



Title: *An Interactive Digital Twin for Visual Querying and Process Mining*

Authors: Spyros Loizou and Dr. Andreas Andreou

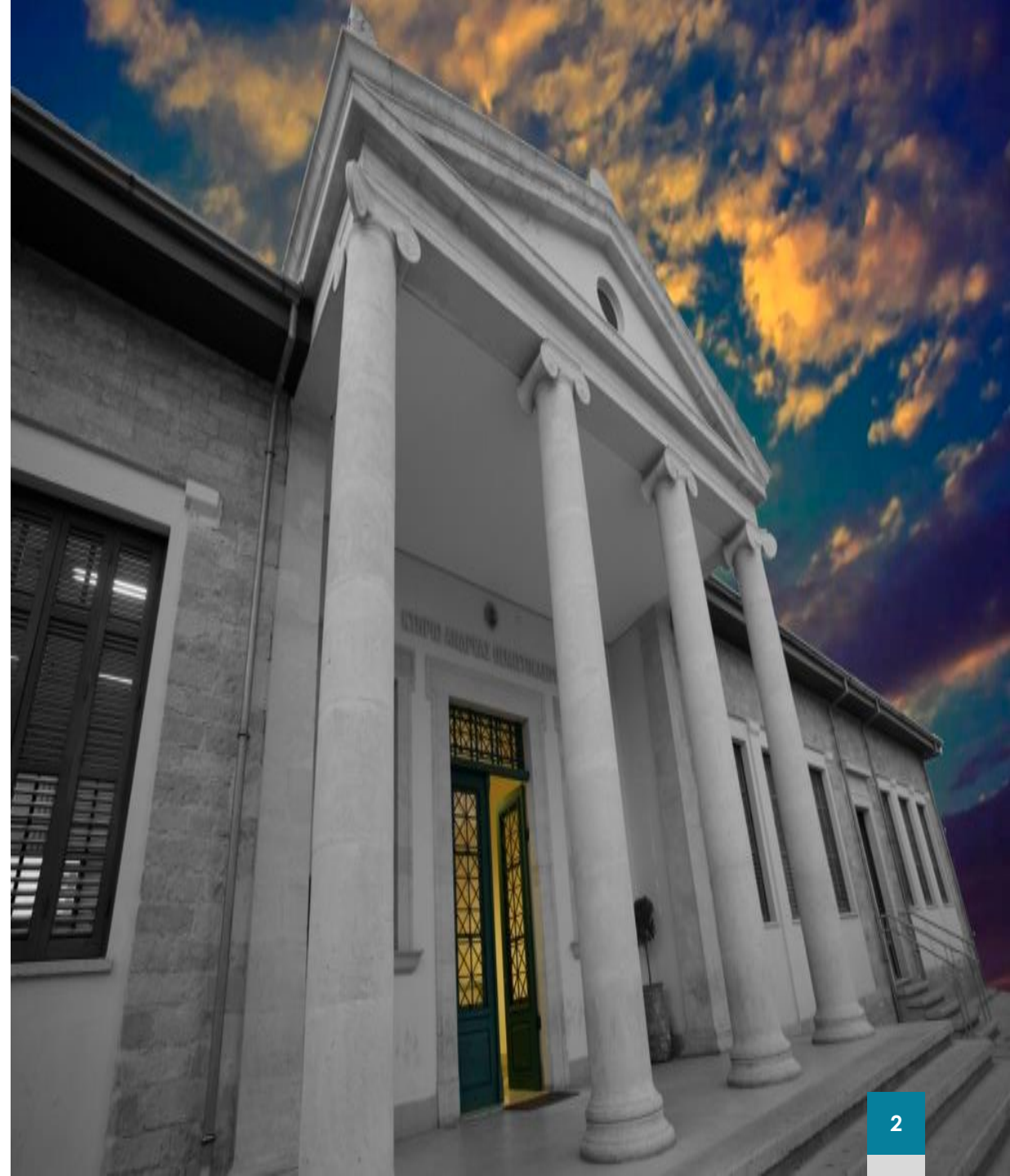
Presenter: PhD Candidate Spyros Loizou

Email: Spyros.Loizou@cut.ac.cy



Short Bios

- **BSc** in *Computer Engineering and Informatics*
(Cyprus University of Technology 2013 – 2017)
- **MSc** in *Data Science and Engineering*
(Cyprus University of Technology 2017 - 2019)
- **Working Experience**
 - Dossier Cloud – Twinning H2020 (2017 – 2019)
 - Epirroi Greece – Cyprus Interreg Programme (2017 – 2019)
 - Destini – Twinning H2020 (2019 – 2022)
 - Digital Logic Lab in CUT (2019 – 2021) – Teaching Assistant
 - VLSI Systems Design Lab in CUT (2019 – 2021) – Teaching Assistant



Contents

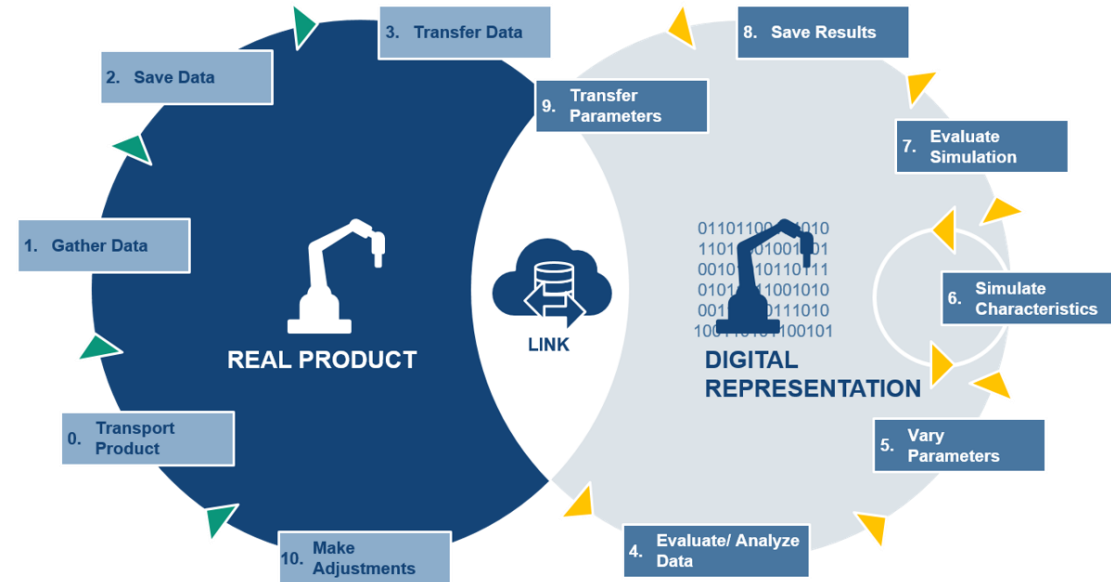
- Introduction
- Technical Background
- Methodology
- System Demonstration
- Future Work
- References

Introduction (1/4)

- Scientific Trends:
 - Internet of Things (IoT)
 - Big Data
 - Cloud Computing
 - Artificial Intelligence (AI)
- Large Volume of data (Structured, Semi-Structured and Unstructured)
- Data processing and Visualization
- Transform primary data to meaningful data

Introduction (2/4)

- Digital Twin
 - Virtual representation
 - Simulation
 - Support Decision Making
 - Predictions



Introduction (3/4)

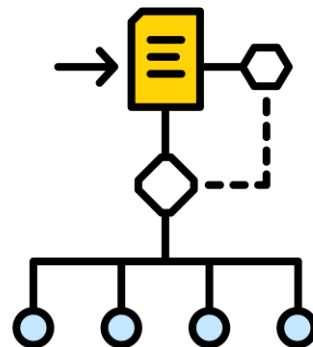
- Process Mining
 - Technique relating the fields of data science and process management
 - Support the analysis of operational processes based on event logs
 - Turn event data into insights and actions
 - Use event data to show what people, machines, and organizations are really doing

Introduction (4/4)

- Open Research Challenge
 - How Digital Twins may contribute to enhancing the applicability and efficiency of process models
 - Prevent costly failures in physical objects or activities
 - Improve quality and productivity, by using advanced analytical, monitoring and predictive capabilities, test processes and services

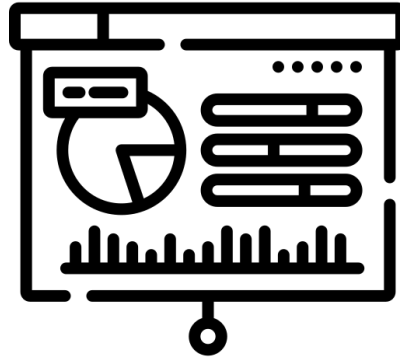
Technical Background (1/4)

- Data Processing
 - Smart Data Processing Systems for data ingestion, data aggregation of an enormous variety of structured, unstructured and semi-structured datasets
 - Knowledge-based meta-data representation techniques for the conversion of raw into smart data, data privacy and protection, and dynamic configuration.



Technical Background (2/4)

- Data Visualization
 - Graphical representation of information and data.
 - Data visualization tools provide an accessible way to see and understand trends, outliers, and patterns in data.
 - Visualization-based data discovery methods allow business users to mash up disparate data sources to create custom analytical views.



Technical Background (3/4)

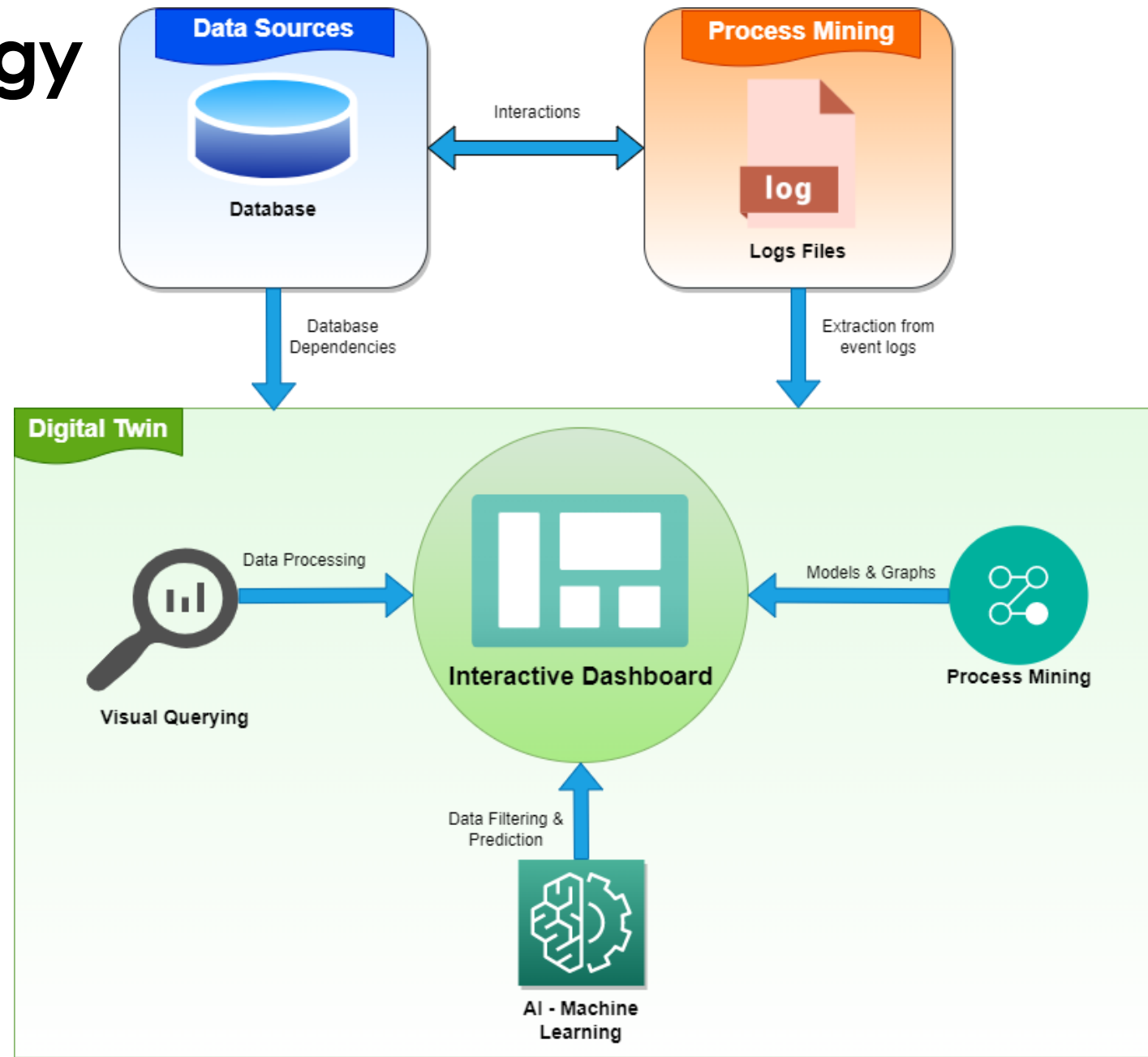
- Digital Twin
 - Traditionally characterized by two-way interactions between the digital and the physical world
 - A DT offers error optimization to save money and time, reduces defects and manages the lifecycle of the Internet of things (IoT)
 - **Idea of DTs to graphically represent and interact with event data and process logs and applies this approach to industrial environments**



Technical Background (4/4)

- Process Mining
 - Technique designed to discover, monitor and improve real processes by extracting readily available knowledge from event logs stored in information systems
 - Use different algorithms to extract and organize data and business flows, with the top 5 mining algorithms being Alpha Miner, Fuzzy miner, Heuristic miner, Inductive Miner and Genetic miner
 - **Provide the means for a totally different user experience based on visual querying and process mining data-driven tasks, which is characterized by simplicity, self-explainability, ease of use and graphical ergonomics**

Methodology



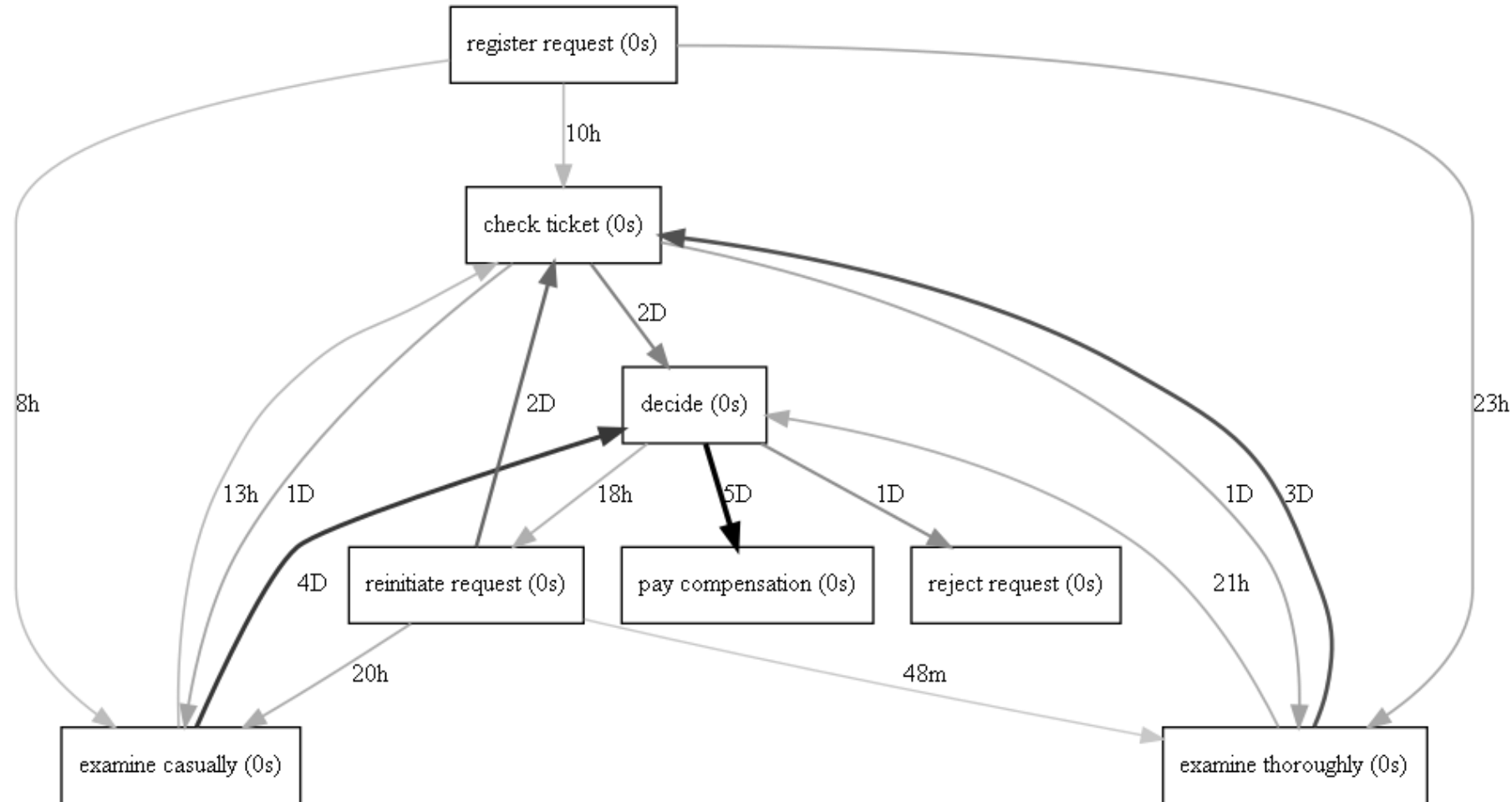
Methodology

- A dedicated software tool was developed to demonstrate the proposed framework
- Python, mainly using pm4py, pandas and Streamlit libraries
- Unity was the environment used to produce playful, aesthetically correct and user-friendly graphics for the presentation and interactive use of the dashboard



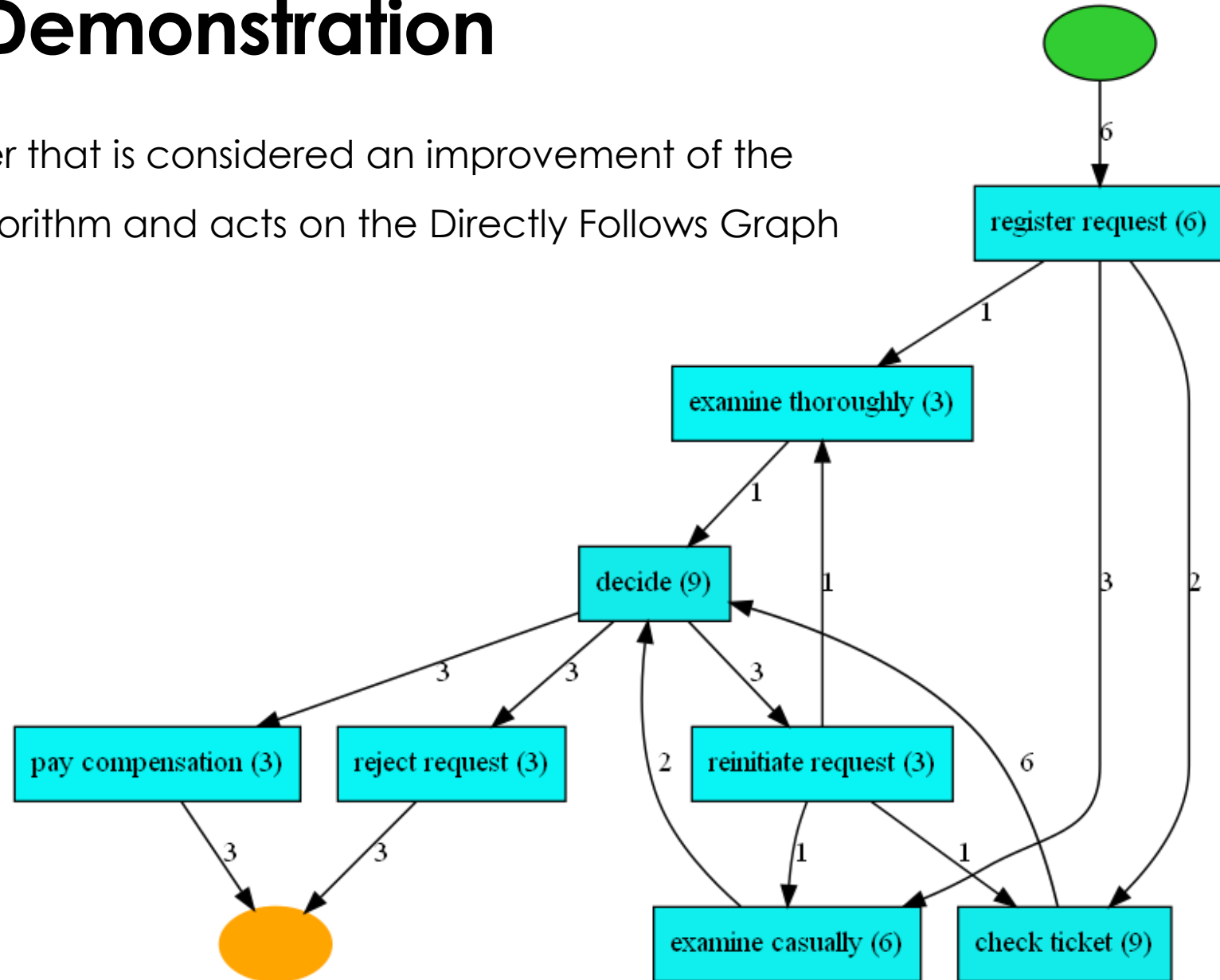
System Demonstration

- Directly Follows Graph based on average time



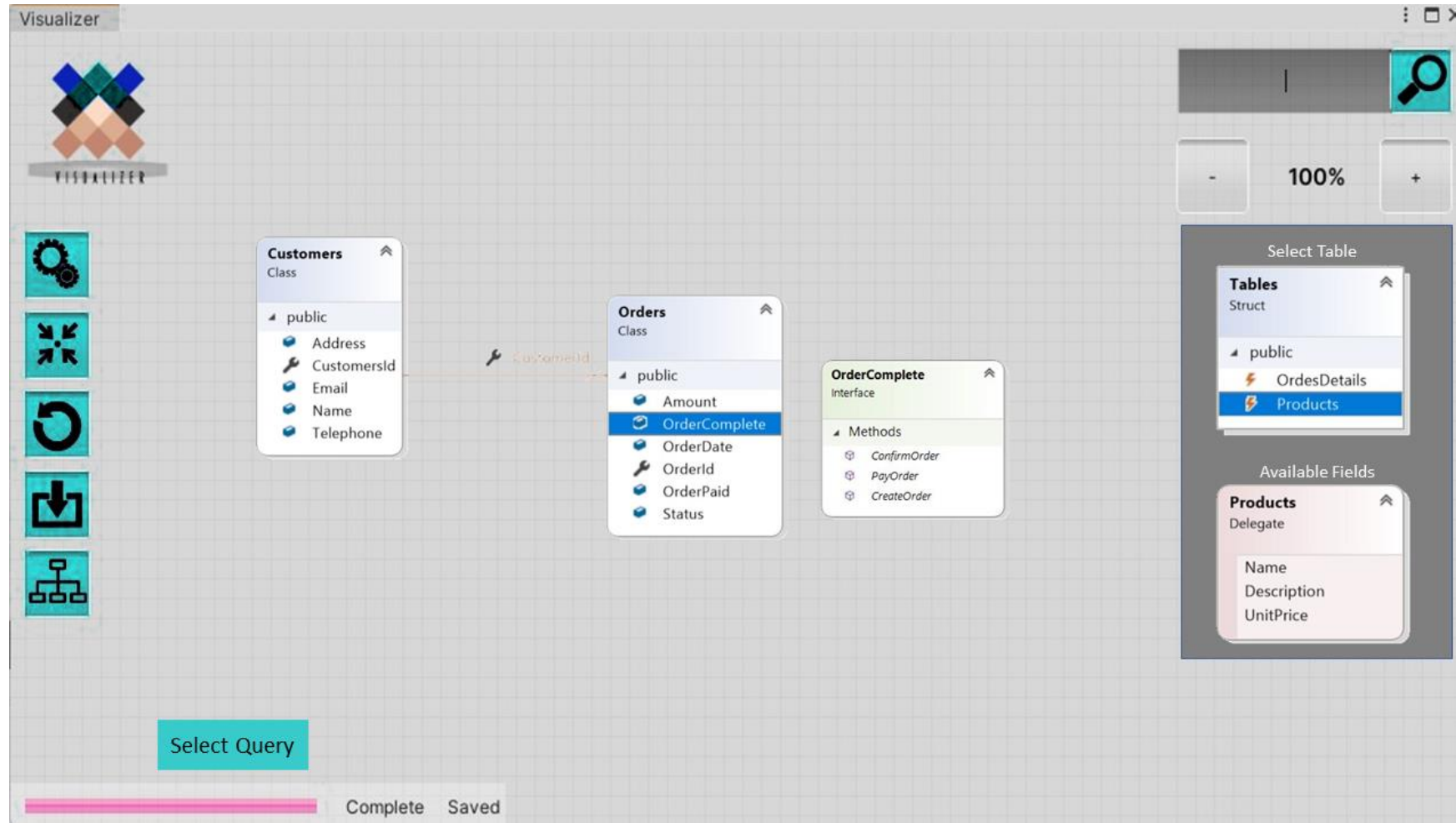
System Demonstration

- Heuristic Miner that is considered an improvement of the Alpha Miner algorithm and acts on the Directly Follows Graph



System Demonstration

- Selection of tables and attributes of interest



System Demonstration

- Selection of logical and arithmetic operators

The screenshot displays the Visualizer application interface, which is a query builder tool. The main workspace shows a diagram with two tables: **Customers** and **Orders**. The **Customers** table has fields: Address, CustomersId, Email, Name, and Telephone. The **Orders** table has fields: OrderDate and OrderId. A relationship line connects the **CustomersId** field in the **Customers** table to the **CustomerId** field in the **Orders** table. On the left side, there is a vertical toolbar with icons for settings, zoom, refresh, save, and a tree view. On the right side, there is a **Select Table** panel showing a list of tables: **public** (expanded), **OrdersDetails**, and **Products**. Below this is an **Available Fields** panel for the **Products** table, listing fields: Name, Description, and UnitPrice. At the bottom, there is a **Select Query** panel with a toolbar containing comparison operators: **>**, **<**, **=**, **>=**, **<=**, and **a<x<b**. A dropdown menu labeled **Between** is open above the **a<x<b** operator. The status bar at the bottom indicates **Complete** and **Saved**.

System Demonstration

- Creation and execution of a visual query

The screenshot displays the Visualizer application interface. At the top left is the 'Visualizer' logo. Below it is a vertical toolbar with icons for settings, zoom, refresh, save, and a hierarchical view. The main workspace shows a visual query diagram with two data classes: 'Customers' and 'Orders'. The 'Customers' class has fields: Address, CustomersId, Email, Name, and Telephone. The 'Orders' class has fields: OrderDate and OrderId. A line connects the 'CustomersId' field of the 'Customers' class to the 'OrderDate' field of the 'Orders' class. On the right side, there is a 'Select Table' panel showing a tree view with 'public' selected, containing 'OrdesDetails' and 'Products'. Below this is an 'Available Fields' panel for the 'Products' table, listing 'Name', 'Description', and 'UnitPrice'. At the bottom, there is an 'Execute Query' button, a date range input field containing '2022-01-01 < x < 2022-04-30', and a status bar showing 'Complete' and 'Saved'.

System Demonstration

- Results obtained from a visual query



The screenshot shows a software window titled 'Visualizer'. In the top-left corner is a logo consisting of a stylized 'X' made of four colored squares (blue, green, orange, black) with the word 'VISUALIZER' below it. To the right of the logo is a search bar with a magnifying glass icon and zoom controls showing '- 100% +'. On the left side of the window is a vertical toolbar with five icons: a gear, a four-way arrow, a circular arrow, a download arrow, and a hierarchical tree structure. The main area of the window displays a table with the following data:

CustomersId	Name	Email	Address	Telephone	OrderId	OrderDate
1451	John	john@example.com	Steet no 1	99111111	147854	2022-01-05
1001	Andrew	Andrew@example.com	Street no 10	99101010	254114	2022-02-10
1018	Chris	chris@example.com	Street no 22	99221122	189588	2022-02-22
1451	John	john@example.com	Street no 1	99111111	156040	2022-04-03
1451	John	john@example.com	Street no 1	99111111	156065	2022-04-10
1991	Helen	jelen@example.com	Street no 9	99121314	221478	2022-04-28

At the bottom of the window, there is a 'Back' button and a status bar with a pink progress indicator, the text 'Complete', and a 'Saved' button.

Future Work

- Use of Data Lakes in which sorted and cleaned data will be hosted
- More sophisticated visualization features supporting predictive analytics
- Suggestion of business flow corrections to achieve better process results





THANK YOU

Email: Spyros.Loizou@cut.ac.cy

References

- L. Wang, G. Wang, and C. A. Alexander, "Big Data and Visualization: Methods, Challenges and Technology Progress," Digit. Technol., vol. 1, no. 1, pp. 33–38, 2015, doi: 10.12691/dt-1-1-7
- A. Rasheed, O. San, and T. Kvamsdal, "Digital twin: Values, challenges and enablers from a modeling perspective," IEEE Access, vol. 8, pp. 21980–22012, 2020, doi: 10.1109/ACCESS.2020.2970143
- C. A. Steed et al., "Big data visual analytics for exploratory earth system simulation analysis," Comput. Geosci., vol. 61, pp. 71–82, 2013, doi: 10.1016/j.cageo.2013.07.025
- J. Vachalek, L. Bartalsky, O. Rovny, D. Sismisova, M. Morhac, and M. Loksik, "The digital twin of an industrial production line within the industry 4.0 concept," Proc. 2017 21st Int. Conf. Process Control. PC 2017, pp. 258–262, 2017, doi: 10.1109/PC.2017.7976223.
- M. Schluse, M. Priggemeyer, L. Atorf, and J. Rossmann, "Experimentable Digital TwinsStreamlining Simulation-Based Systems Engineering for Industry 4.0," IEEE Trans. Ind. Informatics, vol. 14, no. 4, pp. 1722–1731, 2018, doi: 10.1109/TII.2018.2804917
- A. Fuller, Z. Fan, C. Day, and C. Barlow, "Digital Twin: Enabling Technologies, Challenges and Open Research," IEEE Access, vol. 8, pp. 108952–108971, 2020, doi: 10.1109/ACCESS.2020.2998358
- A. Burattin, Process mining techniques in business environments: Theoretical aspects, algorithms, techniques and open challenges in process mining, vol. 207. 2015

References

- R. Accorsi, M. Ullrich, and W. M. P. Van Der Aalst, Process mining, vol. 35, no. 5. 2012
- M. Bozkaya, J. Gabriels, and J. M. Van Der Werf, "Process diagnostics: A method based on process mining," Proc. - Int. Conf. Information, Process. Knowl. Manag. eKNOW 2009, no. 2, pp. 22–27, 2009, doi: 10.1109/eKNOW.2009.29
- G. Park and W. M. P. Van Der Aalst, "Realizing A Digital Twin of An Organization Using Action-oriented Process Mining," Proc. - 2021 3rd Int. Conf. Process Mining, ICPM 2021, pp. 104–111, 2021, doi: 10.1109/ICPM53251.2021.9576846
- Y. Pan and L. Zhang, "A BIM-data mining integrated digital twin framework for advanced project management," Autom. Constr., vol. 124, no. July 2020, p. 103564, 2021, doi: 10.1016/j.autcon.2021.103564
- C. Mehrotra, N. Chitransh, and A. Singh, "Scope and challenges of visual analytics: A survey," Proceeding - IEEE Int. Conf. Comput. Commun. Autom. ICCCA 2017, vol. 2017-Janua, no. 4404, pp. 1229–1234, 2017, doi: 10.1109/CCAA.2017.8229987
- S. Makinen, H. Skogstrom, E. Laaksonen, and T. Mikkonen, "Who needs MLOps: What data scientists seek to accomplish and how can MLOps help?," Proc. -2021 IEEE/ACM 1st Work. AI Eng. - Softw. Eng. AI, WAIN 2021, pp. 109–112, 2021, doi: 10.1109/WAIN52551.2021.00024