CHALMERS

Building Information Models' data for machine learning systems in construction management

DIMOSTHENIS KIFOKERIS, MATTIAS ROUPÉ, MIKAEL JOHANSSON, CHRISTIAN KOCH

Budapest, Hungary

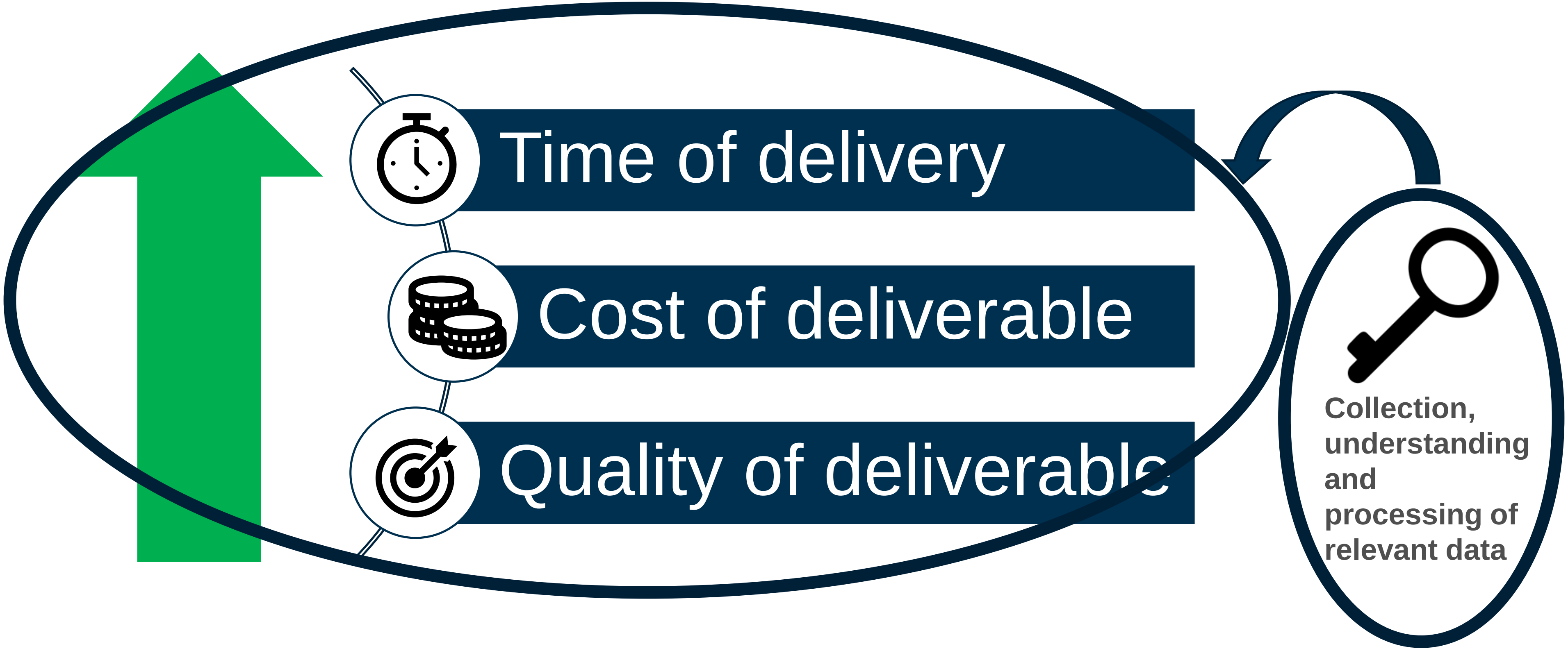
2019-07-01

CCC2019 CREATIVE CONSTRUCTION CONFERENCE

CONSTRUCTION MANAGEMENT

- Time of delivery
- Cost of deliverable
- Quality of deliverable

CONSTRUCTION MANAGEMENT



DATA USE FOR CONSTRUCTION MANAGEMENT

- Aid and enhancement of construction managers' **decision-making** and **action-taking** <u>**regarding project performance assessment**</u>
- Production data **potentially found in information structures and standards within 5-D (or higher) BIM** and **exploited by ML:** <u>**development of an ML system that predicts project delivery time and cost, using data found in future BIM**</u>

CONSTRUCTION INFORMATICS (Turk 2006)

Among others…

- Construction informatics
  - Data mining
  - Machine learning (ML)
    - Supervised (SML)
    - Unsupervised (UML)
    - Hybrid (HML)
  - Building Information Models (BIM)

RESEARCH ON MACHINE LEARNING WITHIN CONSTRUCTION SECTOR

Early attempts:

- Computer-aided acquisition and processing of design- and construction-related expert knowledge

SML:

- Support of bidding processes
  - Analysis of contractual texts for relevant requirements
  - Building code compliance checking
- Estimation of projects or component completion time

RESEARCH ON MACHINE LEARNING WITHIN CONSTRUCTION SECTOR

SML:

- Building energy consumption prediction
- Image recognition for construction-related tasks
    - On-site surveillance issues (e.g. on-site worker movement)
    - Automated on-site detection of workers and heavy equipment
- Identification of occupational accidents
- Development of on-site safety leading indicators

RESEARCH ON MACHINE LEARNING WITHIN CONSTRUCTION SECTOR

SML:

- Recognition of workers’ stress during on-site tasks
- Detection of non-certified work on-site
- Reduction of design error occurrence using sensor data from operating building (interface with Internet of Things)
- Assessment of construction productivity

RESEARCH ON MACHINE LEARNING WITHIN CONSTRUCTION SECTOR

SML:

- Assessment of buildability during design phase
- Facilitation of construction cost related to attributes of building materials

UML:

- Aiding design integration by accounting for construction knowledge experience
- Derivation of construction project risk sources

RESEARCH ON MACHINE LEARNING WITHIN CONSTRUCTION SECTOR

HML:

- Cooling control systems in office buildings
- Integration of constructability and risk analysis

MACHINE LEARNING SYSTEMS UTILIZING 5-D BIM DATA

- Data pre-processing
  - Development of as-built BIM models
  - Textual classification of maintenance work orders for the integration of BIM with facilities management
- Quality control and code compliance
- Knowledge discovery within BIM, as part of cognitive assistance frameworks

Adapted from BIM+

MACHINE LEARNING SYSTEMS UTILIZING 5-D BIM DATA

- Systems for knowledge discovery within BIM are generally the most advanced cases of processing and exploiting BIM data – but:
  - Main focus on improvement and back-propagation of BIM models
  - Only for distinct contexts even when coupled with development of predictive systems, and **not for assessing major indicators of project performance (time, cost, and quality)**

CONSTRUCTABILITY: HOLISTIC FRAMEWORK

- To utilize investigated 5-D (or higher) BIM data for developing a ML system predicting project performance indices (esp. delivery cost and time overheads), the investigated data needs to:
  - Be translated into meaningful independent input variables
  - Be connected with meaningful dependent output variables
- **Thus, it needs to be incorporated in a suitable theoretical and conceptual framework**

CONSTRUCTABILITY: HOLISTIC FRAMEWORK

- Constructability: optimal use of construction knowledge and experience in planning, design, procurement, and field operations to achieve project objectives of time, cost and quality (CII 1986)
- Constructability problems: situations of non-proper implementation of construction knowledge and experience → widening of gap between the “as-designed” and “as-built” project states → sub-optimal objectives (Kifokeris 2018)

DATA STRUCTURES IN INDUSTRY FOUNDATIONS CLASSES (IFCs)

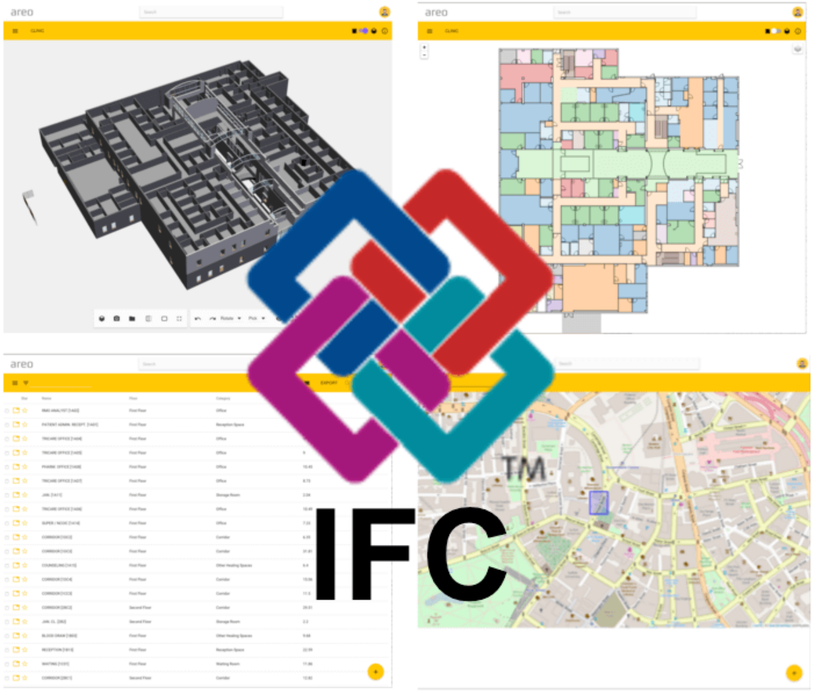
General features

- IFCs: entity-relationship data models, encompassing entities organized into object-based inheritance hierarchies
- Supported building objects: e.g. IfcWall, IfcBeam, IfcWindow, IfcRoof
    - Unlimited sets of properties and quantities can be connected to each object
- GUID: globally unique identifiers for base objects, that can be made persistent for the project
    - Thus, multiple IFCs can be merged deterministically in the project, while keeping their data integrity, without human intervention

DATA STRUCTURES IN INDUSTRY FOUNDATIONS CLASSES (IFCs)

Product-oriented structures

- Instances of IFC objects belong to spatial contexts → IFCs ≠ general 3D-file formats
- Space-enclosing structures: e.g. sites (IfcSite), buildings (IfcBuilding), rooms (IfcSpace)
- IfcRelation: feature forming constraints and relations between building parts



Adapted from AREO

DATA IN INDUSTRY FOUNDATIONS CLASSES (IFCs)

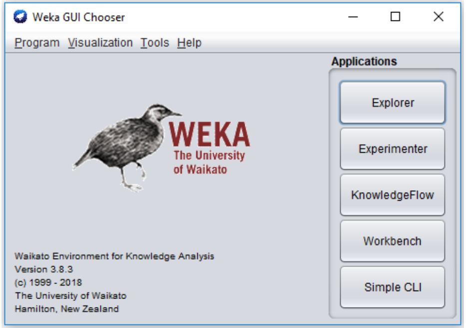
Production-oriented structure

- IfcProcess: base class for processes (e.g. tasks and events) → assigned to products to indicate the output of performed work
- IfcResource: base class for resources (e.g. materials, labor) and their associated cost- and time-related constraints → assigned to processes to indicate tasks performed on behalf of a resource

Until now, limited use and support

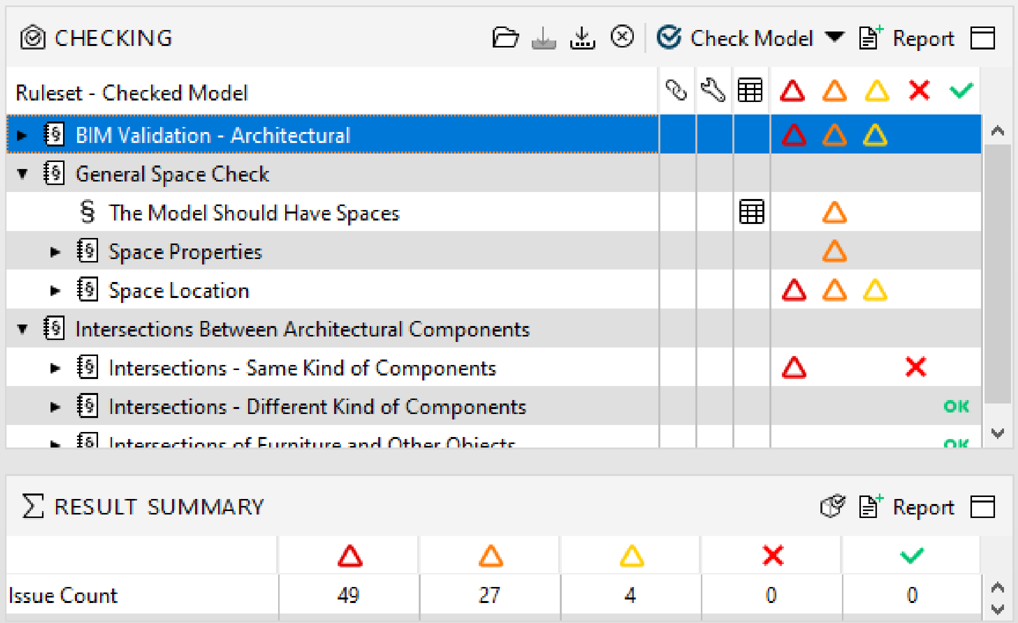
EXAMPLES OF TOOLS FOR 5-D BIM DATA INVESTIGATION

- Solibri Model Checker →
- Customized semantic and/or latent techniques for data mining
- Dedicated data parsers
- Waikato Environment for Knowledge Analysis (WEKA) →
- Surprise Scikit



CONSTRUCTABILITY: HOLISTIC FRAMEWORK

- E.g. elements potentially translated into constructability problems in IFC-ordered 5-D BIM:
  - Geometric and dimensional discrepancies
  - Detected design clashes
  - Construction site spatial and schedule clashes
  - Timeframe conflicts
  - Logistics and material quantity problems
  - Number of reworks



SYSTEM CONCEPTUALIZATION

- Prerequisites:
  - Direct connection of constructability to overall project objectives rather than narrow applications
  - Affiliation of constructability with construction knowledge and experience implementation
  - Capabilities of construction informatics technologies to extract and process data that can be interpreted as constructability problems

SYSTEM CONCEPTUALIZATION

1. DATA COLLECTION
2. VARIABLE FORMULATION
3. SYSTEM FORMULATION
4. INTEGRATION OF RESULTS

SYSTEM CONCEPTUALIZATION

- DATA COLLECTION
- VARIABLE FORMULATION
- SYSTEM FORMULATION
- INTEGRATION OF RESULTS

- **5-D BIM data of a large number of as-designed and as-built building projects**
- Quantitative
  - Components and system types
  - Manufactured products
  - Geometric and dimensional discrepancies
  - Design clashes
- Qualitative
  - Descriptions of spatial and schedule clashes
  - Timeframe conflicts
  - Logistics and material quantity problems
  - Number of reworks.
- Data will reveal constructability problems (input variables of the ML system)
- Extracted and exported into suitable file formats
- **For the same projects, expert input and documented data on the corresponding delivery cost and time overheads (the output variables of the ML system)**

SYSTEM CONCEPTUALIZATION

- **Independent variables**
  - E.g. Number of reworks
  - Measured through values of collected data
  - Produced through UML (e.g. vector quantization, linguistic clustering) **OR** qualitative techniques relying on expert input (e.g. brainstorming sessions)
- **Dependent variables**
  - “Overheads on the intended cost” & “Delay in the time of completion”
  - Discrete or continuous → used for classification or regression

SYSTEM CONCEPTUALIZATION

- **Supervised machine learning**
  - Training and validation depending on data form and amount, and variables’ type and number
  - Multiple experiments with numerous SML schemes
  - E.g. support vector machines (SVM) for binomial classification
  - E.g. support vector regression (SVR) for regression
  - E.g. variations of random forest for multinomial cases

SYSTEM CONCEPTUALIZATION

DATA COLLECTION → VARIABLE FORMULATION → SYSTEM FORMULATION → INTEGRATION OF RESULTS

- **Use of auxiliary mathematical, methodological and software tools**
  - Non-negative matrix factorization for data normalization and pre-processing (steps 1-2)
  - Multi-input Analytical Hierarchy Process (AHP), for variable labelling (step 2)
  - "Kernel trick", to aid in the non-linear function of e.g. SVM or SVR (step 3)
  - N-fold cross-validation, for simultaneous SML training and validation (step 3)
  - WEKA (step 3)
  - Surprise Scikit (steps 2-3)
  - Python (steps 2-3)

SYSTEM CONCEPTUALIZATION

1. DATA COLLECTION
2. VARIABLE FORMULATION
3. SYSTEM FORMULATION
4. INTEGRATION OF RESULTS

- **Integration as a working prototype within 5-D BIMs of new buildings**
  - Verification of its predicting results: delivery cost and time overheads of new buildings, in relation to the detected constructability problems
- Integration through programming routines and/or GUIs (e.g. PyQt)

CONCLUSIONS

- IFC-ordered data in 5-D BIMs: potentially rich and utilizable source for optimizing construction management
  - Understood and processed through the lens of constructability
  - Extracted via the relative tools and methodologies
  - Coupled with expert input
  - Used for training and validating machine learning systems predicting the delivery cost and time overheads of a building
- **Future work: realization of conceptual framework through data mining and expert processes + deployment and experimentation of ML algorithms**

CHALMERS

THANK YOU FOR YOUR ATTENTION!

Chalmers University of Technology
Division of Construction Management, Department of Architecture and Civil Engineering
Sven Hultins Gata 6, 412 96 Gothenburg

Dimosthenis Kifokeris, Ph.D.

Assistant Professor in Construction Production and Management of Construction Projects

Room 371

E-mail: dimkif@chalmers.se

Mikael Johansson, Ph.D.

Research Engineer

Room 375

E-mail: jomi@chalmers.se

Mattias Roupé, Ph.D.

Senior Lecturer

Room 375

E-mail: mattias.roupe@chalmers.se

Christian Koch, Ph.D.

Full Professor in Process Management and Innovation, Research Manager

Room 365

E-mail: christian.Koch@chalmers.se

CHALMERS