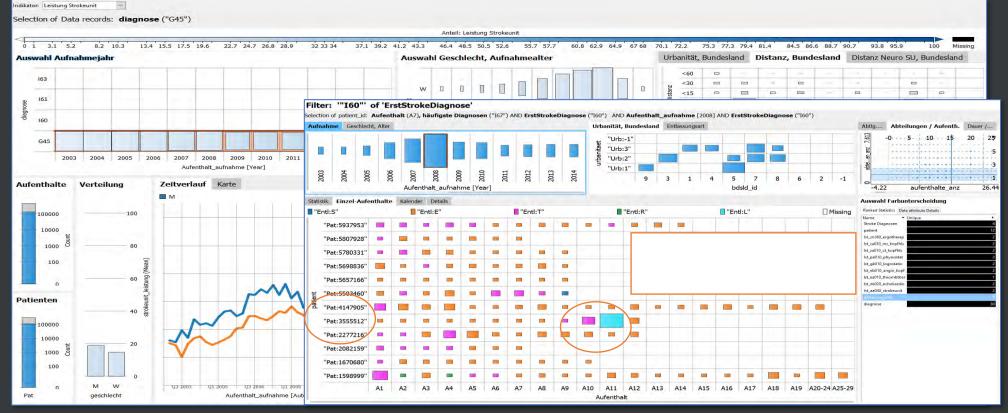
Data on Stroke Treatment



10 Concepts to Integrate

 Methods to Assess and Improve Quality of Data **Applied to:** Data and Trend analysis of stroke treatment

- Percentage of patients being treated in stroke units
- Inspection of patient histories for outlier detection and hypothesis generation



H. Piringer, M. Gyimesi et al: Stroke Analyse und Modellierung, DEXHELPP

Concept 2



10 Concepts to Integrate

- 1. Methods to Assess and Improve Quality of Data
- Potential to Integrate Missing Data

Integration of Different Data Sets

- PatPre-processing at Various Stakeholders
- Unstructured and Different Structures
- Privacy Demands

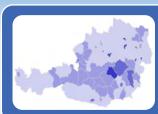
METHOD: Data Processes (Integration & Linkage) & Modelling Tools (Parametrization & Calibration)

Data Levels



10 Concepts to Integrate

- 1. Methods to Assess and Improve Quality of Data
- 2. Potential to Integrate Missing Data



Austrian Health System Data

DEXHELPP



Provincial Data

DEXHELPP 2018 PLUS



Clinical Data





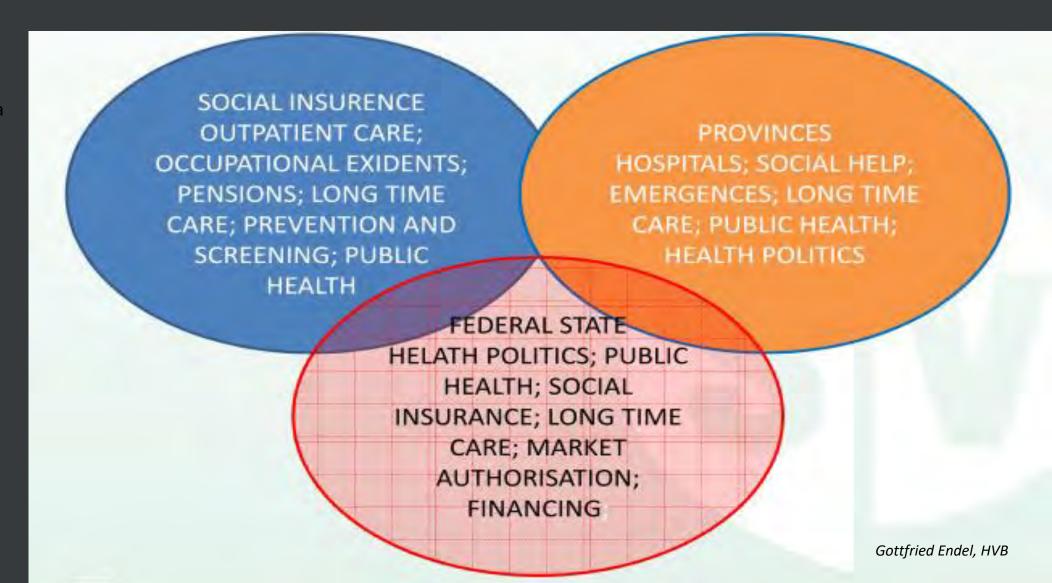
OMICS (excluded at the moment)

DEXHELPP FUTURE

Health System Austria 3 Areas



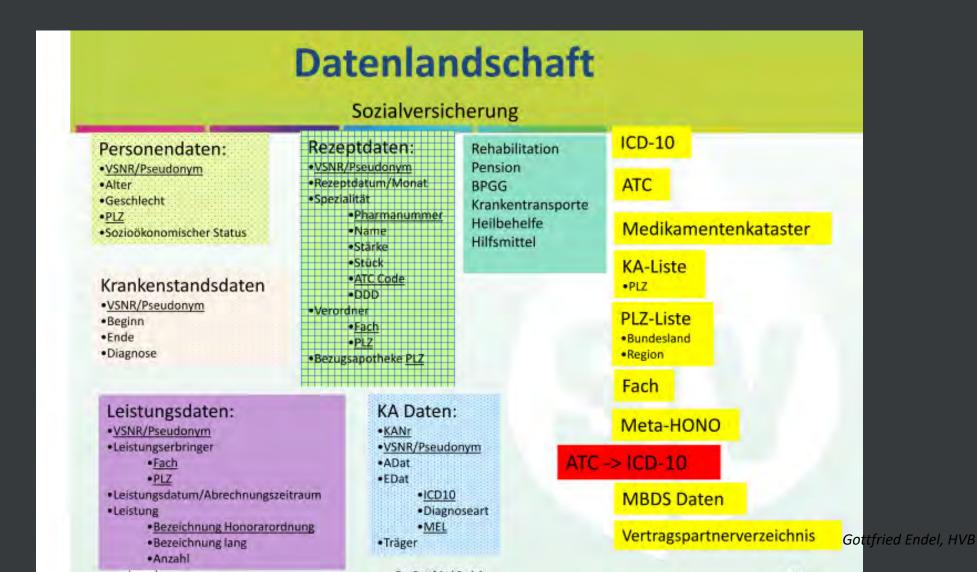
- Methods to Assess and Improve Quality of Data
- 2. Potential to Integrate Missing Data



Data Social Insurancess



- Methods to Assess and Improve Quality of Data
- 2. Potential to Integrate Missing Data



Data Provinces & Austria



- Methods to Assess and Improve Quality of Data
- 2. Potential to Integrate Missing Data

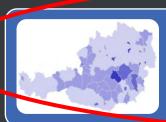


Data Processing DEXHELPP until now....



10 Concepts to Integrate

- Methods to Assess and Improve Quality of Data
- 2. Potential to Integrate Missing Data



Austrian Health System Data



Provincial Data



Clinical Data

Hospitals/Inpatients
Phycisians / Outpatients



OMICS (excluded at the moment)







Diagnosis

Treatment

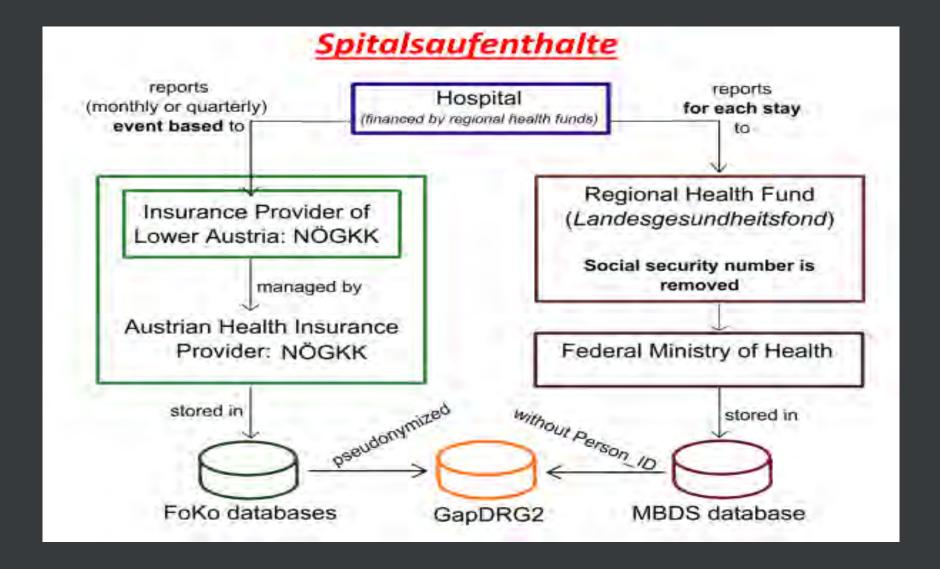
- Therapy
- Medication

Accounting

Example Record Linkage



- Methods to Assess and Improve Quality of Data
- Potential to Integrate Missing Data

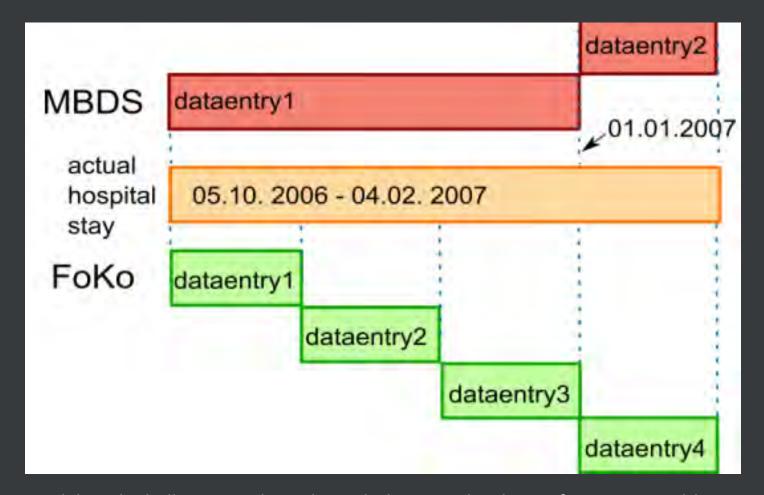


Example Record Linkage



10 Concepts to Integrate

- Methods to Assess and Improve Quality of Data
- 2. Potential to Integrate Missing Data

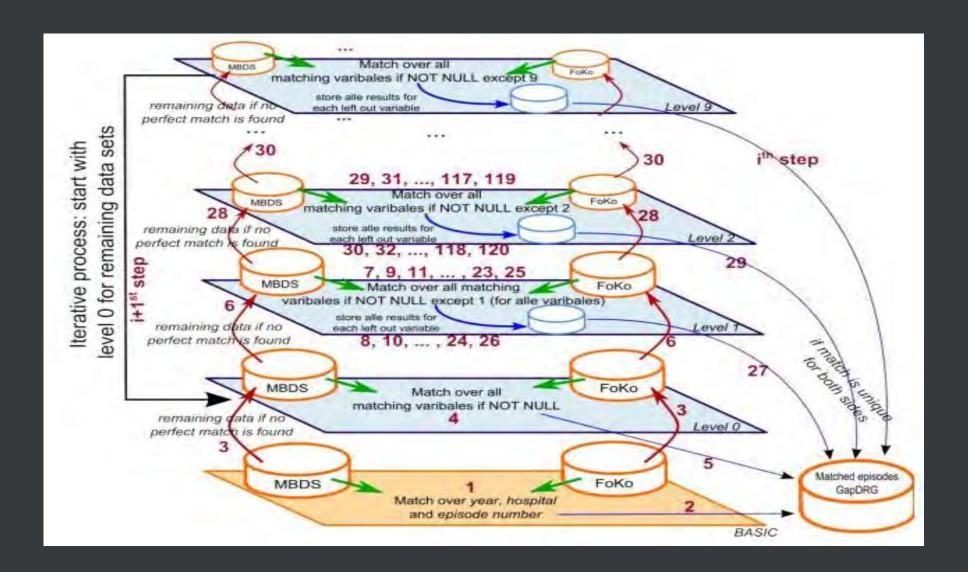


B. Glock, F. Endel et al: Challenges and Results with the Record Linkage of Austrian Health Insurance Data of Different Sources, Informatics for Health Conference 2017 (24 – 26. April, Manchester, UK)





- Methods to Assess and Improve Quality of Data
- 2. Potential to Integrate Missing Data



Concept 3



10 Concepts to Integrate

- 1. Methods to Assess and Improve Quality of Data
- Potential to Integrate Missing Data
- 3. Modular & Efficient Solutions

Modular Concepts for Models

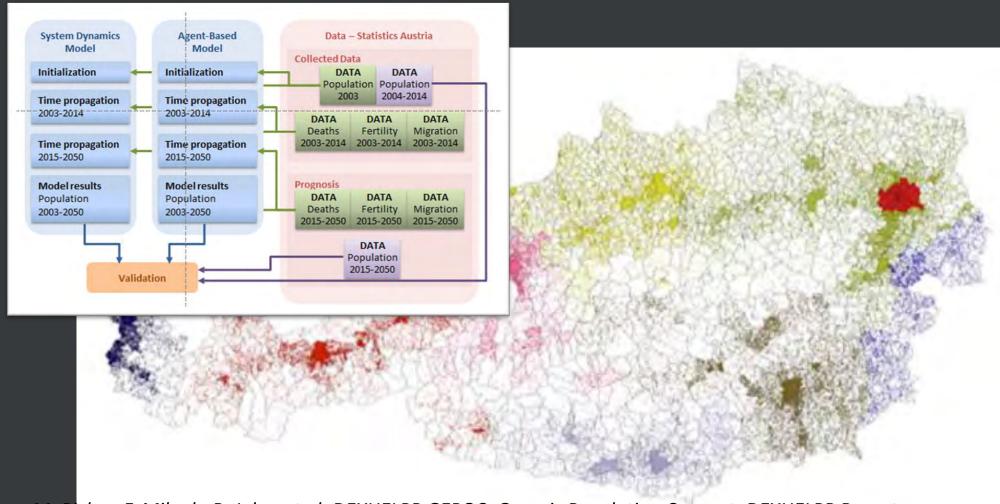
- We can find wrong Data
- We can change wrong Data ...
- ...even over time and when integrated
- We need transparent, "simple" Models

METHOD: Modular Models, Coupling of Models

Virtual Population



- Methods to Assess and Improve Quality of Data
- Potential to Integrate Missing Data
- 3. Modular & Efficient Solutions

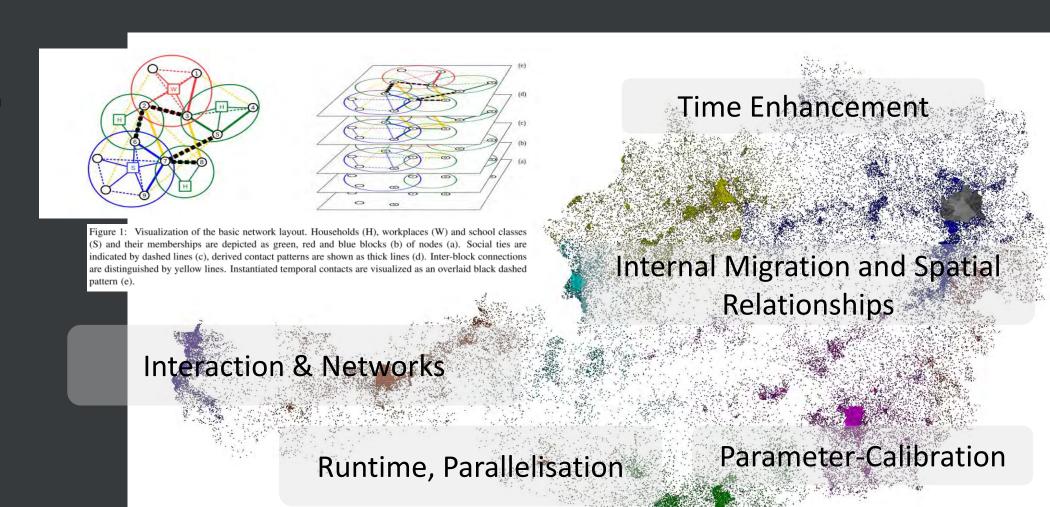


Social Network Layers

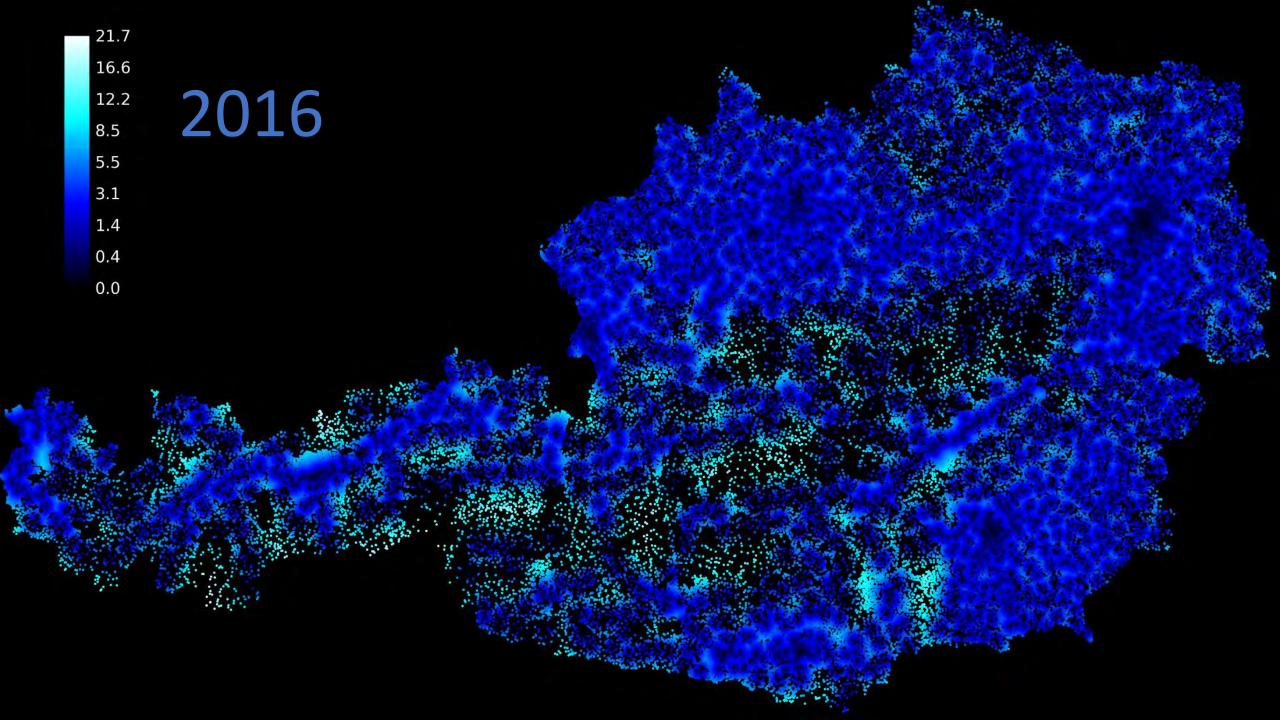


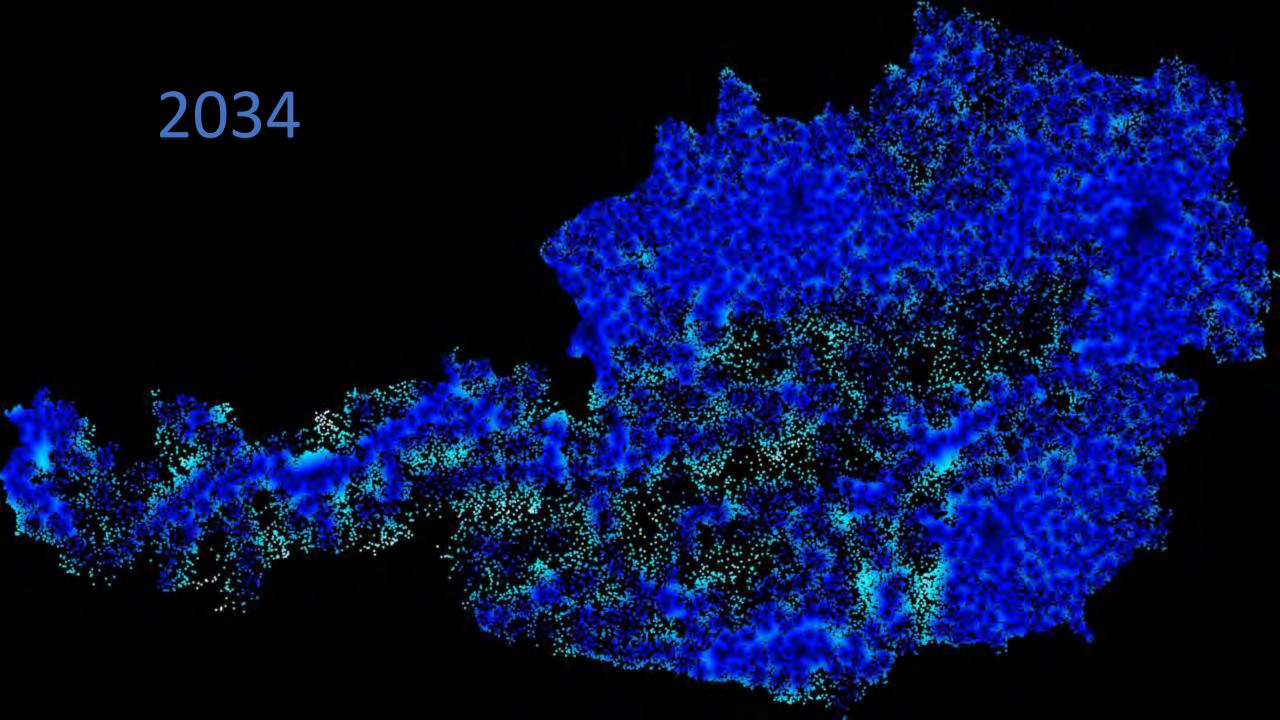
10 Concepts to Integrate

- Methods to Assess and Improve Quality of Data
- Potential to Integrate Missing Data
- 3. Modular & Efficient Solutions



G. Schneckenreither, N. Popper, "Dynamic Multiplex Social Network Models on Multiple Time Scales for Simulating Patterns in Contact Formation and Epidemic Spread" submitted Winter Simulation Conference 2017







10 Concepts to Integrate

- 1. Methods to Assess and Improve Quality of Data
- Potential to IntegrateMissing Data
- 3. Modular & Efficient Solutions
- 4. Reproducible Processes

Reproducible Processes

- Da No "General Model" possible, because of...
- ...different time scales or characteristics.
- Based on Provenance and Modularity
- Managing Tools for Data & Models

METHOD: Validation & Data Citation



10 Concepts to Integrate

- 1. Methods to Assess and Improve Quality of Data
- Potential to IntegrateMissing Data
- 3. Modular & Efficient Solutions
- 4. Reproducible Processes

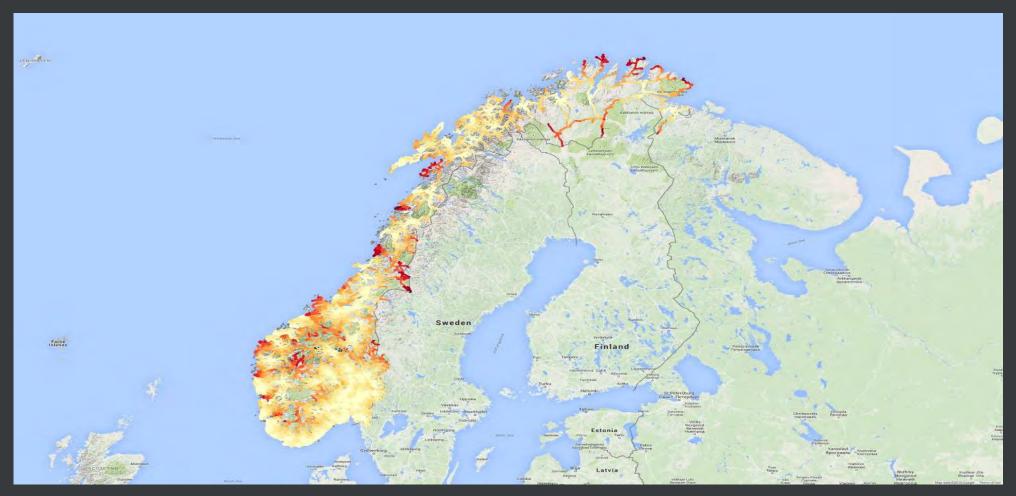
Comparative Effectiveness Research on Psychiatric Hospitalisation

Description: Comparative Effectiveness Research on Psychiatric Hospitalisation by Record Linkage of Large Administrative Data Sets

CEPHOS-LINK is a comparative European register-based study, performed by record linkage technique. It will develop a methods toolkit for conducting record linkage studies in the mental health care field and produce recommendations, guidelines and a set of decision support tools for decision makers in the field of mental health system.



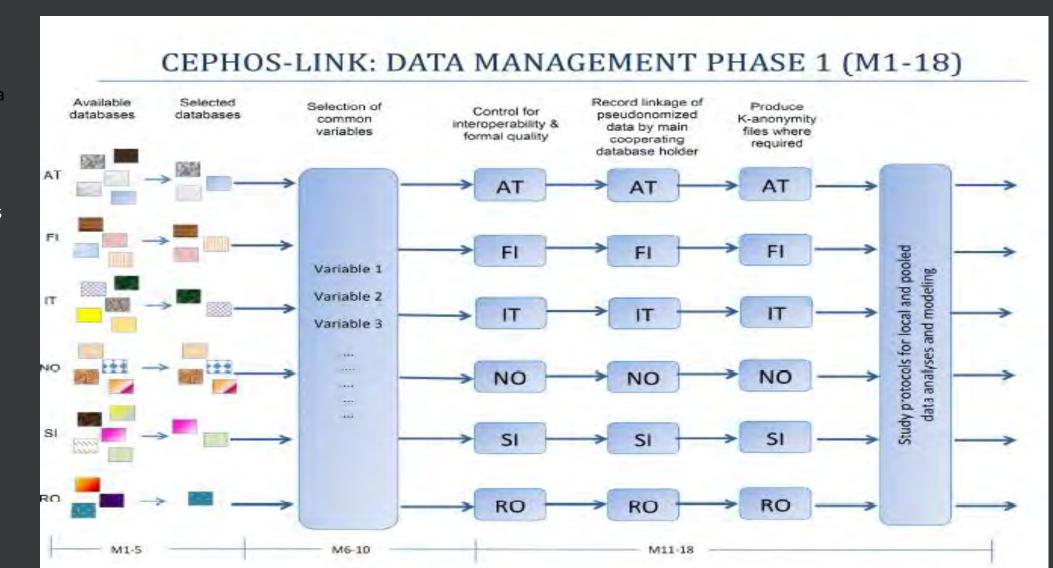
- Methods to Assess and Improve Quality of Data
- 2. Potential to Integrate Missing Data
- 3. Modular & Efficient Solutions
- 4. Reproducible Processes



C. Urach, G. Zauner et al "Statistical methods and modelling techniques for analysing hospital readmission of discharged psychiatric patients: a systematic literature review"; BMC Psychiatry (2016)

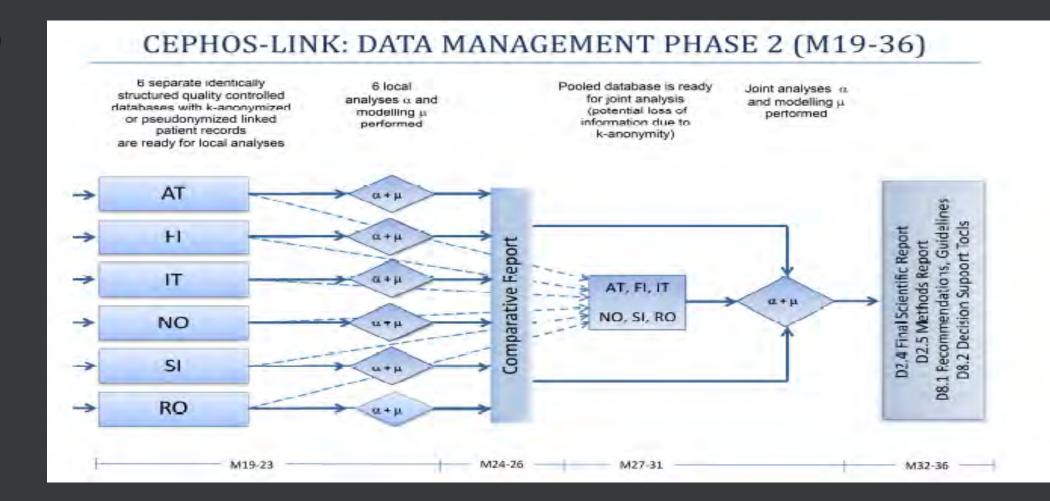


- 1. Methods to Assess and Improve Quality of Data
- Potential to IntegrateMissing Data
- 3. Modular & Efficient Solutions
- 4. Reproducible Processes





- Methods to Assess and Improve Quality of Data
- Potential to IntegrateMissing Data
- 3. Modular & Efficient Solutions
- 4. Reproducible Processes





10 Concepts to Integrate

- 1. Methods to Assess and Improve Quality of Data
- Potential to IntegrateMissing Data
- 3. Modular & Efficient Solutions
- 4. Reproducible Processes
- Different Methods for Different Questions (Complexity)

Different Models for Different Questions

- Da Stable model, already checked
- Better Data doesn't improve
- Model changes don't improve
- Need to Integrate Complexity/Dynamics
- Different (comparable) methods needed

Data & Processes

METHOD: Methods for Choosing Models

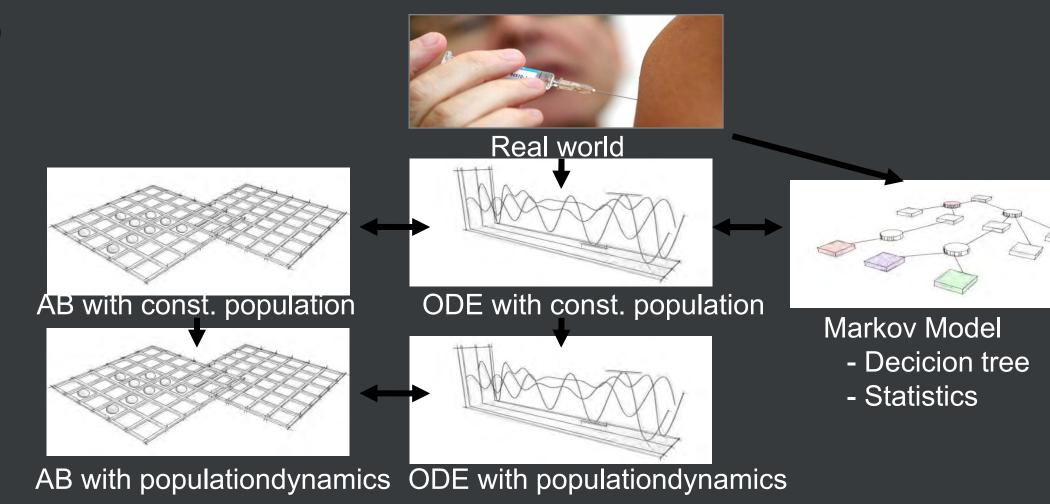
Example Comparison Vaccination



10 Concepts to Integrate

- Methods to Assess and Improve Quality of Data
- 2. Potential to Integrate Missing Data
- 3. Modular & Efficient Solutions
- 4. Reproducible Processes
- Different Methods for Different Questions (Complexity)

Modelling of infectiuous diseases: Pneumococcal modelling systems

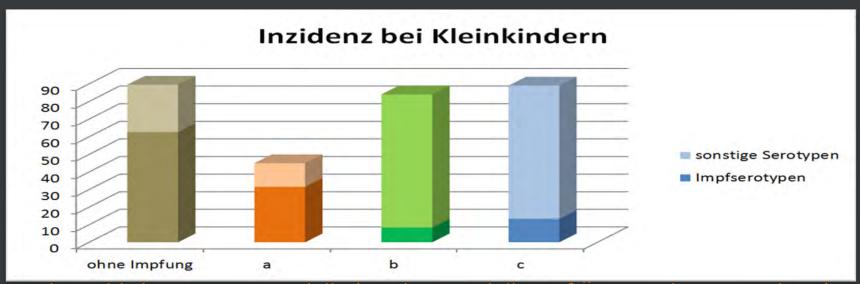


Example Comparison Vaccination



10 Concepts to Integrate

- 1. Methods to Assess and Improve Quality of Data
- 2. Potential to Integrate Missing Data
- 3. Modular & Efficient Solutions
- 4. Reproducible Processes
- Different Methods for Different Questions (Complexity)



a) Gebräuchliche Prognosemodelle (Markov-Modelle, auf Österreich umgerechnet):

E.D.G. McIntosh, P. Conway, J. Willingham, R. Hollingsworth, and A. Lloyd. The cost-burden of paediatric pneumococcal disease in the UK and the potential cost-effectiveness of prevention using 7-valent pneumococcal conjugate vaccine. Vaccine, 2003 Jun 2,21(19-20):2564-72

b) Dynamisches Pneumokokkenmodell, 2009 – Simulationsergebnisse:

u.a.: C. Urach, "Modellierung und Simulation von Impfstrategien gegen Pneumokokkenerkrankungen: Markov- und Differentialgleichungsmodelle im Vergleich" (Diploma Thesis, Inst. f. Analysis und Scientific Computing, Vienna University of Technology, 2009).

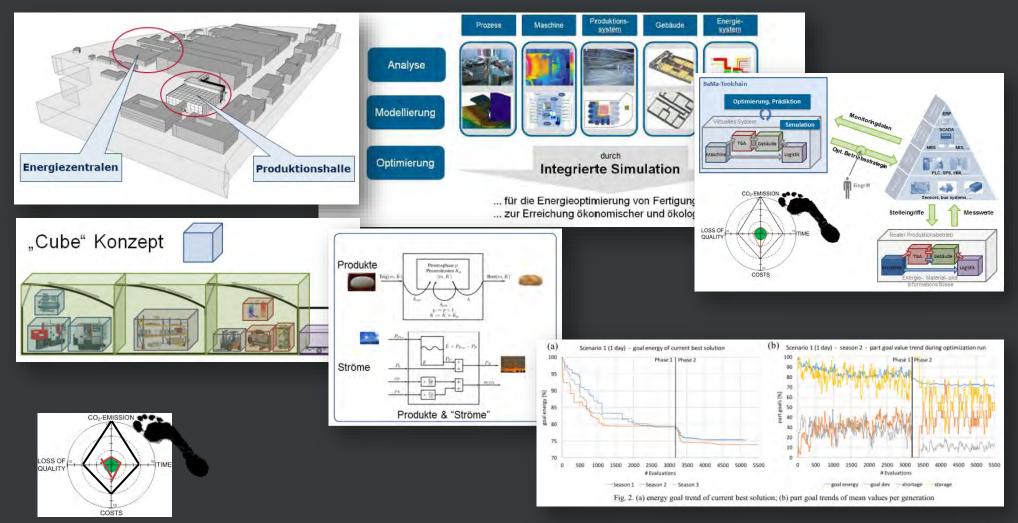
c) Erhebung aus den USA, 2010 (auf Österreich umgerechnet):

Hsu KK et al. Changing serotypes causing childhood invasive pneumococcal disease: Massachusetts, 2001–2007. Pediatr Infect Dis J 2010 Apr; 29:289

Example Combination Balanced Manufacturing



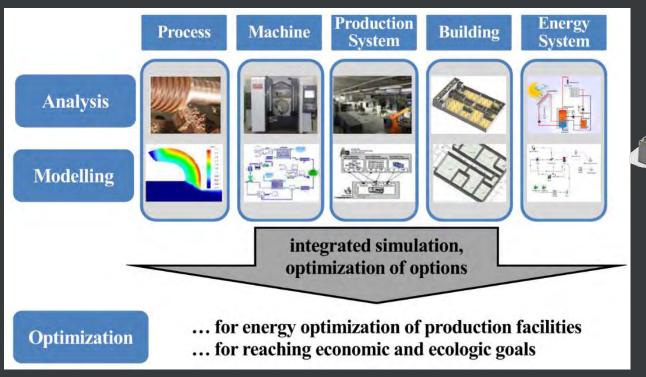
- Methods to Assess and Improve Quality of Data
- 2. Potential to Integrate Missing Data
- 3. Modular & Efficient Solutions
- 4. Reproducible Processes
- Different Methods for Different Questions (Complexity)

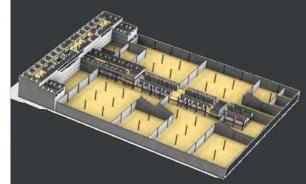


Example Combination Balanced Manufacturing



- Methods to Assess and Improve Quality of Data
- Potential to IntegrateMissing Data
- 3. Modular & Efficient Solutions
- 4. Reproducible Processes
- Different Methods for Different Questions (Complexity)





- Coupled simulation of the overall system enables holistic view of the energy distribution throughout the system.
- Find **optimization** approaches by simulating and comparing different scenarios.



10 Concepts to Integrate

- 1. Methods to Assess and Improve Quality of Data
- Potential to IntegrateMissing Data
- 3. Modular & Efficient Solutions
- 4. Reproducible Processes
- Different Methods for Different Questions (Complexity)
- 6. Comparability of Results

Comparability of Models and Results

- Qualitative Comparison
- Quantitative Comparison including
 Parameter Transformation
- Showing Limitations of Modelling Approaches and Implementation

METHOD: Comparative Modelling

Example: Calibration



10 Concepts to Integrate

- Methods to Assess and Improve Quality of Data
- Potential to Integrate Missing Data
- Modular & Efficient Solutions
- 4. Reproducible Processes
- Different Methods for Different Questions (Complexity)
- 6. Comparability of Results

Calibration algorithms require thousands of simulation runs

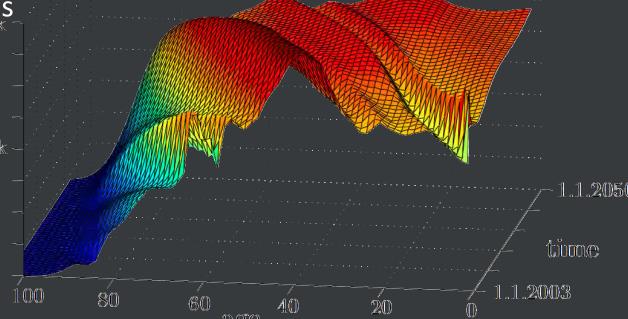
Problematic



Each simulation run contains random elements and takes some time

Formal model analysis

$$\frac{\partial M}{\partial t} - \frac{\partial M}{\partial a} = \alpha_1 + \gamma_1 \mathbb{1}_{[0,dt)}(a)\Psi(a,t) - M\delta_1$$
$$\frac{\partial F}{\partial t} - \frac{\partial F}{\partial a} = \alpha_1 + \gamma_2 \mathbb{1}_{[0,dt)}(a)\Psi(a,t) - F\delta_2$$

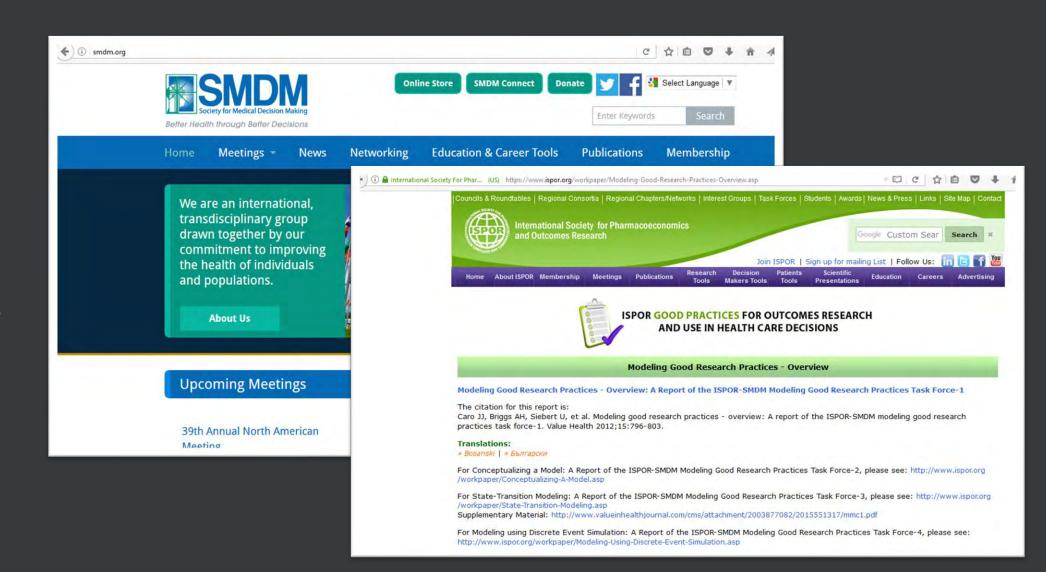


M. Bicher, N. Popper, "Mean-Field Approximation of a Microscopic Population Model for Austria " published Eurosim 2016

Cross Model Validation



- 1. Methods to Assess and Improve Quality of Data
- Potential to Integrate Missing Data
- 3. Modular & Efficient Solutions
- 4. Reproducible Processes
- Different Methods for Different Questions (Complexity)
- 6. Comparability of Results



Exact Comparable Definitions of Models. E.g. Formal Definition of a Cellular Automaton



Cell

Definition

The set of all cells will be denoted M. Consequently a cell is an unique (A set contains by definition unique elements.) element $m \in M$.

There is no limitation in the number of cells. However if only one cell is observed, we call the cellular automaton a trivial cellular automaton.

Special Case

We may call a cellular automaton with a finite number of cells finite or otherwise infinite.

Lattice

Definition

If M is a set of cells, we call M indexed or regularly arranged if there exists a bijective mapping

arrange of these exists a objective mapping $\mathcal{I}: m \to I: m \to \mathcal{I}(m) = i$ between M and an index set I. We call I an index mapping and also use I for mapping tuples of cells onto tuples of indices $I: M^k \to I^k: (m_1, \dots, m_k) \mapsto (\mathcal{I}(m_1), \dots, \mathcal{I}(m_k)) = (i_1, \dots, i_k)$ where $k \in \mathbb{N} \setminus \{0\}$.

 $(m_{\mathcal{I}(m)})_{m\in M}$ is the natural indexing for a set of indexed cells

Definition

Since by now all cells are "arranged" or indexed using an index set $I \subset \mathbb{Z}^d$, we call d the **dimension** of the cellular automaton.

Definitio

A subset I of \mathbb{Z}^d is called **connected** if for each two elements $a,b\in I$ there exists a series of elements $(z_\alpha)_{\alpha\in\mathbb{N}}\subset I$ for which $\|z_\alpha-z_{\alpha+1}\|=1\ \forall \alpha\in\mathbb{N}$ and for which $a,b\in(z_\alpha)_{\alpha\in\mathbb{N}}$.

Definition

We call a connected subset $I \subseteq \mathbb{Z}^d$, $d \in \mathbb{N} \setminus \{0\}$ an index set.

Special Case

Despite requiring a connected index set, this definition does not exclude exotic or even absurd index sets (i.e. lattices). (TODO)

G. Schneckenreither, Thesis: "Developing Mathematical Formalisms for Cellular Automata in Modelling and Simulation"; 2014

Neighbours

Let $k \in \mathbb{N} \setminus \{0\}$.

Definition

A tuple $J:=(j_1,\ldots,j_k)\in(\mathbb{Z}^d)^k$ where $j_o\neq j_\beta$ is called a **relative index tuple** and for $i\in\mathbb{Z}^d$ the addition respectively subtraction $J\pm i:=(j_1\pm i,\ldots,j_k\pm i)$ is well-defined.

Definition

Given a relative index tuple J we define the index translation \mathcal{T}_J of an index i by $\mathcal{T}_J:I\to (\mathbb{Z}^d)^k:i\mapsto i+J$ and call the result an absolute index tuple.

Note that i + J is not necessarily a subset of I.

heorem

I is a vector space? (TODO)

Definition

For a cell m_i from an indexed set of cells with index set I and a relative index tuple J we use the resulting absolute index tuple $T_J(i) = (i_1,\ldots,i_k) \in (\mathbb{Z}^d)^k$, to define the neighbourhood of m_i as $N_{m_i,J} := (n_1,\ldots,n_k) \in (M \cup \{\emptyset\})^k$ where

$$n_{\alpha} := \begin{cases} m_{i_{\alpha}} = \mathcal{I}^{-1}(i_{\alpha}) & i_{\alpha} \in I \\ \emptyset & i_{\alpha} \notin I \end{cases}$$
 $\alpha \in \{1, ..., k\}$

Furthermore we call k the neighbourhood dimension.

The non-existant cell \emptyset is required in order to maintain the original tuple structure of the neighbourhood and to be able to indicate that indices which are outside the index set do not refer to a cell.

Special Case

- A cell lies in its own neighbourhood (reflexive) if and only if 0 ∈ Z^d is part of the relative index tuple.
- An index tuple respectively neighbourhood is neither neccessarily symmetric, bidirectional nor local.

heorem

It is not unusual that neighbourhoods are of local character, which means that the neighbourhood relation is defined by the distance between cells. In this case, the index translation can be replaced by the appropriate function $T_{\rm mextic}$. It is however always possible to find an equivalent relative index tuple J and use T_J .

Definitio

For an indexed set of cells $(M,I,\mathcal{I},\mathcal{I}^{-1})$ and an index translation \mathcal{T} the neighbourhood mapping is defined by $\mathcal{N}: \mathcal{I}^{-1} \circ \mathcal{T} \circ \mathcal{I} : \mathcal{M} \to I \to \mathcal{Z}^k \to (M \cup \{\emptyset\})^k : n_t \mapsto i \mapsto (i_1,\dots,i_k) \mapsto (n_1,\dots,n_k).$

G. Schneckenreither, N. Popper, F. Breitenecker: "Methods for Cellular Automata and Evolution Systems in Modelling and Simulation"; IFAC PapersOnLine, 48 (2015), 1; S. 1 - 944.

Border

Two types of "special cells" can be destinguised:

Definition

A border-cell is a cell, which is located at the boundary of the lattice.

Definition

If the absolute index tuple of a cell m_i does not lie completely within the index set $(\mathcal{T}_I(i) \nsubseteq I)$, we talk of (a cell with) a degraded neighbourhood.

Optional boundary conditions must be applied to cells with a degraded neighbourhood!

In order to manipulate the geometry of the lattice (periodic boundary conditions for example) we modify the index translation:

Definition

Given a relative index tuple J, the generalised index translation is defined by $T_J: I \to I^k: i \mapsto (i_1, \dots, i_k)$ where

$$\alpha := \begin{cases} i+j_{\alpha} & i+j_{\alpha} \in I \\ \tau(i+j_{\alpha}) & i+j_{\alpha} \notin I \end{cases}$$
 $\alpha \in \{1,...,k\}$

and $\tau: \mathbb{Z}^d \setminus I \to I$.

Special Case

What kind of preconditions to or characteristics of τ generate which type geometry? (TODO)

- a toroid geometry for a two-dimensional cellular automaton (coll. "periodic boundary condition" a) can be achieved by using the modulus function...
- the same is true for a cylindrical geometry...
- what about a spherical surface?

actually in this situation ther is no boundary!

State

Let $k \in \mathbb{N} \setminus \{0\}$ be the neighbourhood dimension.

Definitio

There exists a (temporary) state mapping from the set of all cells M to the set of all possible states $\mathbb S$, which assigns a state to each cell $S:M\to\mathbb S:m\to S(m)=s$. We also use S as $S:(M\cup\{\emptyset\})^k\to(\mathbb S\cup\{\emptyset\})^k:(m_1,\ldots,m_k)\mapsto(s_1,\ldots,s_k)$ where

$$s_{\alpha} := \left\{ \begin{array}{ll} \mathcal{S}(m_{\alpha}) & m_{\alpha} \in M \\ \emptyset & m_{\alpha} \notin M \Longleftrightarrow m_{\alpha} = \emptyset \end{array} \right. \qquad \alpha \in \{1, \dots, k\}$$

The non-existent state \(\text{)} is required to maintaine the tuple structure and to indicate a degraded neighbourhood.

Update

Let $k \in \mathbb{N} \setminus \{0\}$.

Definition

An update rule (also update rules) is a mapping $\mathcal{F}: (\mathbb{S} \cup \{\emptyset\})^k \to \mathbb{S}: (s_1, \dots, s_k) \mapsto s$

To calculate a new state for a cell:

 $M \to (M \cup \{\emptyset\})^k \to (\mathbb{S} \cup \{\emptyset\})^k \to \mathbb{S}:$ $m \mapsto (m_1, \dots, m_k) \mapsto (s_1, \dots, s_k) \mapsto s$

actually, in detail: $F \circ S \circ T^{-1} \circ T \circ T \circ T$: $M \to I \to \mathbb{Z}^k \to (M \cup \{\emptyset\})^k \to (\mathbb{S} \cup \{\emptyset\})^k \to \mathbb{S} :$ $m \mapsto i \mapsto (i_1, \dots, i_k) \mapsto (m_1, \dots, m_k) \mapsto (s_1, \dots, s_k) \mapsto s_k$

Remark

- An update rule (update rule set) can be the explicit definition of a mapping but also a (continuous) function or a combination of functions.
- Since degraded neighbourhoods contain non-existent cells respectively states (Ø), an update rule must react on a degraded neighbourhood and implement (arbitrary) "boundary conditions".

Warning

- An update rule never defines the geometry of the lattice!
- We exclude stochastic update rules from a basic definition since the neccessary introduction of a probability space would be an extension to our formal definition (TODO).

Definition

An update rule must be definied for every possible neighbourhood configuration. Otherwise we deal with undefined behaviour.

- An update rule must be compatible with the index set. All occuring degradations of neighbourhoods must be taken into account.
- An update rule must be self-contained since all possible state-configurations^a arrise from the update rule.

"except for the initial condition, see later

Special Case

- The set of all possible states S may contain a finite or infinite number of states
- A "state-space" (coll.) can be a vector space, a ring or any other algebraic structure.
- By introducing a partitioning on the set of all possible states, different cell types can be distinguished.

A non-trivial cellular automata features more than one different element in $\mathbb{S}. \label{eq:section}$

Global State

efinition

The state of all cells is accumulated in the temporary state mapping S which can be identified with an element of $\mathfrak{S} := \mathbb{S}^M$. We then also call S the (temporary) global state.

Definitio

Given a neighbourhood mapping \mathcal{N} , a temporary state mapping \mathcal{S} and an update rule \mathcal{F} we define the local evolution operator $\mathcal{S} := \mathcal{F} \circ \mathcal{S} \circ \mathcal{N} : M \to (M \cup \{\emptyset\})^k \to \mathbb{S} : m \mapsto (m_1, \ldots, m_k) \mapsto (s_1, \ldots, s_k) \mapsto s.$

We can see that a local evolution operator is a state mapping and a global state.

Iteration

Definition

A (global) evolution operator is a mapping $\mathcal{E}: \mathfrak{S} \to \mathfrak{S}: \mathcal{S} \mapsto \tilde{\mathcal{S}} := \mathcal{F} \circ \mathcal{S} \circ \mathcal{N}$.

Definition

An iterative process can be obtained by defining $S_{t+1} := \mathcal{E}(S_t) = \mathcal{F} \circ S_t \circ \mathcal{N}$ where $n \in \mathbb{N}$.

Definition

For an iteration, an initial state or initial condition S_0 must be given. It is neccessary that the initial condition is compatible with the update rules.

Only the states of the cells and accordingly the state mapping may change during iteration!

Automaton

Definition

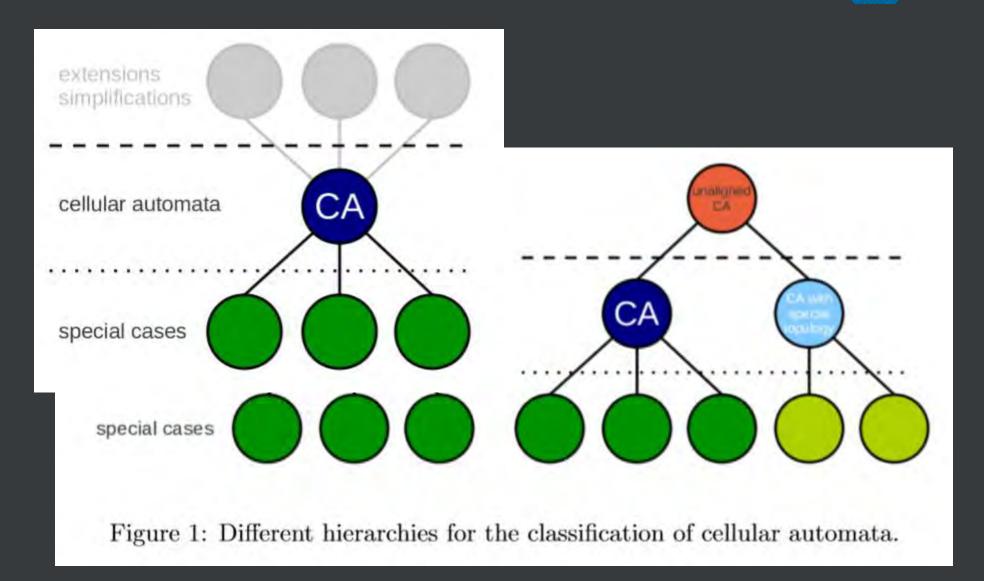
A cellular automaton comprises ...

- (i) an indexed set of cells
- (ii) an initial state mapping S_0
- (iii) a ralative index tuple J
 (iv) a generalised index translation (only if the geometry shall be "manipulated")
- (v) an update rule F
- (vi) the iterative application of an evolution operator

Exact Comparable Definitions of Models. E.g. Formal Definition of a Cellular Automaton



- Methods to Assess and Improve Quality of Data
- Potential to IntegrateMissing Data
- 3. Modular & Efficient Solutions
- 4. Reproducible Processes
- Different Methods for Different Questions (Complexity)
- 6. Comparability of Results





10 Concepts to Integrate

- 1. Methods to Assess and Improve Quality of Data
- Potential to Integrate Missing Data
- 3. Modular & Efficient Solutions
- 4. Reproducible Processes
- Different Methods for Different Questions (Complexity)
- 6. Comparability of Results
- 7. Make it Understandable

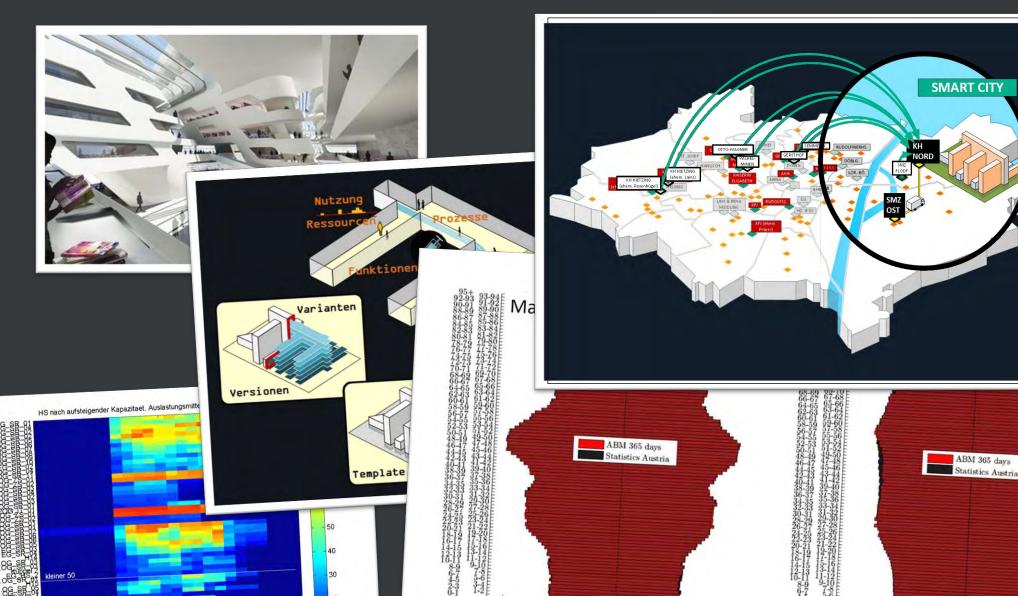
Complicated Models and Various Results

- Good data & an "improvable" Model
- Right Model & good Results face
- Nobody understands.
- -> Change Management & Interdisciplinarity

METHOD: Data Representation & Human Computer Interfaces (HCI)



- 1. Methods to Assess and Improve Quality of Data
- Potential to IntegrateMissing Data
- 3. Modular & Efficient Solutions
- 4. Reproducible Processes
- Different Methods for Different Questions (Complexity)
- 6. Comparability of Results
- 7. Make it Understandable



Having Fun with Epidemics and Herd Immunity...



10 Concepts to Integrate

- Methods to Assess and Improve Quality of Data
- 2. Potential to Integrate Missing Data
- 3. Modular & Efficient Solutions
- 4. Reproducible Processes
- Different Methods for Different Questions (Complexity)
- 6. Comparability of Results
- 7. Make it Understandable



"On the benefits of explaining herd immunity in vaccine advocacy", Nature Human Behaviour, 6.3.2017



10 Concepts to Integrate

- Methods to Assess and Improve Quality of Data
- Potential to IntegrateMissing Data
- 3. Modular & Efficient Solutions
- 4. Reproducible Processes
- 5. Different Methods for Different Questions (Complexity)
- 6. Comparability of Results
- 7. Make it Understandable
- Open and Independent Solutions

Open and Independent Solutions

- Publication limited because of Economic
 - Exploitation & Stakeholder Interests
- Lack of Comparability of Different Models
- Rules and Guidelines needed

METHOD: Open Access & Public Domain

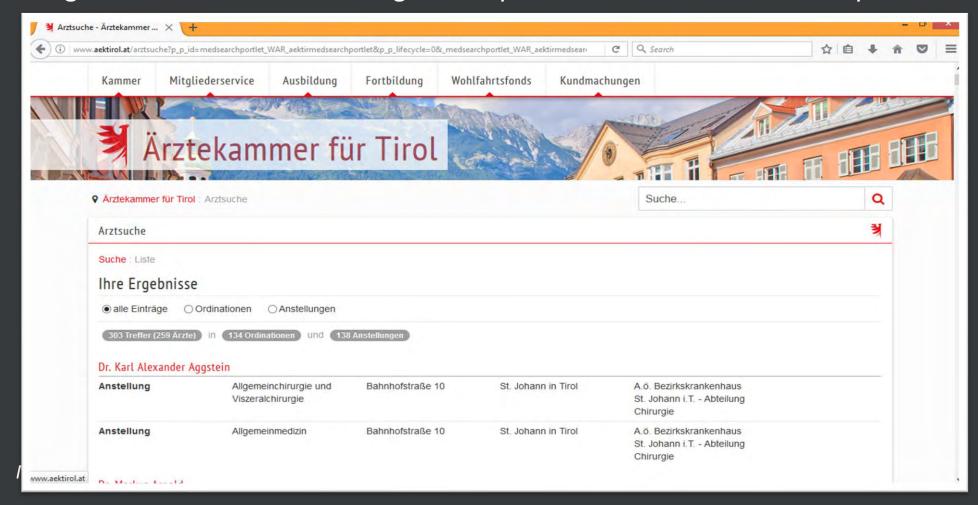
Additional Sources: Example Web Scraping



10 Concepts to Integrate

- 1. Methods to Assess and Improve Quality of Data
- Potential to Integrate Missing Data
- 3. Modular & Efficient Solutions
- 4. Reproducible Processes
- Different Methods for Different Questions (Complexity)
- 6. Comparability of Results
- 7. Make it Understandable
- Open and Independent Solutions

Integration of Data from Web Pages of Physicians – Actualisation monthly



Health Care: Physicians



- Methods to Assess and Improve Quality of Data
- Potential to IntegrateMissing Data
- 3. Modular & Efficient Solutions
- 4. Reproducible Processes
- Different Methods for Different Questions (Complexity)
- 6. Comparability of Results
- 7. Make it Understandable
- 8. Open and Independent Solutions

- Physicians total: 16.500
- Medical Practices: 20.219
- Medical Practices with GPS Coordinates: 13.607
- Total Hours per Week: 193.899 h
- GPs:
 - Medical Practices: 8.062
 - Contracts with Social Incurances: 3841
 - Total Hours per Week: 89.594 h





10 Concepts to Integrate

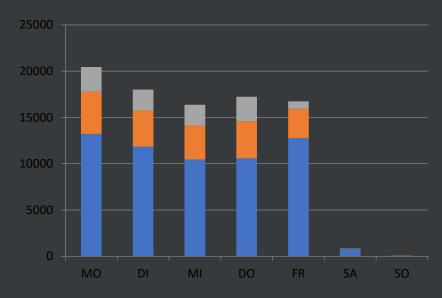
- 1. Methods to Assess and Improve Quality of Data
- Potential to IntegrateMissing Data
- 3. Modular & Efficient Solutions
- 4. Reproducible Processes
- Different Methods for Different Questions (Complexity)
- 6. Comparability of Results
- 7. Make it Understandable
- 8. Open and Independent Solutions

Example Medical Practices Hours:





absolut





10 Concepts to Integrate

- Methods to Assess and Improve Quality of Data
- Potential to IntegrateMissing Data
- 3. Modular & Efficient Solutions
- 4. Reproducible Processes
- Different Methods for Different Questions (Complexity)
- 6. Comparability of Results
- 7. Make it Understandable
- Open and Independent Solutions
- 9. Priority for Data Security and Stake Holder Interests

Priority for Data Security and Stakholders

- Da Additional Requirements e.g. with the new
 - EU General Data Protection Regulation
- Quality can be reduced by missing data
- Transparent Processes are Needed in

advance

Data & Processes

METHOD: Data Security & Governance

DEXHELPP Research Server



10 Concepts to Integrate

- Methods to Assess and Improve Quality of Data
- Potential to IntegrateMissing Data
- 3. Modular & Efficient Solutions
- 4. Reproducible Processes
- 5. Different Methods for Different Questions (Complexity)
- 6. Comparability of Results
- 7. Make it Understandable
- 8. Open and Independent Solutions
- 9. Priority for Data Security and Stake Holder Interests



- 1. Data: Temporary and restricted to well defined Research
- 2. Access: Restricted and Reproducible
- 3. Methods: Usage of all Modelling & Simulation Methods
- 4. Export: Well defined Rules for Export and Usage

Processes according to Guidelines of Research Data Alliance (RDA)



10 Concepts to Integrate

- Methods to Assess and Improve Quality of Data
- Potential to IntegrateMissing Data
- 3. Modular & Efficient Solutions
- 4. Reproducible Processes
- Different Methods for Different Questions (Complexity)
- 6. Comparability of Results
- 7. Make it Understandable
- Open and Independent Solutions
- 9. Priority for Data Security and Stake Holder Interests
- 10. Broad Applications (Health System, Energy, Industry, Energy, Mobility, Infrastructure)

Broad Applications for Models

- Re-Use of Models necessary, because of...
- ...Ressources & Quality
- Research Questions become Guidelines interdisciplinary and "communicate"

METHOD: Co-Simulation, Multi Method

Guidelines

Modelling



10 Concepts to Integrate

- 1. Methods to Assess and Improve Quality of Data
- Potential to Integrate Missing Data
- 3. Modular & Efficient Solutions
- 4. Reproducible Processes
- Different Methods for Different Questions (Complexity)
- 6. Comparability of Results
- 7. Make it Understandable
- 8. Open and Independent Solutions
- 9. Priority for Data Security and Stake Holder Interests
- 10. Broad Applications (Health System, Energy, Industry, Energy, Mobility, Infrastructure)

Exploring Advantages of Multi-Method Modelling and its applications in large socio-technical infrastructure systems Glock et al, ASIM 2016





10 Concepts to Integrate

- Methods to Assess and Improve Quality of Data
- Potential to Integrate Missing Data
- 3. Modular & Efficient Solutions
- 4. Reproducible Processes
- Different Methods for Different Questions (Complexity)
- 6. Comparability of Results
- 7. Make it Understandable
- 8. Open and Independent Solutions
- Priority for Data Security and Stake Holder Interests
- 10. Broad Applications (Health System, Energy, Industry, Energy, Mobility, Infrastructure)

The airport city - A large complex socio-technical system

Technical Infrastructure Energy Telecom Water, Waste Sewage

Landside Access Access Roads Railway Access Landside Transportation Internal Road System

Terminals Passenger terminals Air Cargo Terminals & Warehouses Air Mail Facilities

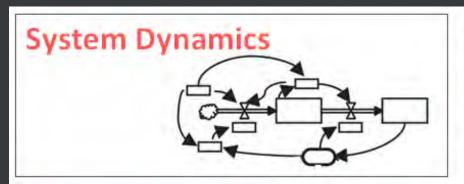
Airside & Access Airport Control Zones ATC **Facilities** Aprons, Airside Road System Safety & Security Facilities

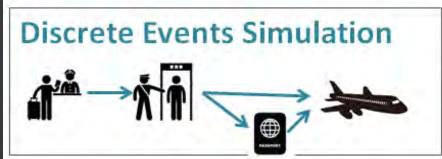
International Civil Aviation Organization: **Growth Rate** Passenger Transport increased (2012) from 4.9% to 6.3% (2015)!!

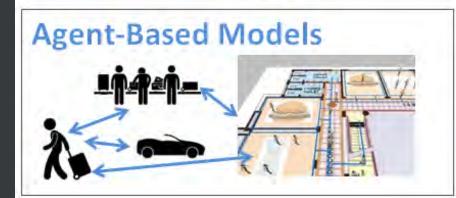


10 Concepts to Integrate

- Methods to Assess and Improve Quality of Data
- Potential to IntegrateMissing Data
- 3. Modular & Efficient Solutions
- 4. Reproducible Processes
- Different Methods for Different Questions (Complexity)
- 6. Comparability of Results
- 7. Make it Understandable
- 8. Open and Independent Solutions
- 9. Priority for Data Security and Stake Holder Interests
- 10. Broad Applications (Health System, Energy, Industry, Energy, Mobility, Infrastructure)







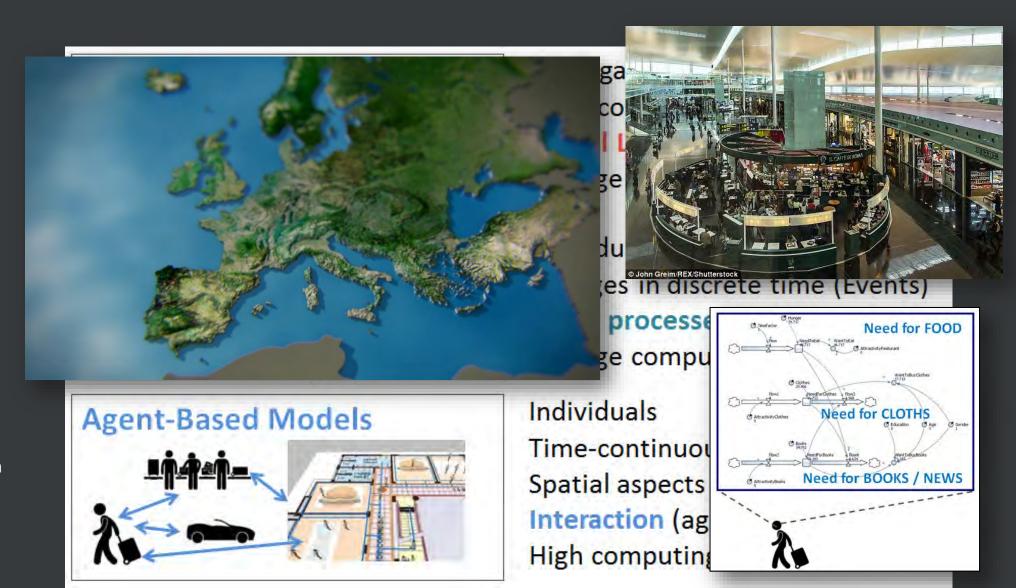
Aggregated states
Time-continuous
Causal Links & Feedback
average computing times

Individuals (Entities) & Resources
Changes in discrete time (Events)
Given processes
average computing times

Individuals
Time-continuous
Spatial aspects
Interaction (agents, environment)
High computing times



- Methods to Assess and Improve Quality of Data
- Potential to Integrate Missing Data
- 3. Modular & Efficient Solutions
- 4. Reproducible Processes
- Different Methods for Different Questions (Complexity)
- 6. Comparability of Results
- 7. Make it Understandable
- 8. Open and Independent Solutions
- 9. Priority for Data Security and Stake Holder Interests
- 10. Broad Applications (Health System, Energy, Industry, Energy, Mobility, Infrastructure)



10 Concepts to Integrate



10 Concepts to Integrate

- 1. Methods to Assess and Improve Quality of Data
- 2. Potential to Integrate Missing Data
- 3. Modular & Efficient Solutions
- Reproducible Processes
- 5. Different Methods for **Different Questions** (Complexity)
- Comparability of Results
- Make it Understandable
- Open and Independent Solutions
- 9. Priority for Data Security and Stake Holder Interests
- 10. Broad Applications (Health System, Energy, Industry, Energy, Mobility, Infrastructure)

Methods & Concepts to be Developed and Integrated

- Explorative Visual Computing –Visual Analytics and Statistics
- Data Processes (Integration & Linkage) & Modelling Tools (Parametrization & Calibration)
- Modular Models, Coupling of Models
- Validation & Data Citation
- Methods for Choosing Models
- Comparative Modelling
- Data Representation & Human Computer Interfaces (HCI
- Open Access & Public Domain
- Data Security & Governance
- Co-Simulation, Multi Method Modelling

Guidelines

Interface

Guidelines

Implementation DEXHELPP

Formal & Technological

Processes

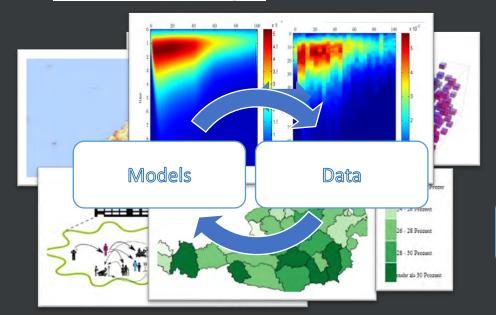


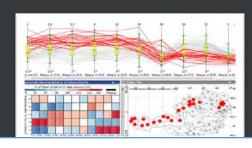
10 Concepts to Integrate

- Methods to Assess and Improve Quality of Data
- Potential to Integrate Missing Data
- 3. Modular & Efficient Solutions
- 4. Reproducible Process
- Different Methods fo Different Questions (Complexity)
- 6. Comparability of Resu
- 7. Make it Understanda
- 8. Open and Independence
 Solutions
- Priority for Data Security and Stake Holder Interests
- 10. Broad Applications (Health System, Energy, Industry, Energy, Mobility, Infrastructure)









Decision Support & Process Integration (e.g.: HTA)

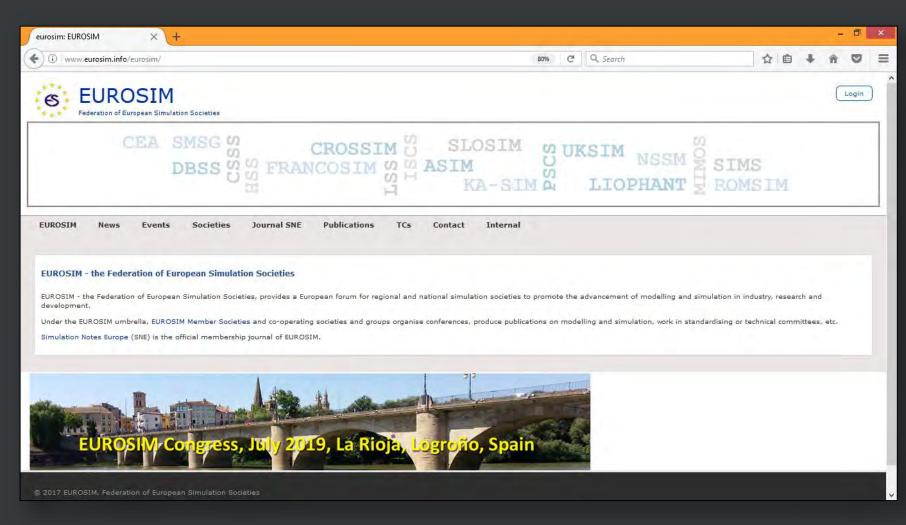


Infrastructur
DEXHELPP Server & Services

EUROSIM Initiative DDSS Technical Committee



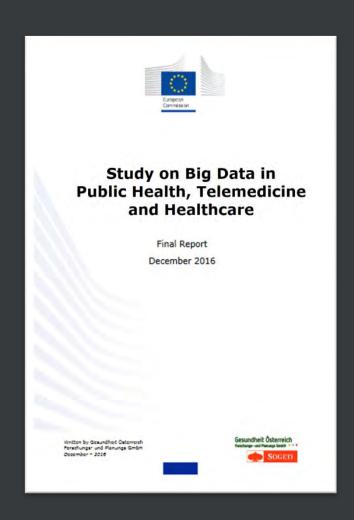
- Methods to Assess and Improve Quality of Data
- 2. Potential to Integrate Missing Data
- 3. Modular & Efficient Solutions
- 4. Reproducible Processes
- Different Methods for Different Questions (Complexity)
- 6. Comparability of Results
- 7. Make it Understandable
- 8. Open and Independent Solutions
- 9. Priority for Data Security and Stake Holder Interests
- 10. Broad Applications (Health System, Energy, Industry, Energy, Mobility, Infrastructure)

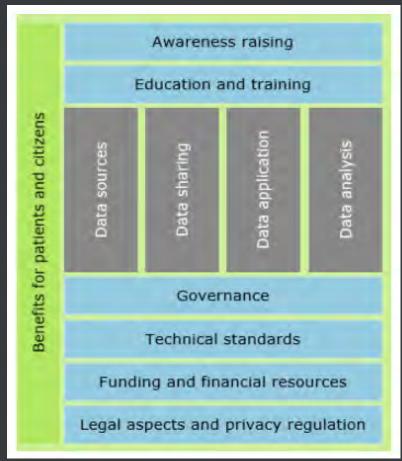




10 Concepts to Integrate

- Methods to Assess and Improve Quality of Data
- Potential to Integrate Missing Data
- 3. Modular & Efficient Solutions
- 4. Reproducible Processes
- Different Methods for Different Questions (Complexity)
- 6. Comparability of Results
- 7. Make it Understandable
- Open and Independent Solutions
- Priority for Data Security and Stake Holder Interests
- 10. Broad Applications (Health System, Energy, Industry, Energy, Mobility, Infrastructure)





Overview of fields of policy recommendations

https://ec.europa.eu/health/sites/health/files/ehealth/docs/bigdata_report_en.pdf



10 Concepts to Integrate

- Methods to Assess and Improve Quality of Data
- 2. Potential to Integrate Missing Data
- 3. Modular & Efficient Solutions
- 4. Reproducible Processes
- Different Methods for Different Questions (Complexity)
- 6. Comparability of Results
- 7. Make it Understandable
- 8. Open and Independent Solutions
- 9. Priority for Data Security and Stake Holder Interests
- 10. Broad Applications (Health System, Energy, Industry, Energy, Mobility, Infrastructure)

The following twenty examples of use of Big Data in Health were identified and selected for further analysis:

- Comet K-Project DEXHELPP AT
- The Shared Care Platform DK
- E-Estonia National Identity Scheme EE
- AEGLE (An analytics framework for integrated and personalized healthcare services in Europe) – UK, IT, GR, SE, BE, NL, PT, FR
- The Business Intelligence database system GR
- PASSI (Progressi delle Aziende Sanitarie) IT
- Arno Observatory IT
- The Swedish Big Data Analytic Network SE
- Clinical Practice Research Datalink (CPRD) UK
- Sentinel Stroke National Audit Programme (SSNAP) UK
- Hospital Episode Statistics (HES) UK (England)
- The YODA Project (Yale University open data access) US
- FDA Adverse Event Network Analyser US
- CEPHOS-LINK FI, AT, RO, NO, SI, IT
- Twitter (Adverse drug reactions and public health) International
- Flatiron US
- UK Biobank UK
- Semantic Data Platform for Healthcare (SEMCARE) DE, NL, AT, UK, ES
- Integrated BioBank of Luxembourg (IBBL) LU
- Spanish Rare Diseases Registries Research Network (SpainRDR) ES



10 Concepts to Integrate

- 1. Methods to Assess and Improve Quality of Data
- Potential to Integrate Missing Data
- 3. Modular & Efficient Solutions
- 4. Reproducible Processes
- Different Methods for Different Questions (Complexity)
- 6. Comparability of Results
- 7. Make it Understandable
- 8. Open and Independent Solutions
- 9. Priority for Data Security and Stake Holder Interests
- 10. Broad Applications (Health System, Energy, Industry, Energy, Mobility, Infrastructure)

Recommendations

- Recommendation 1 on Awareness Raising:
 Develop and implement a communication strategy to increase the awareness of the added value of Big Data in Health and encourage a positive public mind set towards Big Data in Health
- Recommendation 2 on Education and Training
 Strengthen human capital with respect to the increasing need for a workforce that can utilize the potential of Big Data in Health
- Recommendation 3 on Data Sources:
 Expand existing and explore new sources of Big Data in Health and secure their quality and safety
- Recommendation 4 on Open Data and Data Sharing:
 Promote open use and sharing of Big Data in Health without compromising patients' rights to privacy and confidentiality
- Recommendation 5 on Applications and Purposes:
 Increase target-oriented application of Big Data analysis in health based on the needs and interests of stakeholders including patients



10 Concepts to Integrate

- 1. Methods to Assess and Improve Quality of Data
- Potential to Integrate Missing Data
- 3. Modular & Efficient Solutions
- 4. Reproducible Processes
- Different Methods for Different Questions (Complexity)
- 6. Comparability of Results
- 7. Make it Understandable
- 8. Open and Independent Solutions
- 9. Priority for Data Security and Stake Holder Interests
- 10. Broad Applications (Health System, Energy, Industry, Energy, Mobility, Infrastructure)

Recommendations

- Recommendation 6 on Data Analysis:
 Identify the potentials of Big Data analysis, improve analytical methods and facilitate the use of new and innovative analytical methods
- Recommendation 7 on Governance of Data Access and Use:
 Implement governance mechanisms to ensure secure and fair access and use of Big Data for research in health
- Recommendation 8 on Standards:
 Develop standards for Big Data in Health to enhance and simplify its application and improve interoperability
- Recommendation 9 on Funding and Financial Resources:
 Ensure purposeful investment steered by the European Commission to warrant cost-effectiveness and sustainability
- Recommendation 10 on Legal Aspects and Privacy Regulations:
 Clarify and align existing legal and privacy regulation of Big Data in Health

10 Concepts to Integrate



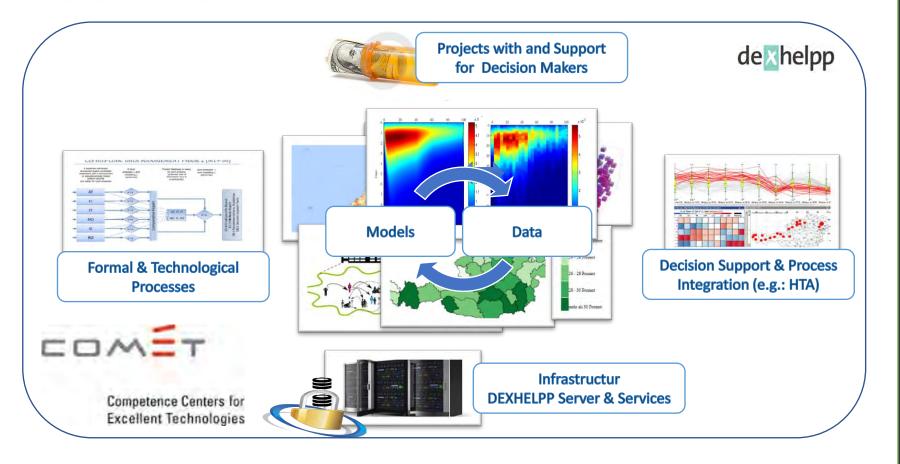
10 Concepts to Integrate

- Methods to Assess and Improve Quality of Data
- Potential to Integrate Missing Data
- 3. Modular & Efficient Solutions
- 4. Reproducible Processes
- Different Methods for Different Questions (Complexity)
- 6. Comparability of Results
- 7. Make it Understandable
- Open and Independent Solutions
- 9. Priority for Data Security and Stake Holder Interests
- 10. Broad Applications (Health System, Energy, Industry, Energy, Mobility, Infrastructure)

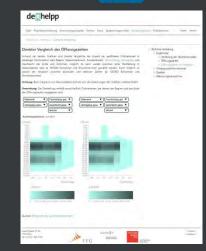
Outlook

- Putting it all together a huge challenge, but necessary because of external reasons
- A new "Process for Processes" instead of "Modelling & Simulation" integrating all areas of research (4th Paradigm of Jim Gray)
- Not only for models, but also for the modelling process: not static
 anymore -> dynamic approaches also for smaller models
- Huge methodolical Differences between "Data Driven" and "Model Driven" Approaches – e.g. Data Driven Journalism
- No "Generic World Model" possible (e.g. Co-Morbidities), but there are Processes to generate controllable complex and dynamic models

deXhelpp









ISPOR Worldwide Chapters



ISPOR - The Professional Society for Health Economics and Outcomes Research is a multistakeholder organization with more than 20,000 individual and chapter members from 110+ countries worldwide.

Upcoming Conferences and Education



ISPOR Warsaw 2019 27-28 MARCH 2019 | WARSAW | POLAND



ISPOR 2019 MAY 18-22, 2019 | NEW ORLEANS, LA | USA



ISPOR Latin America 2019 12-14 SEPTEMBER 2019 | BOGOTÁ | COLOMBIA



ISPOR Europe 2019 2-6 NOVEMBER 2019 | COPENHAGEN DENMARK







Dr. Niki Popper niki.popper@dwh.at

More Information on:

dexhelpp.at

dwh.at

cocos.tuwien.ac.at