



University of Groningen

ECIDA

Blaauw, Frank; Overbeek, Roy; Albers, Toon; Vlek, Jeroen; Maessen, Mario; Gooijer, Jan; Lazovik, Elena; Arbab, Farhad; Lazovik, Alexander

10.13140/RG.2.2.33143.47524

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version Publisher's PDF, also known as Version of record

Publication date: 2019

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):

Blaauw, F., Overbeek, R., Albers, T., Vlek, J., Maessen, M., Gooijer, J., Lazovik, E., Arbab, F., & Lazovik, A. (2019). *ECiDA: Evolutionary Changes in Data Analysis*. Poster session presented at ICT.Open, Hilversum, Netherlands. https://doi.org/10.13140/RG.2.2.33143.47524

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

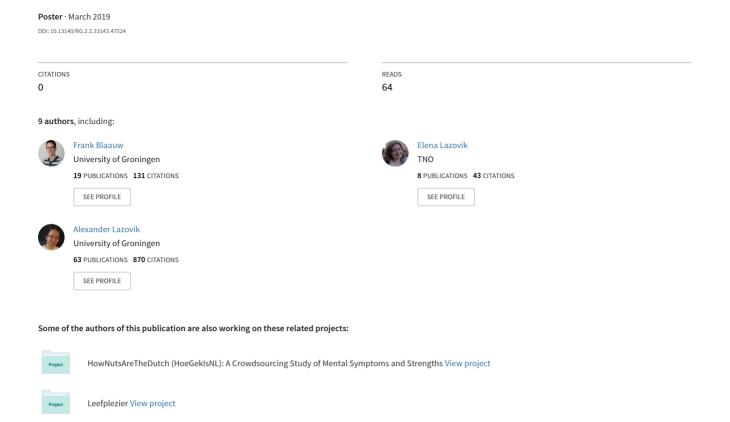
The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: https://www.rug.nl/library/open-access/self-archiving-pure/taverneamendment.

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Download date: 07-06-2022

ECiDA: Evolutionary Changes in Distributed Analysis



Evolutionary Changes in Distributed Analysis



The project

Current data analysis platforms all too often rely on the fact that the analyis and underlying data is static. In reality this is nearly never the case: data scientists come up with new methods to analyze data all the time, and data sources are almost by definition dynamic.

- ECiDA revolves around three main concepts:
- **Dynamicity:** components in ECiDA must be dynamic.
- Consistency: changes in ECiDA can never result in a faulty state.
- Abstraction: ECiDA should offer a usable system, hiding any unneccessary complexity.

Procedure





1. Research question

Data scientist comes

up with a research

topic.





2. Building blocks Data scientist selects a number of relevant building blocks in ECiDA and specifies their relationship.





4. Results 3. Pipeline Results at the sink of ECiDA uses artificial intelligence and conthe pipeline can be analyzed and used to straint programming to build a pipeline. steer decision making.

After the initial pipeline has been built, the network topology can always be modified.

Architecture of change Information Machine Analysis **Data sources** pipelines sinks learning **Dynamic pipeline Automatic verification**

Collaboration

ECiDA is being built as a general purpose data science tool. If you would like to learn more, or would like to collaborate, feel free to contact us at:

f.j.blaauw@rug.nl

Use cases

- Water top sector: Water distribution monitoring and automation.
- Life Sciences and Health top sector: Prediction and maintaining of water quality.
- HTSM/Smart Industry top sector: Structural reliability of pipes.

Timeline



Start project Inception of ECiDA, illuminate all paths and set up a roadmap.



Requirements phase Elicitate the requirements of the different stakeholders in ECiDA.

'20)

'20



Design phase

Provide initial architecture design and technical use case descriptions. Implementation of first predictive model.



Algorithm phase

Implement a package for detecting inaccurate predictions and a package for establishing chemical fingerprints in water.



Development phase 1

Implement language extensions for automated consistency checks and create initial version of ECiDA. Implement network simulation and an algorithm for structural quality of pipes.



Development phase 2

Provide dynamic data processing as a service, implement use cases and design anomaly detection algorithms.



Development phase 3

Improve language extensions, build ECiDA platform, and implement algorithm to optimize water distribution networks.



Application phase

Apply ECiDA to the use cases, finalize last implementation details of ECiDA.



Finalize project

Finish project and its documentation.



Partners





Netherlands Organisation NWO for Scientific Research









universityof groningen





