# *An Inside Look at Customer Journey Analytics*

Periklis Andritsos

Faculty of Information

University of Toronto

periklis.Andritsos@utoronto.ca

Credits: Gaël Bernard, PostDoc, University of Toronto

# About Me

- BSc, from National Technical University Of Athens (Timos Sellis' group)
- MSc/PhD, from University of Toronto
    - LIMBO: Scalable Clustering of Categorical Data
    - Structure Discovery in Large Data Sets
- Taught in Italy and Switzerland
- Co-founded two start-up companies
    - Thoora.com
    - Odaia.ai
- Currently faculty at the University of Toronto

# Agenda

- Introduction
- Contributions
    1. Customer Journey Discovery
    2. CJM-explorer (CJM-ex tool)
    3. Truncated Trace Classifier
    4. Path Prediction
- Conclusion

# *Introduction*

[6] Bernard, G., & Andritsos, P. (2017). A process mining based model for customer journey mapping. In Forum and Doctoral Consortium Papers Presented at the 29th International Conference on Advanced Information Systems Engineering (CAiSE 2017) (Vol. 1848, pp. 49-56). CEUR Workshop Proceedings.

# Motivations

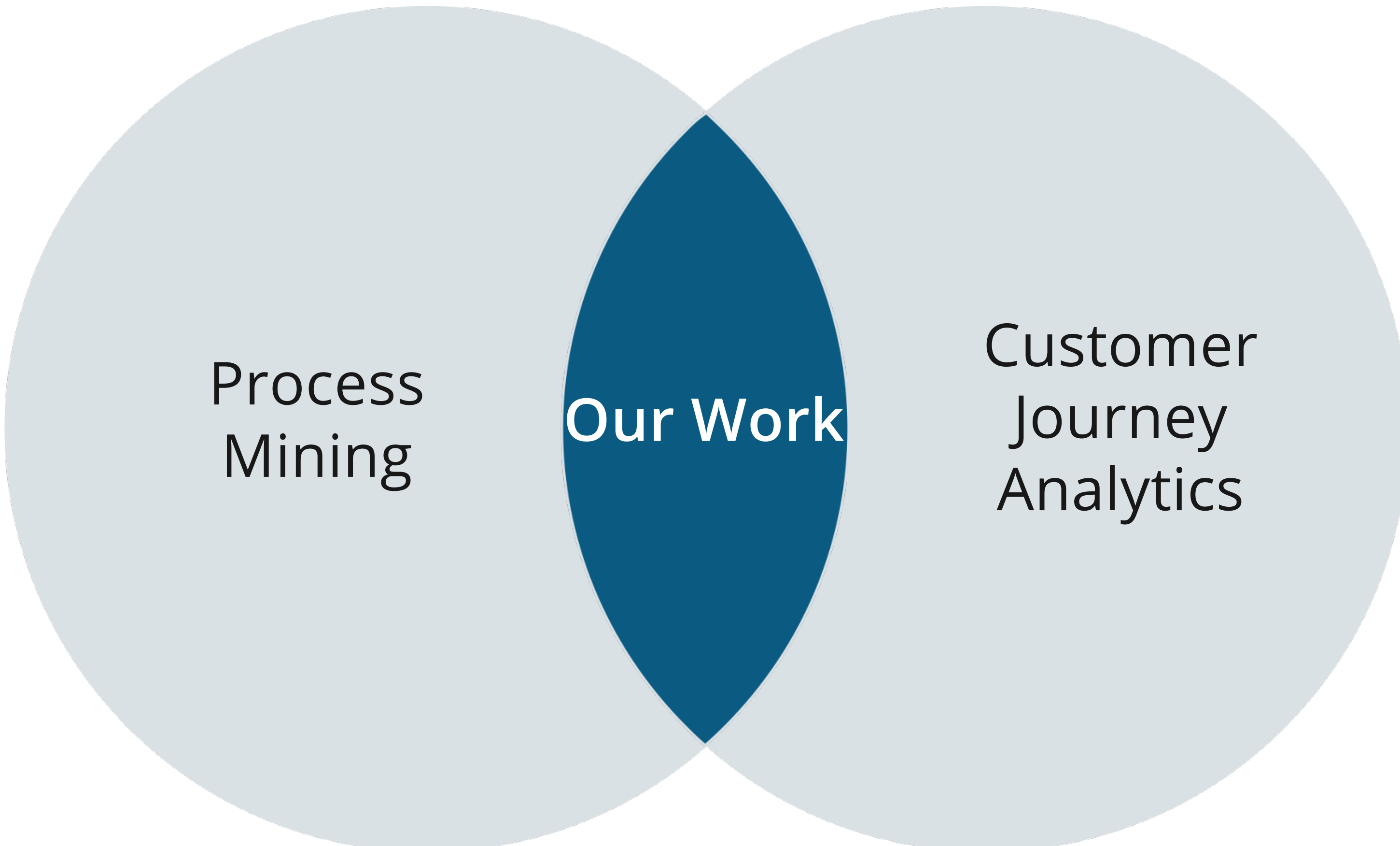
## Customer Journey Analytics

- Increasing number of:
  - Devices
  - Channels
  - Technologies

- Customer's interactions are complex and unique
  - Growing importance of understanding them [61]

[61] Lemon, K. N. and Verhoef, P. C. (2016). Understanding customer experience throughout the customer journey. Journal of Marketing, 80(6):69–96.

# Introduction



# Process Mining

## Introduction

- Data science
  - Process agnostic
- Process science
  - Disregard evidences hidden in the data
- Process mining bridges this gap [86].

[86] van der Aalst, W. (2016). Process Mining: Data Science in Action. Springer.

# Process Mining

## Event Log

**Trace**

**Trace**

**Trace**

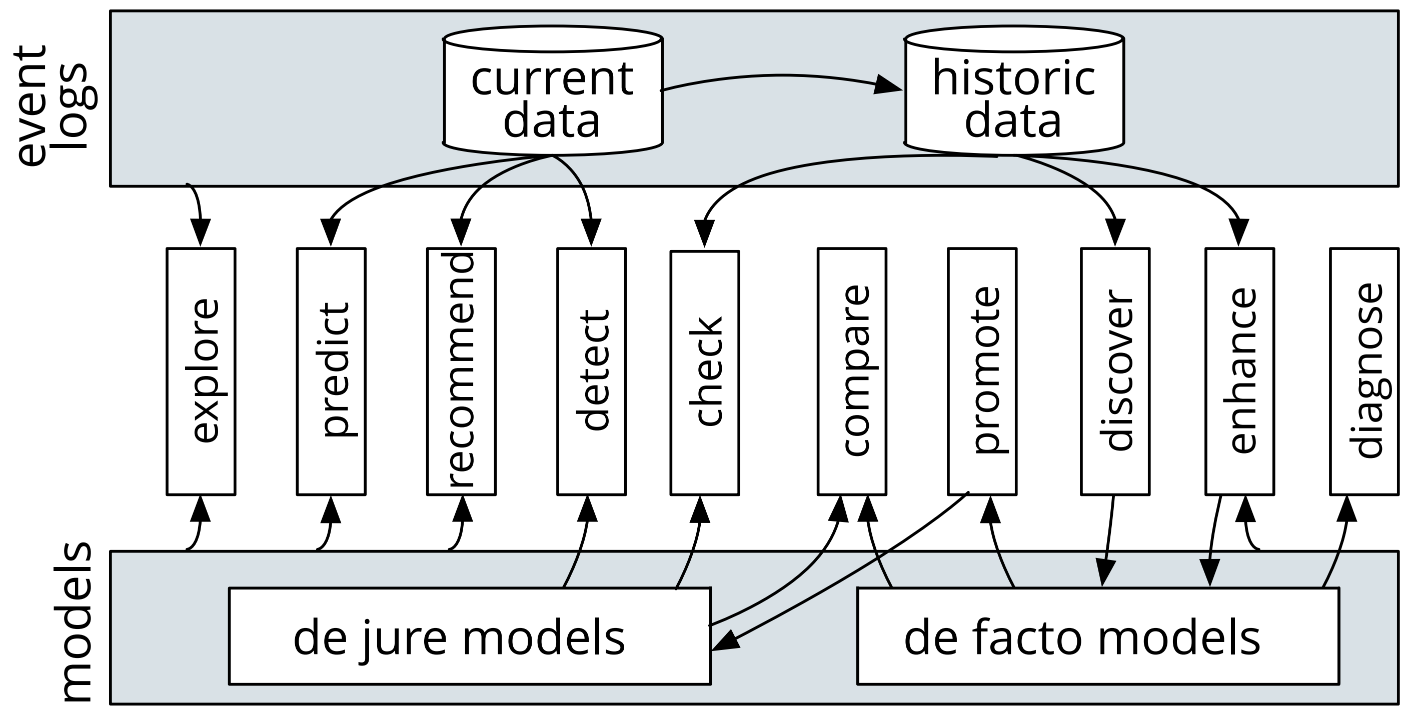
**Events**

09.09.20 - 16:35 Renting a car at the service desk
09.09.20 - 17:37 Booking an insurance
20.09.20 - 13:12 Picking up the car
28.09.20 - 17:14 Returning the car

http://xes-standard.org

# Process Mining

## Framework [86]

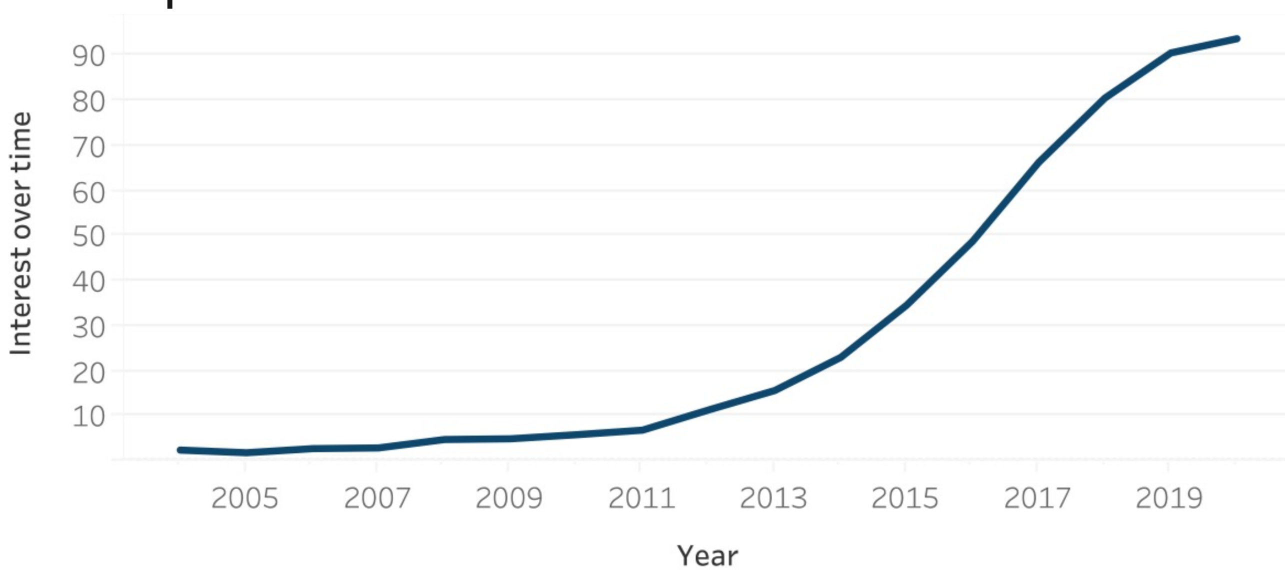



[86] van der Aalst, W. (2016). Process Mining: Data Science in Action. Springer.

# Customer Journey

## Introduction

- Customer journey
  - Sequence of touchpoints
- Touchpoints
  - Interactions between the customer and the service provider.
- Popular topic

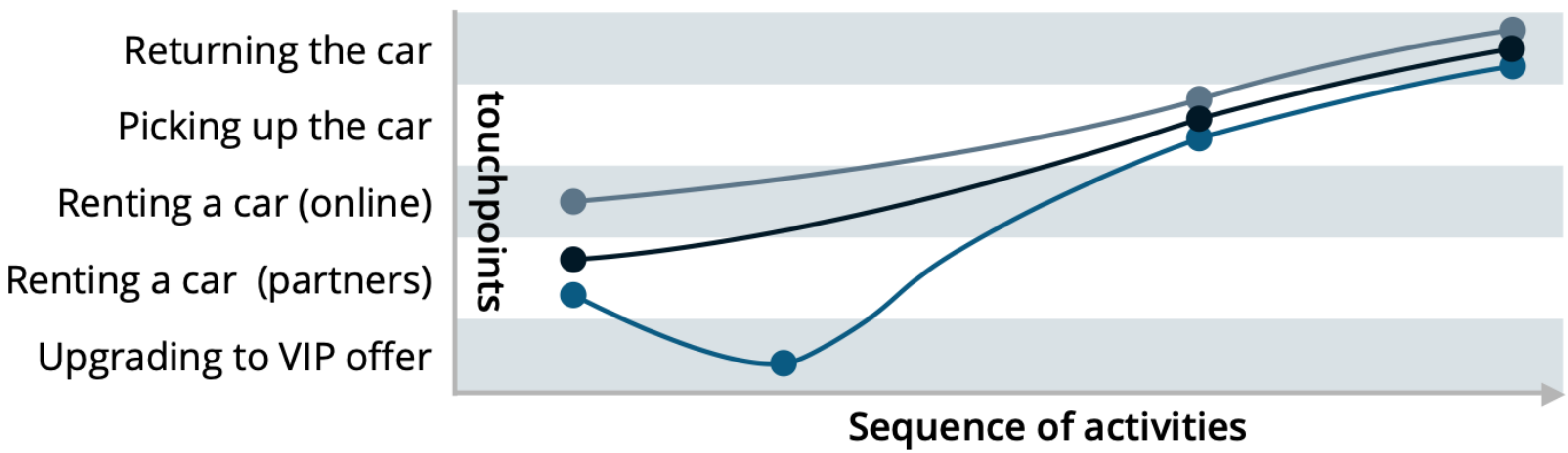



Source: https://trends.google.com/trends/explore?date=all&q=customer%20journey

# Customer Journey Map (CJM)

## Introduction

> *Customer journey mapping is used to understand a customer's behaviour, feelings, motivations and attitudes while using a service. [54]*



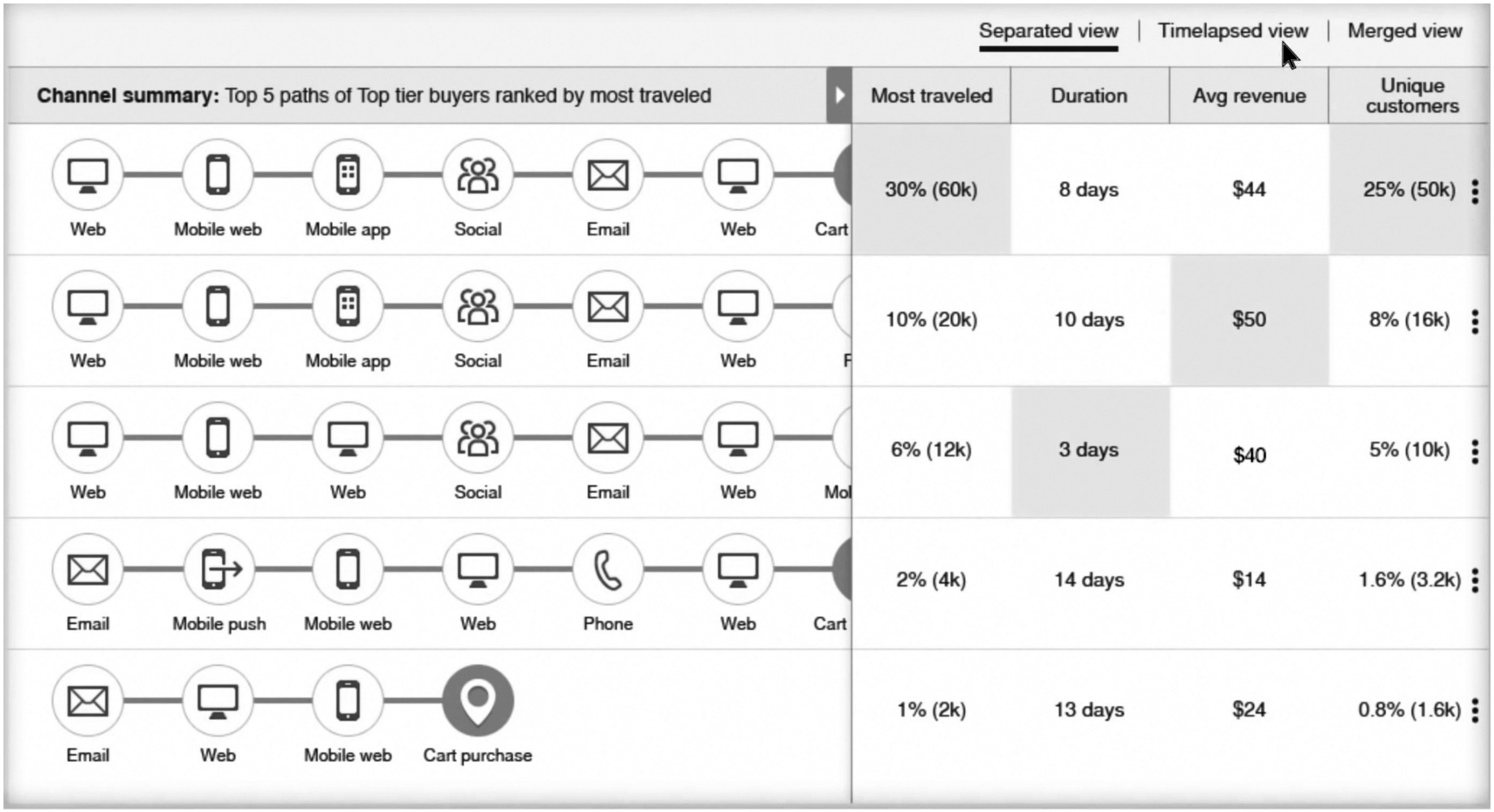

[54] Kojo, I., Heiskala, M., & Virtanen, J. P. (2014). Customer Journey Mapping of an Experience-Centric Service by Mobile Self-reporting: Testing the Qualiwall Tool. In International Conference of Design, User Experience, and Usability. Springer International Publishing.

# Expected CJM [40]

Deloitte Nederland - https://www.youtube.com/watch?v=JBF2dwCP80Y&t=84s

[40] Følstad, A., Kvale, K., and Halvorsrud, R. (2013). Customer journey measures - state of the art research and best practices. SINTEF Rapport A24488, 28 p. SINTEF, 2013. http://hdl.handle.net/11250/2390670

# Actual CJM [40]




Source: IBM Watson Marketing - https://www.youtube.com/watch?v=QaM6-x4Wfv4

[40] Følstad, A., Kvale, K., and Halvorsrud, R. (2013). Customer journey measures- state of the art research and best practices. Oslo, Norway: Report A, 24488.

# Research Questions

**How can customer journey maps be discovered, explored, and enhanced from event logs?**

1 Genetic Customer Journey Discovery

2 CJM-explorer

CJM-abstractor

**How can the touchpoints of a customer journey be predicted?**

3 Truncated Trace Classifier

4 Path Prediction

# Contributions

## Acknowledged by peers

> *Bernard and Andritsos are the ones who first highlighted the* ***prospective value of process mining*** *for customer by illustrating how it can be used to analyze the* ***customer journey****. Moreover, they demonstrate a* ***perfect correspondence*** *between the components of a customer journey map and the XES format, in which process mining event logs are stored. [A]*

[A] Terragni, A., & Hassani, M. (2019, April). Optimizing customer journey using process mining and sequence-aware recommendation. In Proceedings of the 34th ACM/SIGAPP Symposium on Applied Computing (pp. 57-65).

# Contributions

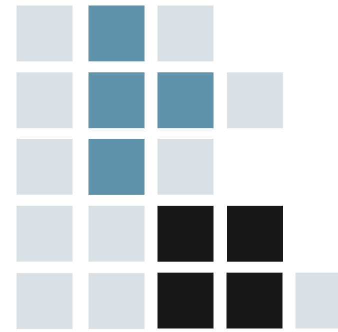
## Mapping [6]

| Process Mining | CJM |
|---|---|
| Trace | Journey |
| Event | Touchpoint |
| De Jure Process Model | Expected CJM |
| De Facto Process Model | Actual CJM |

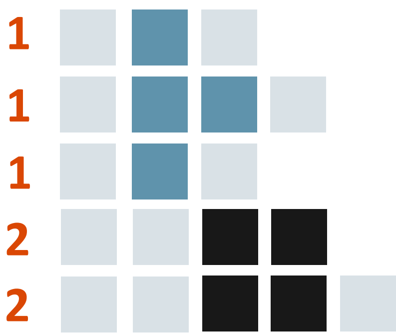
[6] Bernard, G., Andritsos, P. (2017). A process mining based model for customer journey mapping. In: Forum and Doctoral Consortium Papers Presented at the 29th International Conference on Advanced Information Systems Engineering . CAiSE 2017 (Forum) (Vol. 1848, pp. 49-56). CEUR Workshop Proceedings.

# Contributions

## Illustrated with a dotted chart [B]



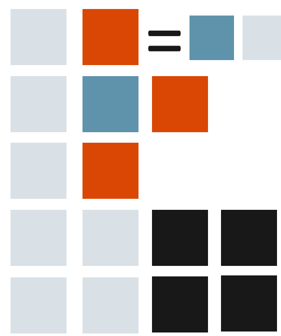
### Horizontal Clustering




1 Genetic Customer Journey Discovery

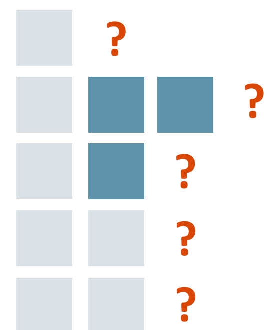
2 CJM-explorer

### Vertical Clustering




CJM-abstractor

### Event Predictions




3 Truncated Trace Classifier

4 Event Predictions

[B] Song, M., & van der Aalst, W. (2007, December). Supporting process mining by showing events at a glance. In *Proceedings of the 17th Annual Workshop on Information Technologies and Systems (WITS)* (pp. 139-145).

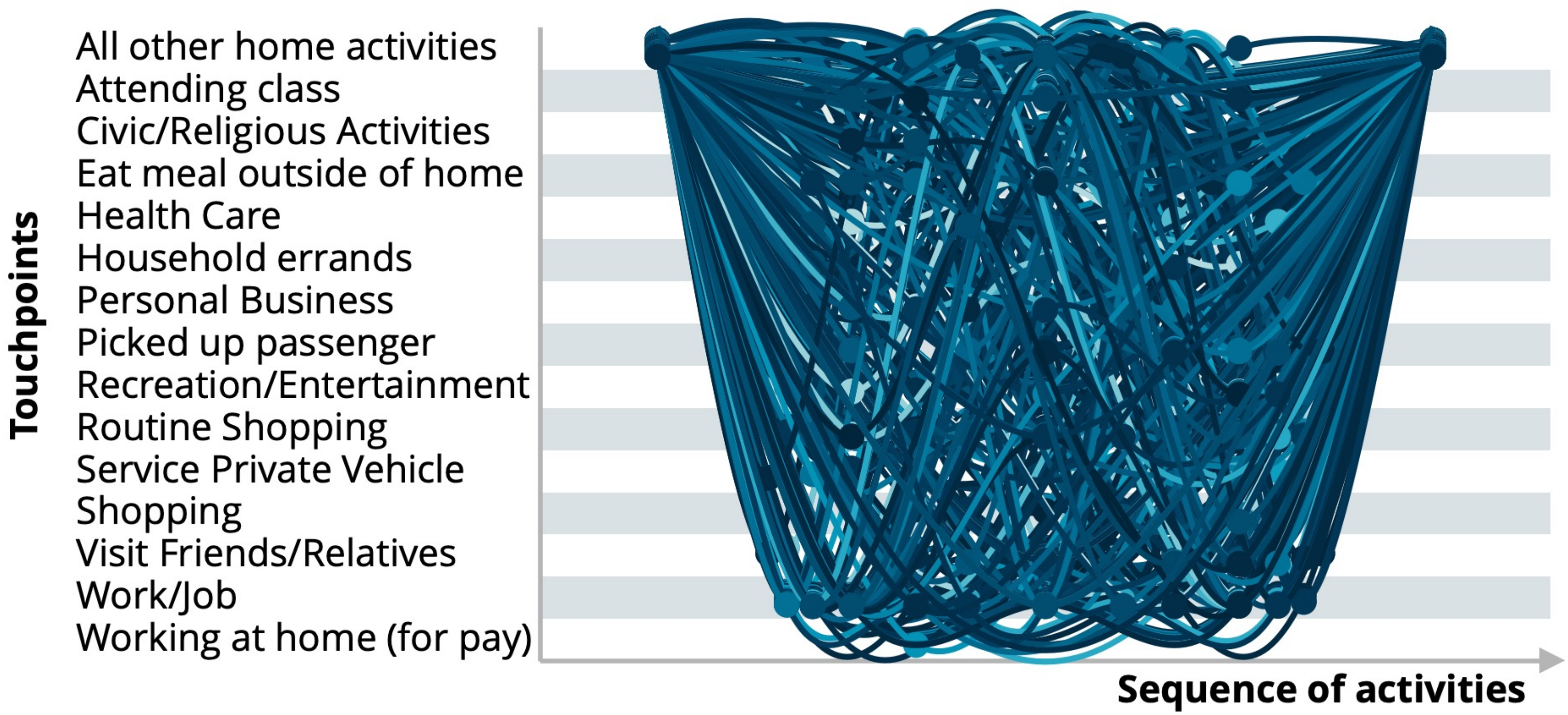
# 1 *Customer Journey Discovery*

[9] Bernard G., Andritsos P. (2019) Contextual and Behavioral Customer Journey Discovery Using a Genetic Approach. In: Welzer T., Eder J., Podgorelec V., Kamišalić Latifić A. (eds) Advances in Databases and Information Systems. ADBIS 2019. Lecture Notes in Computer Science, vol 11695. Springer, Cham. https://doi.org/10.1007/978-3-030-28730-6_16

# Introduction

## Chicago, my daily travel survey [C]

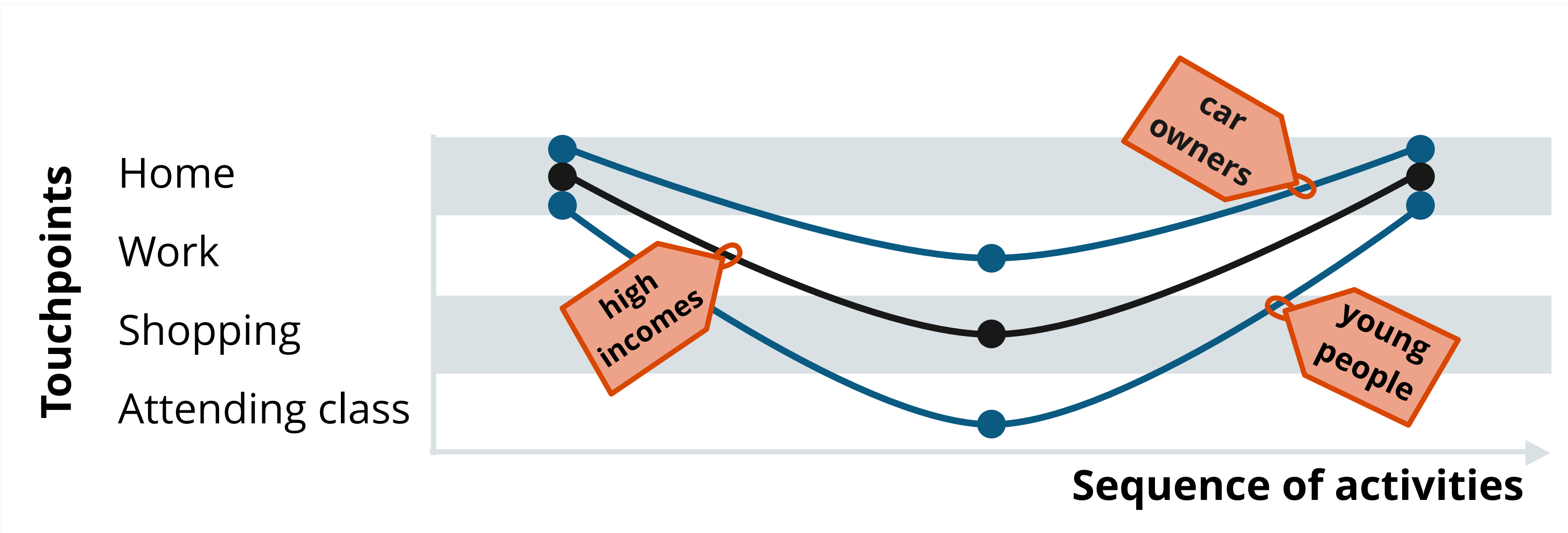
15 unique activities
124K activities
29.5K journeys




[C] Dataset: https://www.cmap.illinois.gov/data/transportation/travel-survey

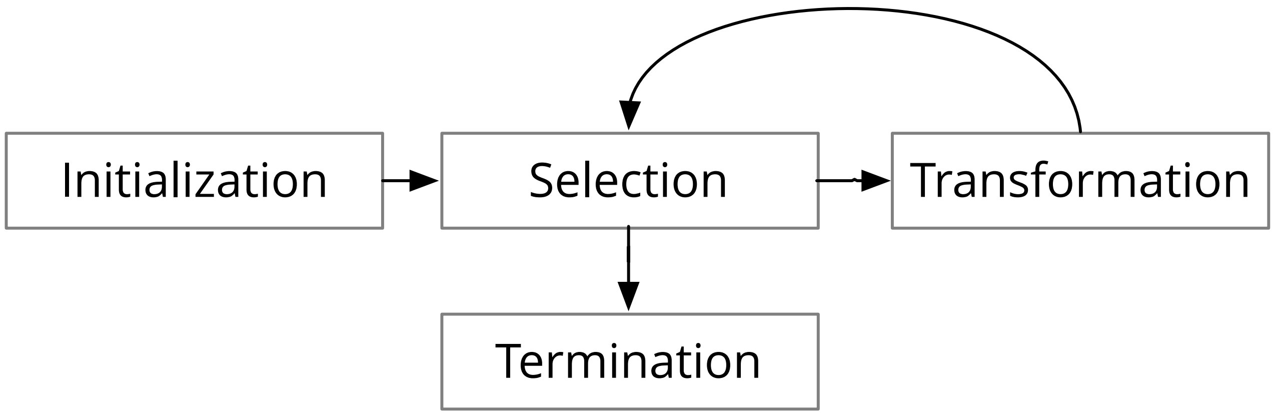
# Introduction

## Contextual and behavioral representative journeys



# Genetic Approach

- Converge to an optimal solution given personalized evaluation criteria.
- Approach




- Inspired by process mining [20,31,91]

[20] Buijs, J. C., van Dongen, B. F., and van der Aalst, W. M. (2012). A genetic algorithm for discovering process trees. In Evolutionary Computation (CEC), 2012 IEEE Congress on, pages 1–8. IEEE.

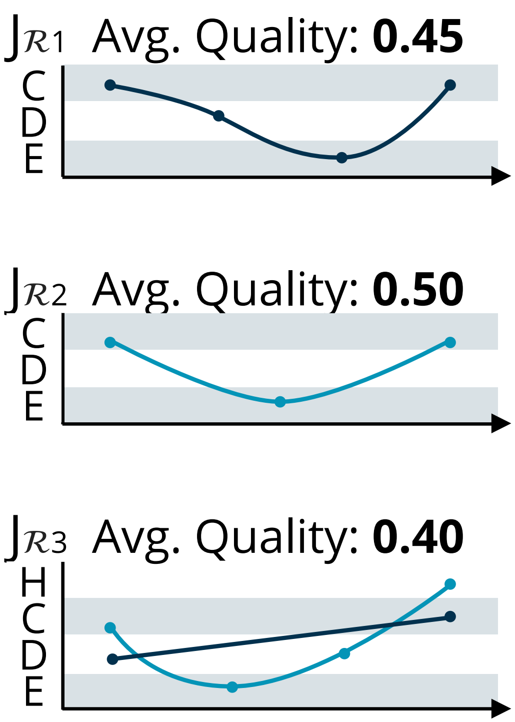
[31] De Medeiros, A.A., Weijters, A.: Genetic process mining. In: Applications and Theory of Petri Nets 2005, Volume 3536 of Lecture Notes in Computer Science. Citeseer (2005)

[91] Vazquez-Barreiros, B., Mucientes, M., Lama, M.: Prodigen: Mining complete, precise and minimal structure process models with a genetic algorithm. Information Sciences 294, 315–333 (2015)

# Genetic Approach

## Initial population and evaluation

**Generation 1**




**Evaluation:**

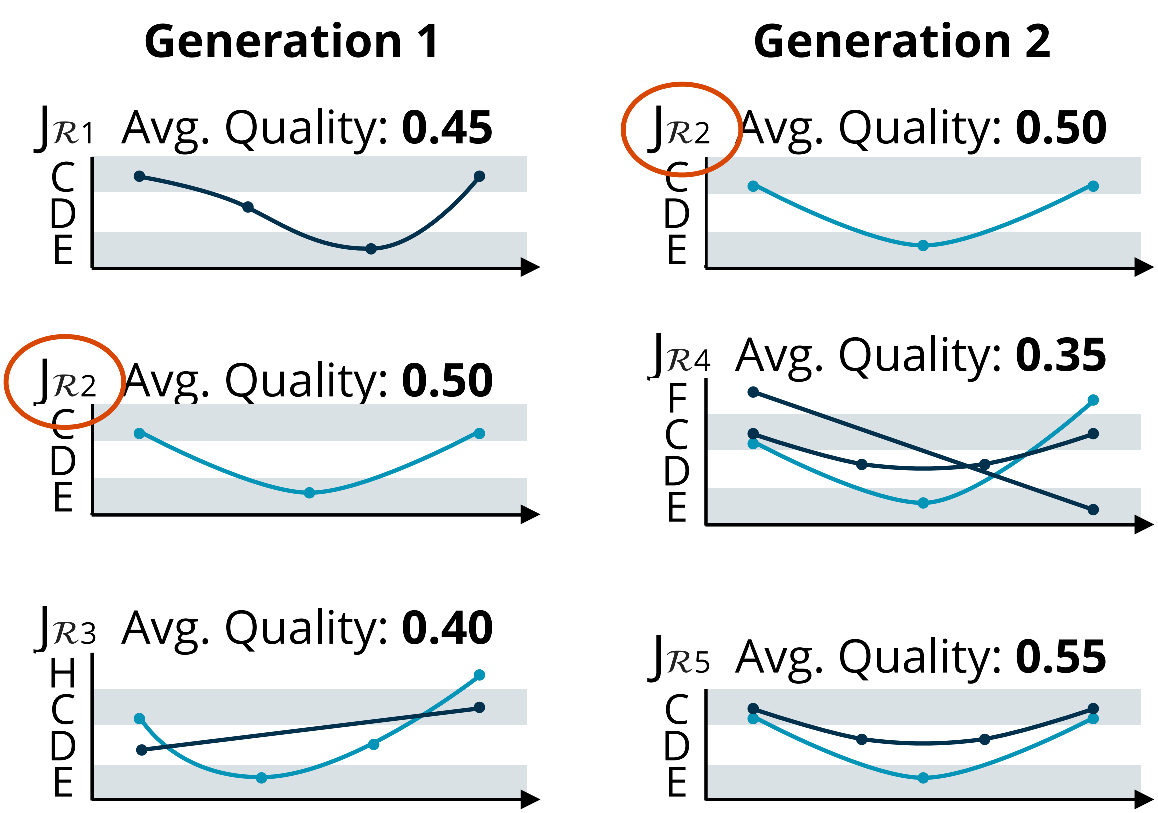
1. Fitness $$1 - \frac{\sum_{i=1}^{|J_{\mathcal{A}}|} min_{j=1}^{|J_{\mathcal{R}}|}(Levenshtein(\sigma_{\mathcal{A}_i}; \sigma_{\mathcal{R}_j}))}{\sum_{i=1}^{|J_{\mathcal{A}}|} Length(\sigma_{\mathcal{A}_i})}$$

2. Distance from $K_h$ $$\frac{1}{1 + (\frac{|k_{\mathcal{R}} - k_h|}{x_0})^2}$$

3. Contextual Distance $$\frac{v_1 \cdot v_2}{||v_1|| \cdot ||v_2||}$$

# Genetic Approach

## Elite population and transformations

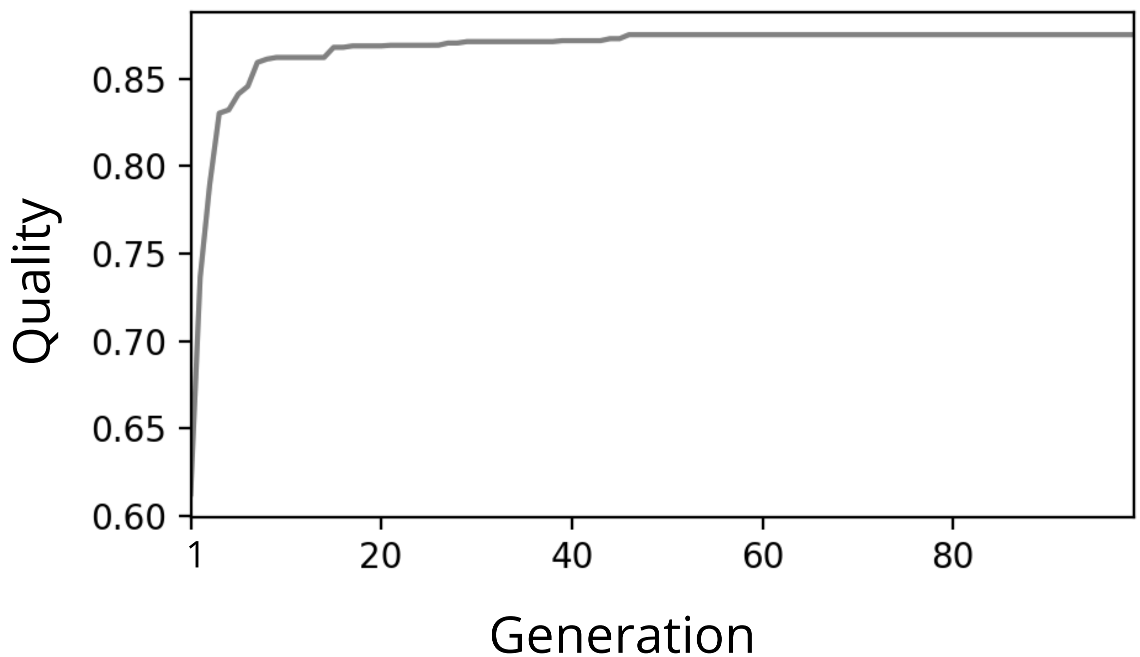


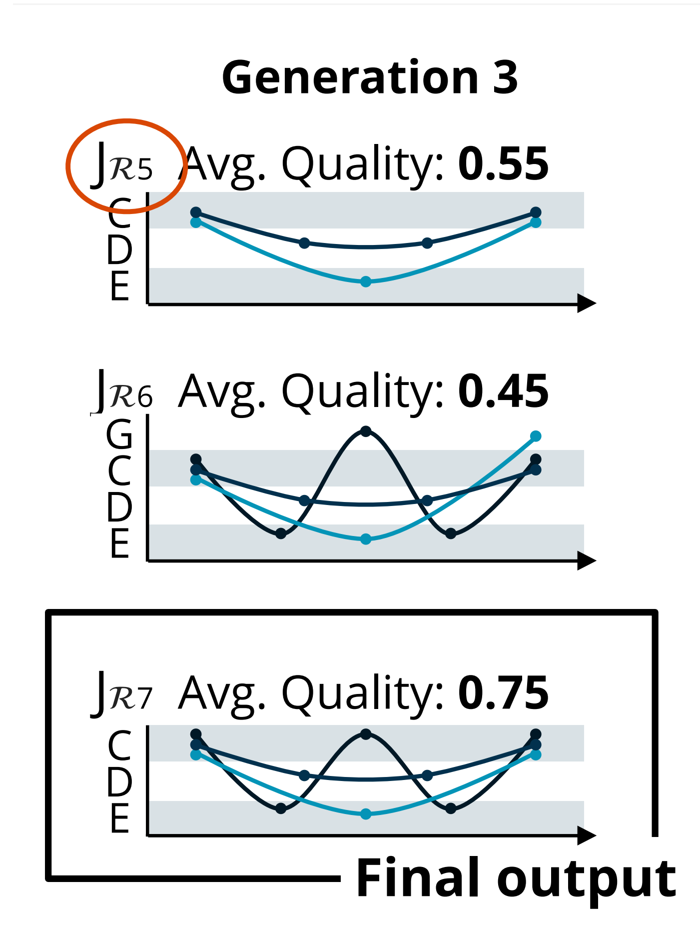

**Transformations:**

1. Add a touchpoint
2. Remove a touchpoint
3. Add a journey
4. Remove a journey
5. Crossover

# Genetic Approach

## Stop



# Evaluation

## Settings

- Datasets
  - 40 Synthetic CJMs [D]
  - Generative journeys used to create slightly altered journeys
  - Goal: find back generative journeys from altered journeys
  - Evaluation: Jaccard Distance
- Traminer [41]

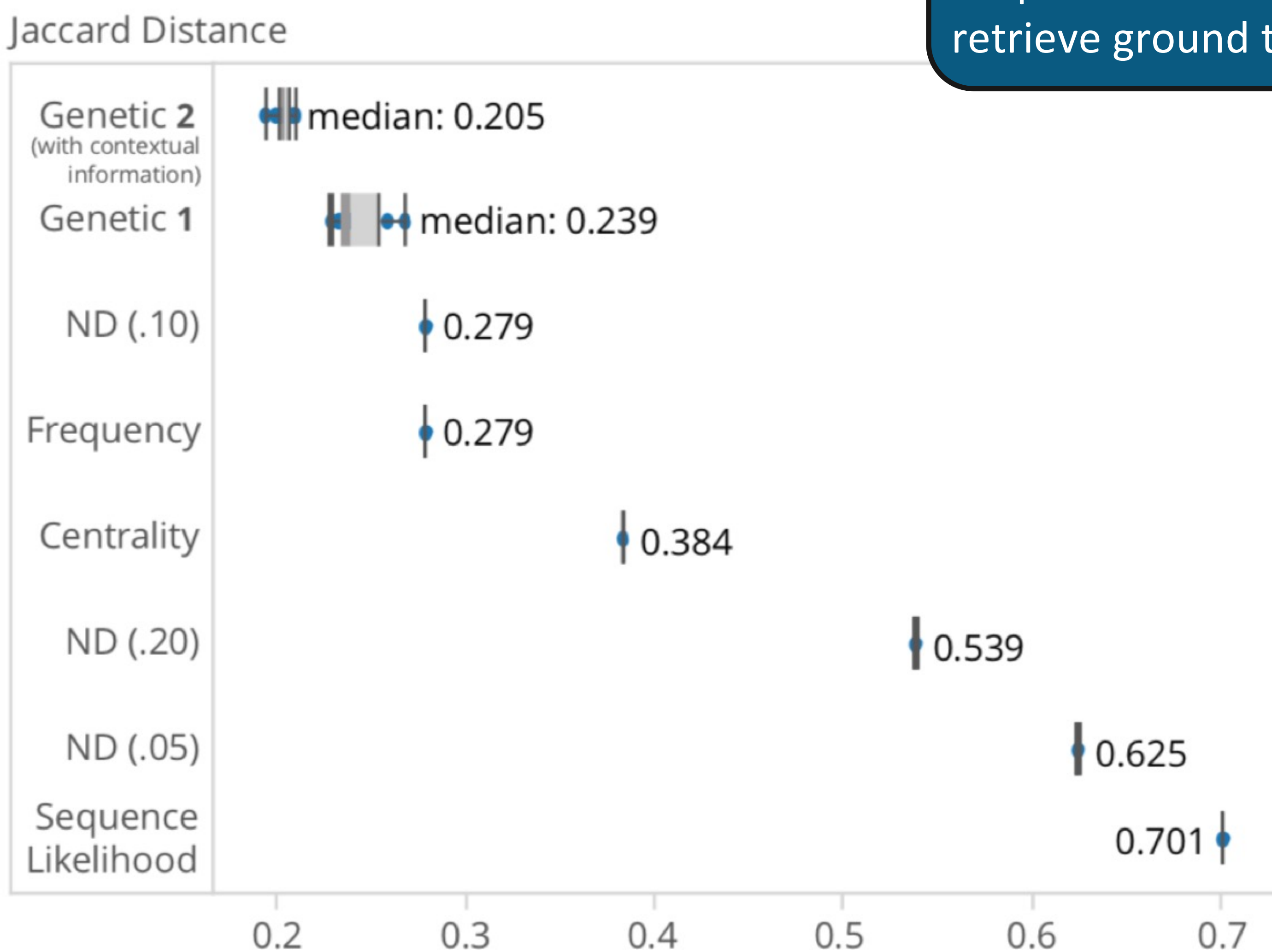
[D] http://customer-journey.unil.ch/datasets/

[41] Gabadinho, A. & Ritschard, G. (2013), "Searching for typical life trajectories applied to childbirth histories", In Levy, R. & Widmer, E. (eds)Gendered life courses - Between individualization and standardization. A European approach applied to Switzerland, pp. 287-312. Vienna: LIT.

# Evaluation

## Results

**Takeaway**
Our approach outperforms Traminer to retrieve ground truth CJMs



# Lessons Learned

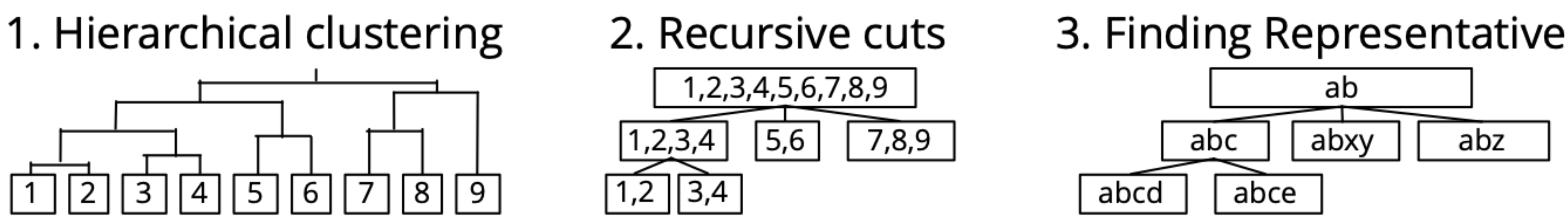
- CJM discovery inspired by process mining
- Domain-agnostic
  - Might be expended for Industry-specific needs.
- Limitation
  - Levenshtein distance is expensive.

# 2 CJM-explorer

[5] Bernard, G., Andritsos, P. (2017). CJM-ex: Goal-oriented exploration of customer journey maps using event logs and data analytics. In Proceedings of the BPM Demo Track and BPM Dissertation Award co-located with 15th International Conference on Business Process Management (BPM 2017). http://ceur-ws.org/Vol-1920/

# Introduction

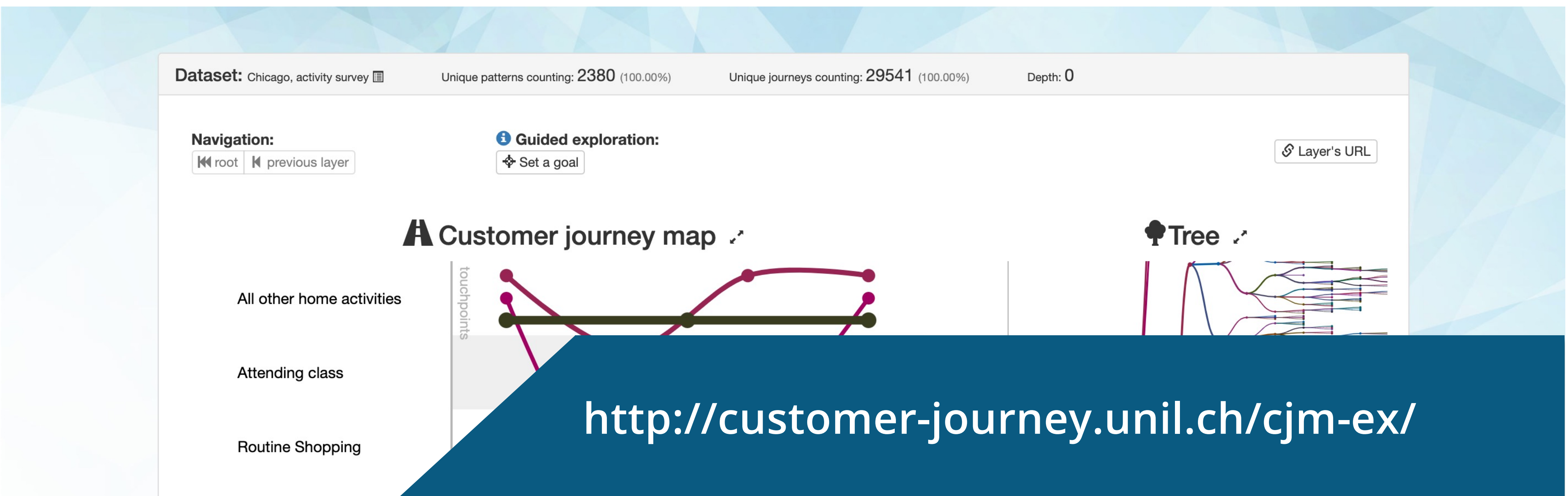
- Web interface to navigate in CJMs
- Hierarchical clustering
  - Top-down navigation



# Contributions

- Three synchronized views
- Goal-oriented navigation

will be shown during the demo



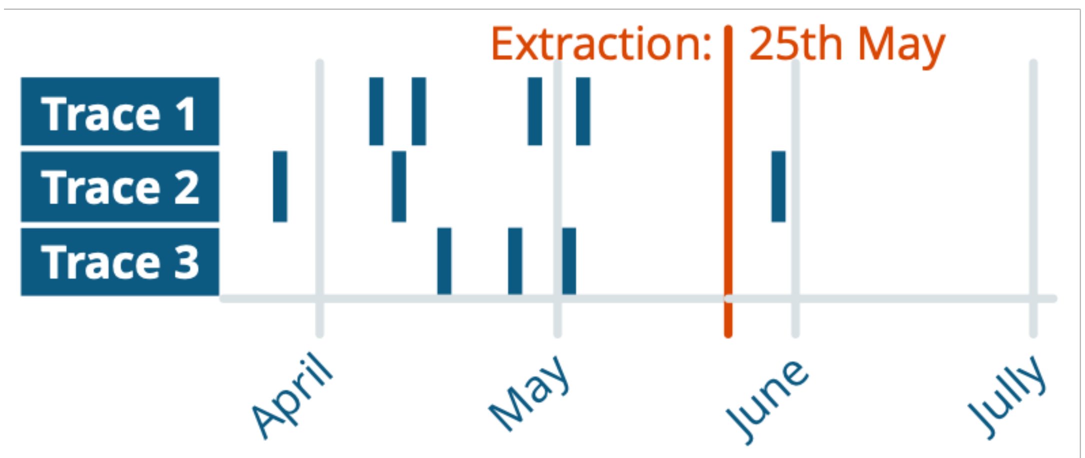

http://customer-journey.unil.ch/cjm-ex/

# 3 *Truncated Trace Classifier*

[E] Bernard, G., Andritsos, P. (2020). Truncated Trace Classifier. Removal of Incomplete Traces from Event Logs. In Business Process Modeling, Development and Support (BPMDS). Springer, Cham. Not published yet.

# Introduction

- Truncated trace
  - Trace where the last events are missing.
- Why?
  - Snapshots challenge [86, chapter 5.3]




[86] Van Der Aalst, W.: Data science in action. In: Process Mining, pp. 3–23. Springer (2016)

# Introduction

**Takeaway**
A TTC can do more than just removing truncated traces.

- Truncated Trace Classifier (TTC)
  Predicts if a trace is truncated

- Why is it useful?
  - Make process models more precise.
  - Repair the process discovery contest [F] event logs.
  - Improve operational efficiency
    e.g., "this ticket is supposed to be closed but is not”.

[F] https://www.tf-pm.org/competitions-awards/discovery-contest

# Baseline

> *Information about case completion may be recorded in the log explicitly, with a* ***dedicated end event****. Otherwise,* ***we need to apply manual rules*** *to filter out incomplete cases. For example, in the Traffic fines log, we consider traces where the last recorded event is Send Fine to be pending and therefore incomplete [92]*

[92] Verenich, I.: Explainable predictive monitoring of temporal measures of business processes. Ph.D. thesis, Queensland University of Technology (2018)

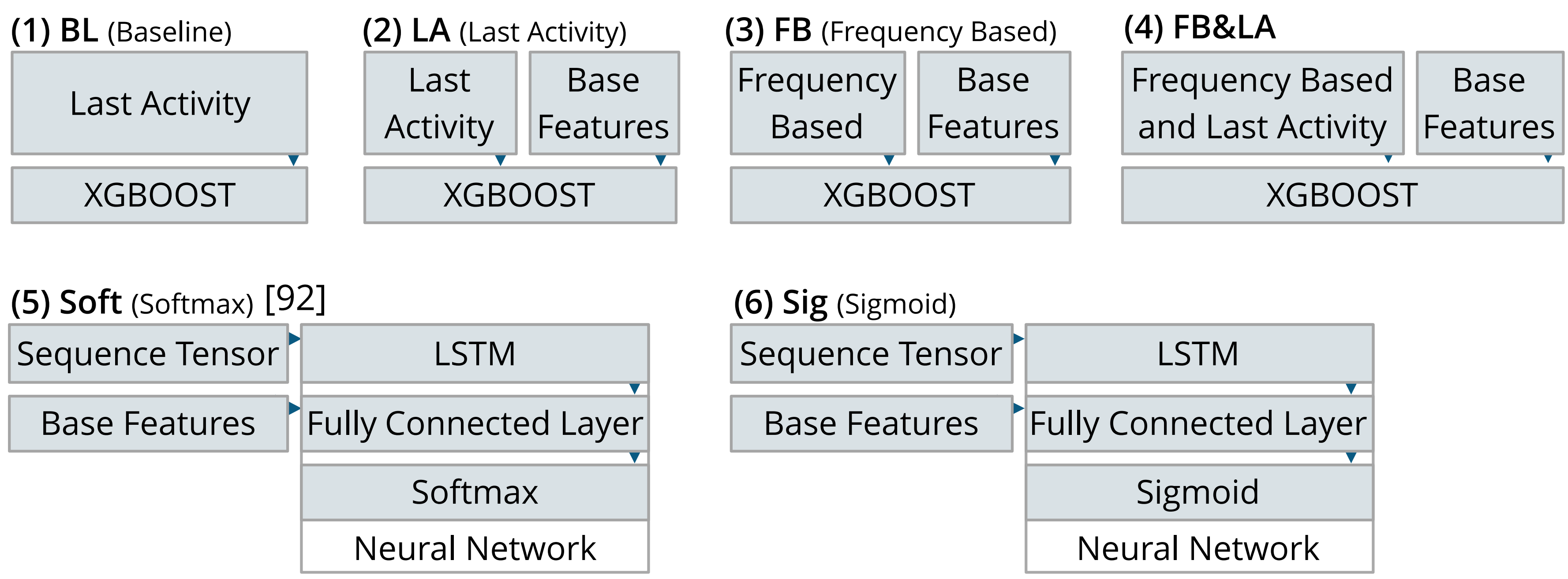
# Labelling

| **Trace** | **Input Sample** | **Target: (Truncated?)** |
|---|---|---|
| <abc> | a | true |
| | ab | true |
| | abc | false |
| <acadd> | a | true |
| | ac | true |
| | aca | true |
| | acad | true |
| | acadd | false |

# Implementations

Base Features

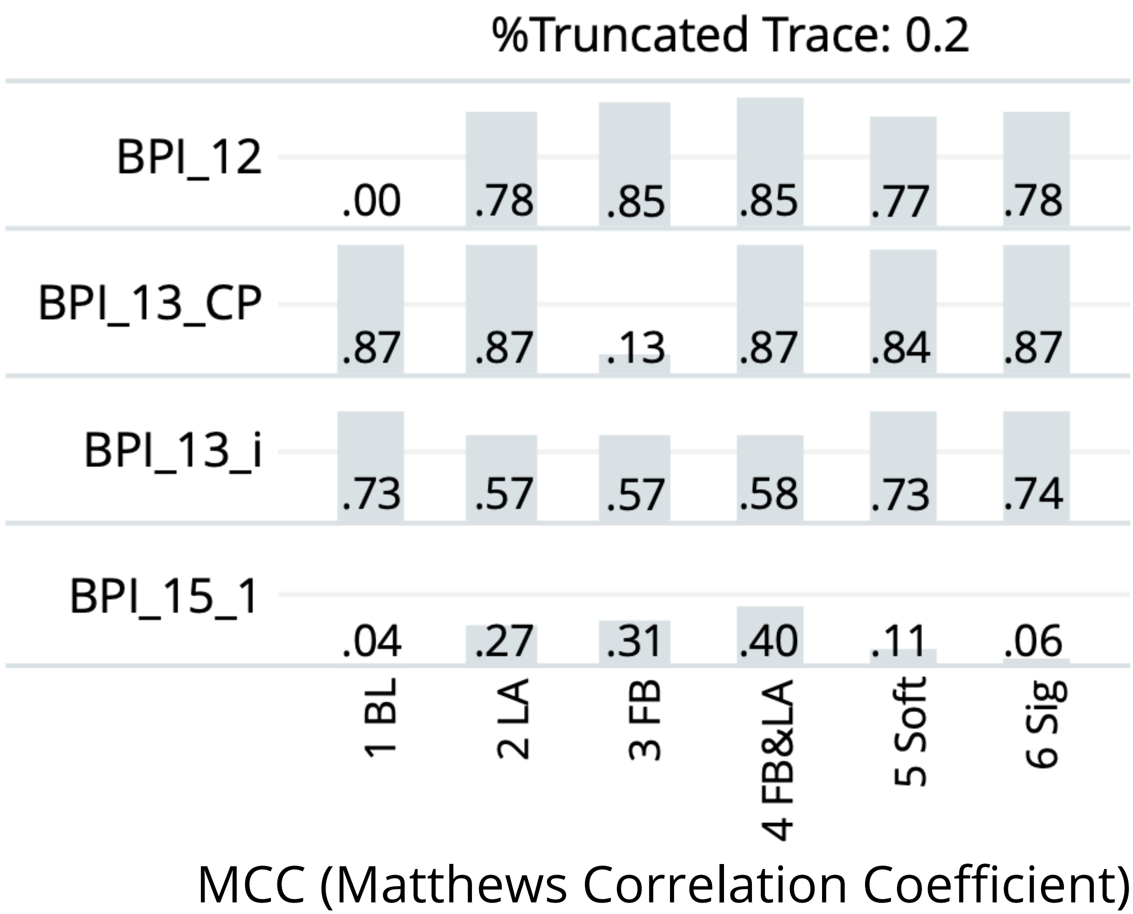
(1) #Activities in the prefix
(2) Seconds since the first event in the trace
(3) Seconds since the previous event in the trace




[92] Tax, N., Verenich, I., La Rosa, M., & Dumas, M. (2017, June). Predictive business process monitoring with LSTM neural networks. In International Conference on Advanced Information Systems Engineering (pp. 477-492). Springer, Cham.

# Results

**Takeaway**
'4FB&LA' achieve highest average MCC



# Lessons Learned

- Relying only on the last activity should not be a default choice.
- Advantage of using a TTC
  - Improving precision of discovered process models.
  - Improving next event predictions.

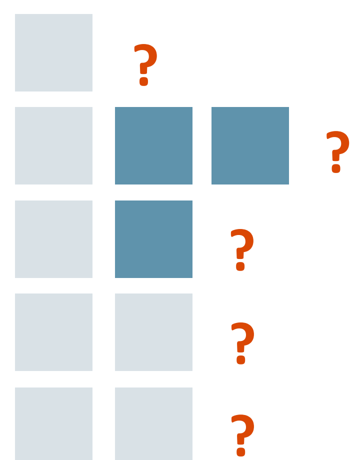
**Best Paper**

# 4 *Path Prediction*

[8] Bernard G., Andritsos P. (2019) Accurate and Transparent Path Prediction Using Process Mining. In: Welzer T., Eder J., Podgorelec V., Kamišalić Latifić A. (eds) Advances in Databases and Information Systems. ADBIS 2019. Lecture Notes in Computer Science, vol 11695. Springer, Cham. https://doi.org/10.1007/978-3-030-28730-6_15

# Introduction

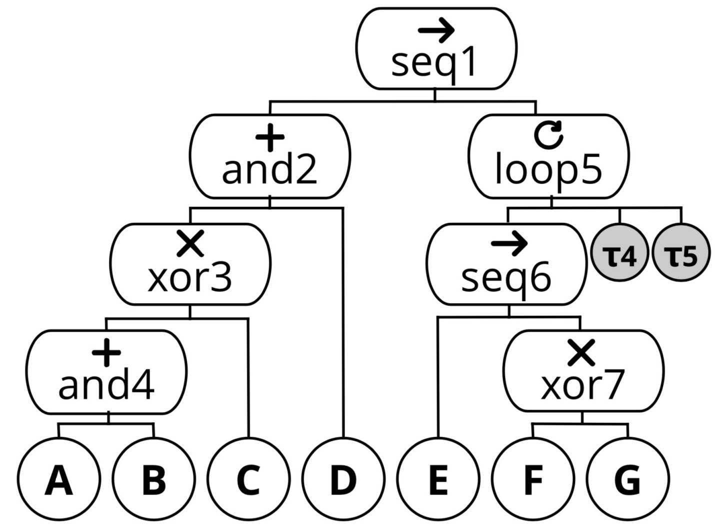
**What are the most likely next touchpoints?**



# Algorithm

## Building the footprint

Inductive Miner [58]




LaFM

Loop aware Footprint Matrix

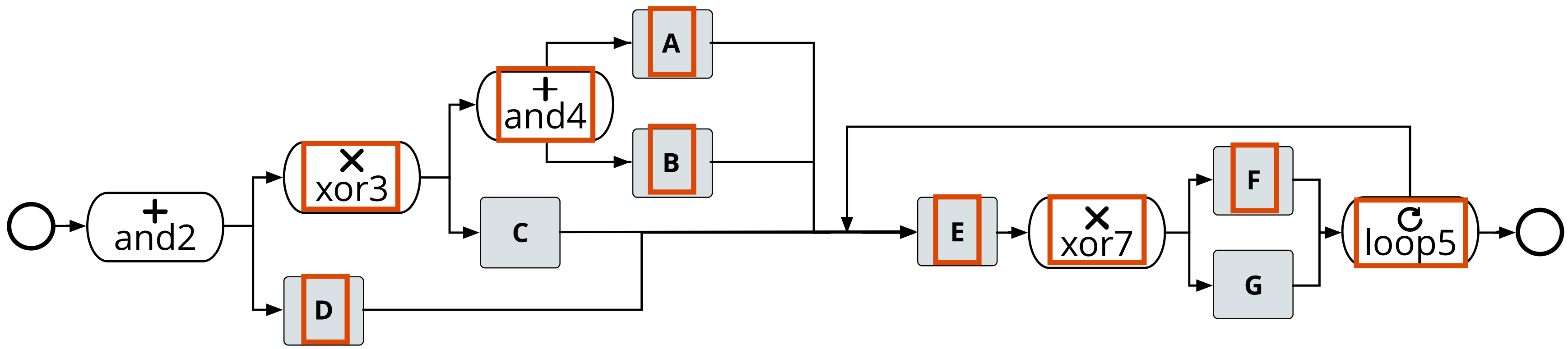
| Traces | and2(1)\| | and2(2)\| | and2(3)\| | and4(1)\| | and4(2)\| | loop5\| | xor7\|loop5{1} | xor7\|loop5{2} | xor3\| |
|---|---|---|---|---|---|---|---|---|---|
| ABDEF | 1 | 1 | 2 | 1 | 2 | 1 | 1 | ∅ | 1 |
| BDAEGEF | 1 | 2 | 1 | 2 | 1 | 2 | 2 | 1 | 1 |
| DCEFEG | 2 | 1 | ∅ | ∅ | ∅ | 2 | 1 | 2 | 2 |
| ... | | | | | | | | | |

[58] Leemans, S.J., Fahland, D., van der Aalst, W.M.: Discovering block-structured process models from event logs-a constructive approach. In: International conference on applications and theory of Petri nets and concurrency. pp. 311–329. Springer (2013)

# Algorithm

## Making predictions

Prefix: D
Predicted Path: A B E F E F

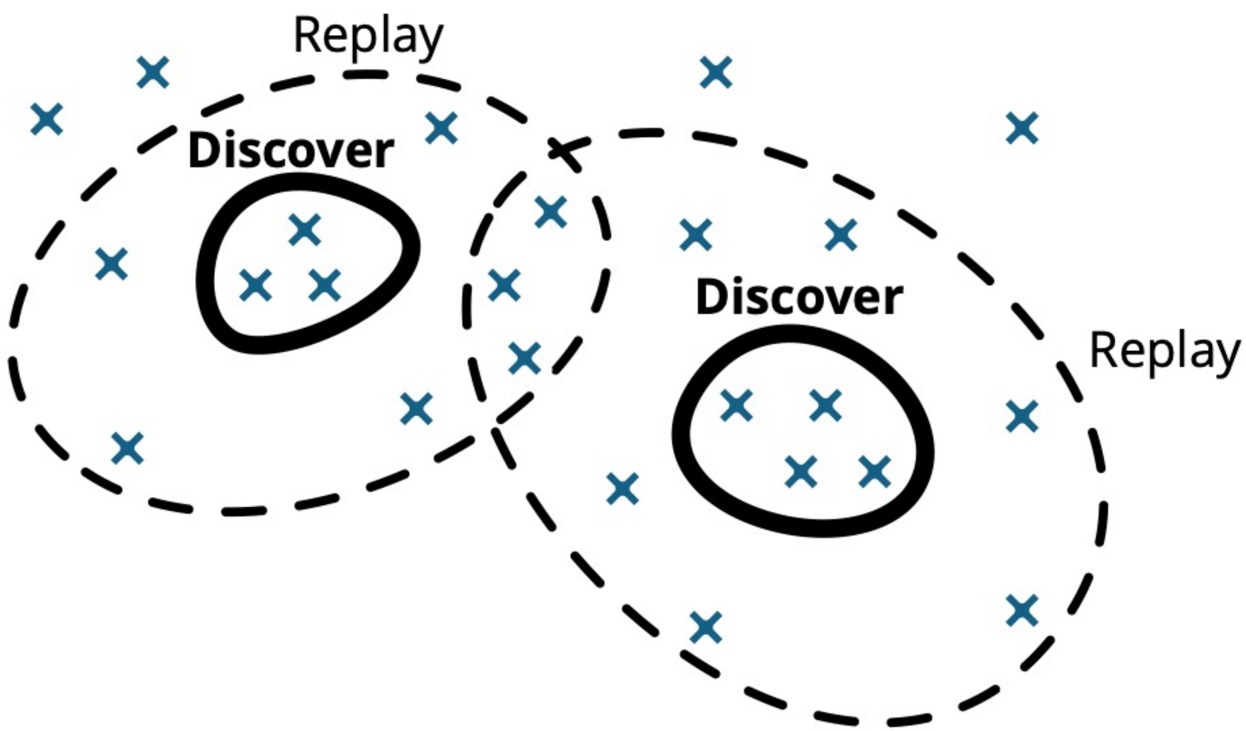



| Traces | and2(1)\| | and2(2)\| | and2(3)\| | and4(1)\| | and4(2)\| | loop5\| | xor7\|loop5{1} | xor7\|loop5{2} | xor3\| |
|---|---|---|---|---|---|---|---|---|---|
| ABDEF | 1 | 1 | 2 | 1 | 2 | 1 | 1 | ∅ | 1 |
| BDAEGEF | 1 | 2 | 1 | 2 | 1 | 2 | 2 | 1 | 1 |
| DCEFEG | 2 | 1 | ∅ | ∅ | ∅ | 2 | 1 | 2 | 2 |
| … | | | | | | | | | |

# Clustered LaFM

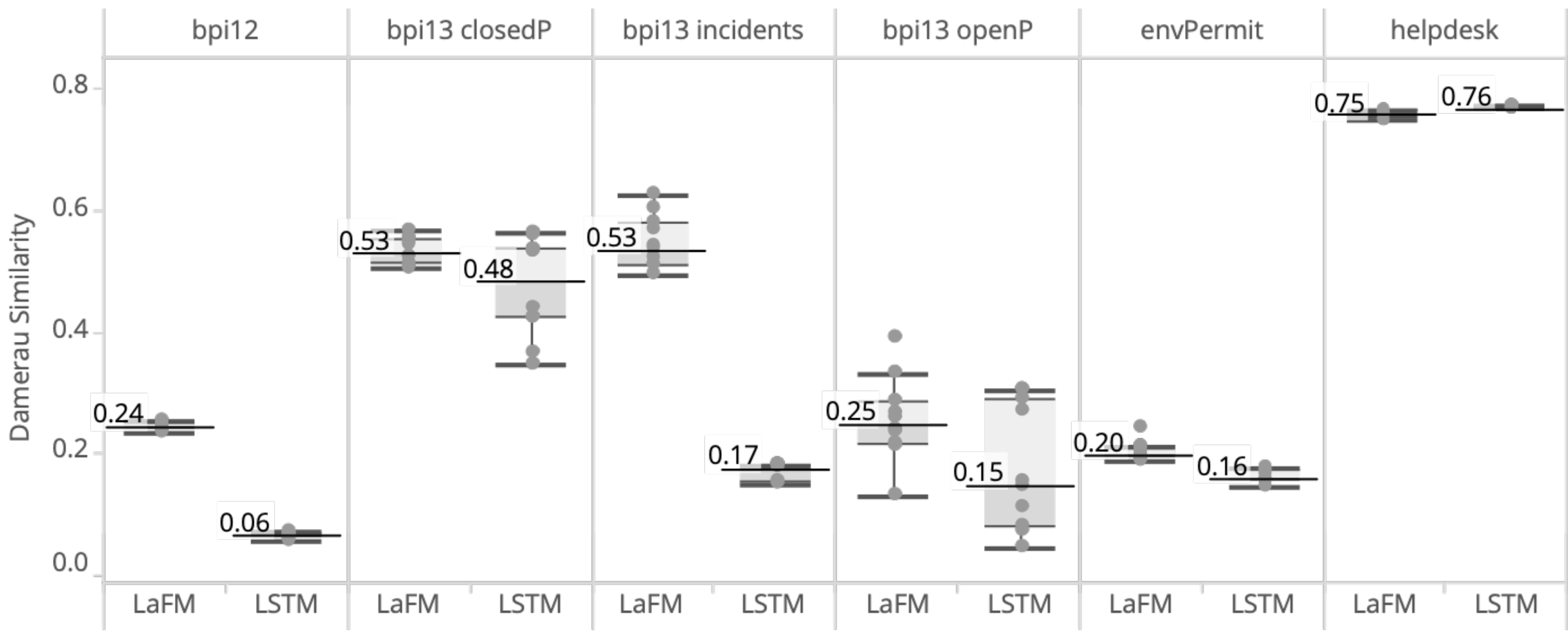
## For complex event logs

**Takeaway**
Soft clustering is the key to handle complex event logs



# Clustered LaFM

## For complex event logs



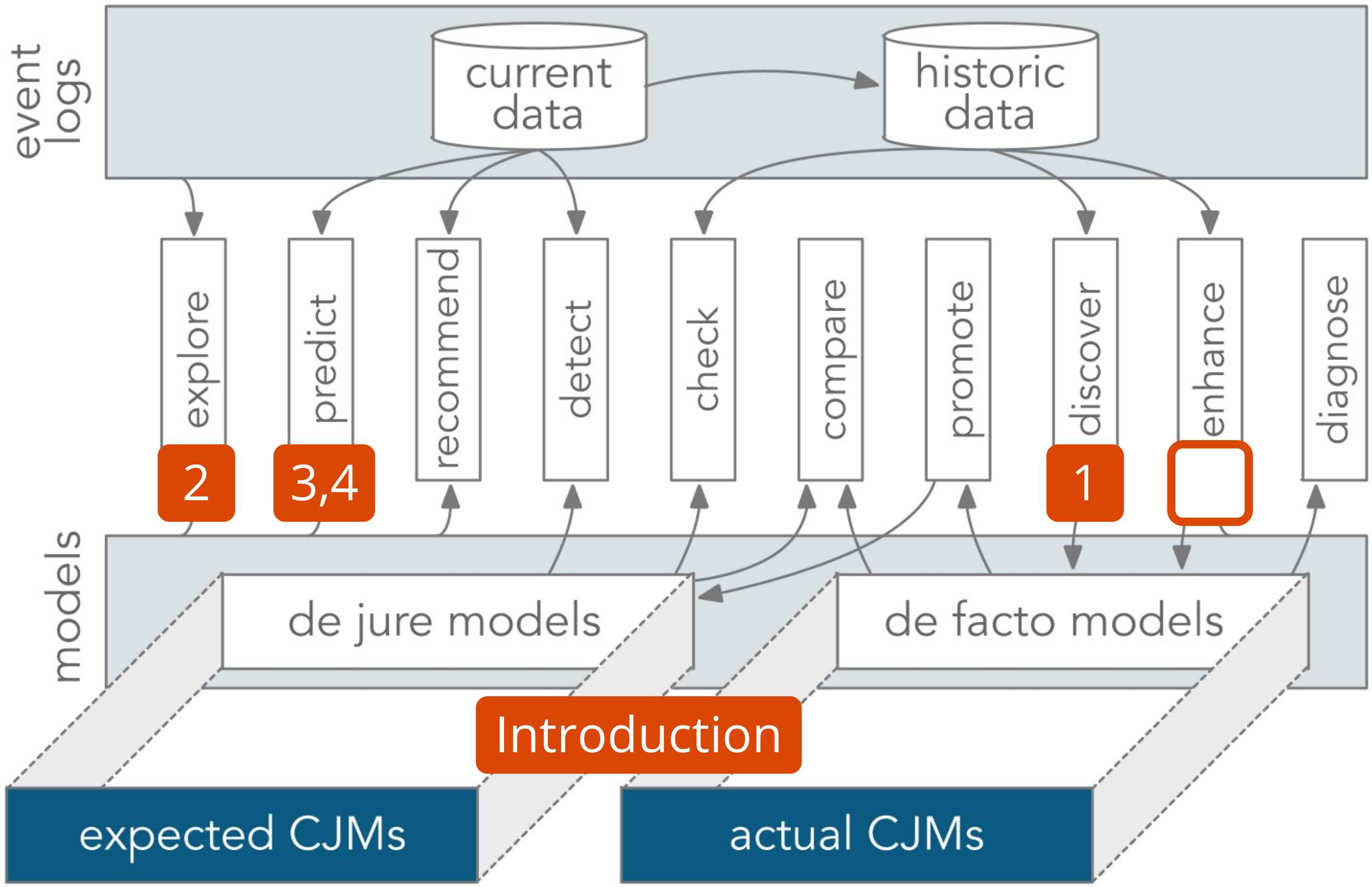
# Lessons Learned

- Transparent predictions
  - Accurate
  - Interpretable output
- Soft clustering
  - Works with complex event logs

# *Conclusion*

# Contributions

## Adjusted process mining framework [86]



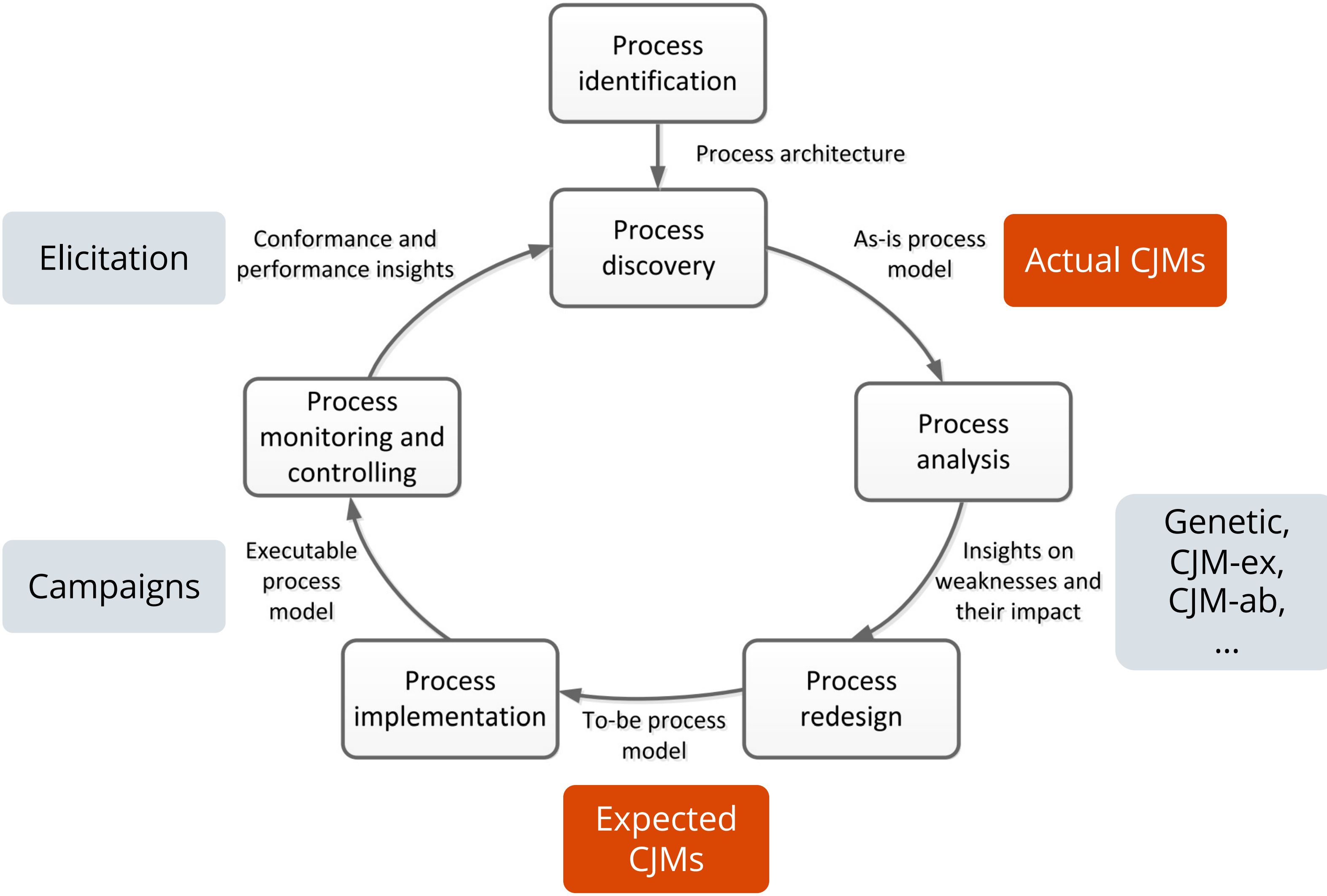

[86] Original framework: van der Aalst, W. (2016). Process Mining: Data Science in Action. Springer.

Adjusted: [6] Bernard, G., Andritsos, P. (2017). A process mining based model for customer journey mapping. In: Forum and Doctoral Consortium Papers Presented at the 29th International Conference on Advanced Information Systems Engineering . CAiSE 2017 (Forum) (Vol. 1848, pp. 49-56). CEUR Workshop Proceedings.

# Future Works

- Work-in-progress
  - Partitionning of unlabeled customer journeys.
- Investigate the relevance of each process mining activities for customer journeys.
- New visualization «VisuEL», Best Demo Award

# Future Works




[36] Dumas, M., La Rosa, M., Mendling, J., & Reijers, H. A. (2013). Fundamentals of business process management (Vol. 1, p. 2). Heidelberg: Springer.