How to make reliable decisions from data: The ExtremeXP paradigm

Iyad Alshabani

CEO, R&D Investigator, Manager

Iyad.alshabani@bitsparkles.com

Sophia Antipolis Sep. 19th 2025

BitSparkles — Driving Digital Transformation & Innovation

Who We Are

- SME
- Digital Transformation
- Consultancy

How We Create Value

- From Data
- Research and Innovation Management
- Technology transfer
- Open innovation
- Spin-off research results

What We Do

- Digital Transformation
- AI Process & Infrastructure (AIOps)
- Data Sharing & Governance
- Automation & Integration

On-going Projects

- ExtremeXP
  - Experiment driven and user eXPerience oriented Analytics for eXtremely Precise outcomes and decisions
  - 2023-2026
  - Architecture modelling, Exploitation, Innovation and business modelling
  - https://extremexp.eu/
- CIPHER
  - Cybersecurity Intelligence, Protection, and Holistic Enterprise Resilience
  - 2025-2028
  - Data Management, Exploitation, innovation and business model

Experiment driven and user eXPerience oriented

Analytics for eXtremely Precise outcomes and decisions

The ExtremeXP paradigm: Humans in the center of the AI processes for explainable, accurate, precise and fit for purpose insights

ExtremeXP

Funded by the European Union

Co-funded by the European Union Horizon Programme Call HORIZON-CL4-2022-DATA-01-01, under Grant Agreement No. 101093164

ExtremeXP Vision

To provide

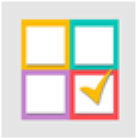
accurate, precise, fit-for-purpose, and trustworthy data-driven insights

via

evaluating different complex analytic variants

considering

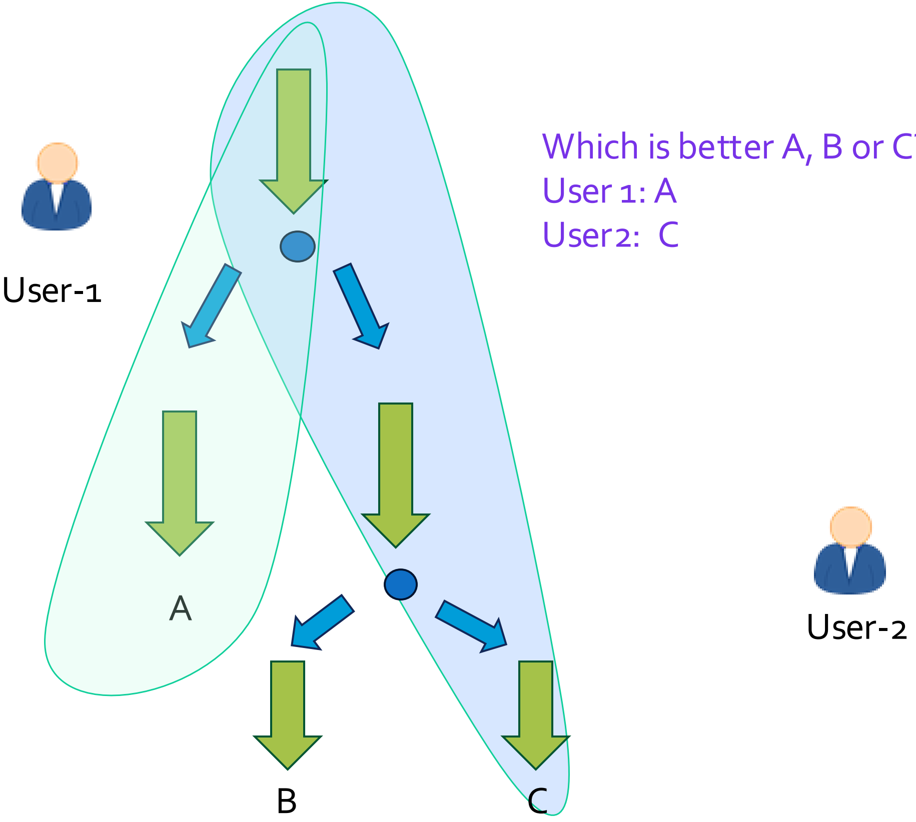
diverse user intents, constraints and feedback



Funded by the European Union

Decision Making: The Data-driven impact (ML-based)

- Challenges:
  - Extreme data scale, low quality, different modalities, and ownership.
  - Need for accurate and precise analytics.
  - Trustworthiness is essential for decision adoption.
  - Human factor and purpose-fit.
- Solution: learn from experimenting different configurations, models, data sources/data sets, user profiles, user feedback, previous experiments



Which is better A, B or C?
User 1: A
User2: C

Funded by the European Union

Example

Objective:

- Detect imminent failures in machine electrical and mechanical components using deep learning models

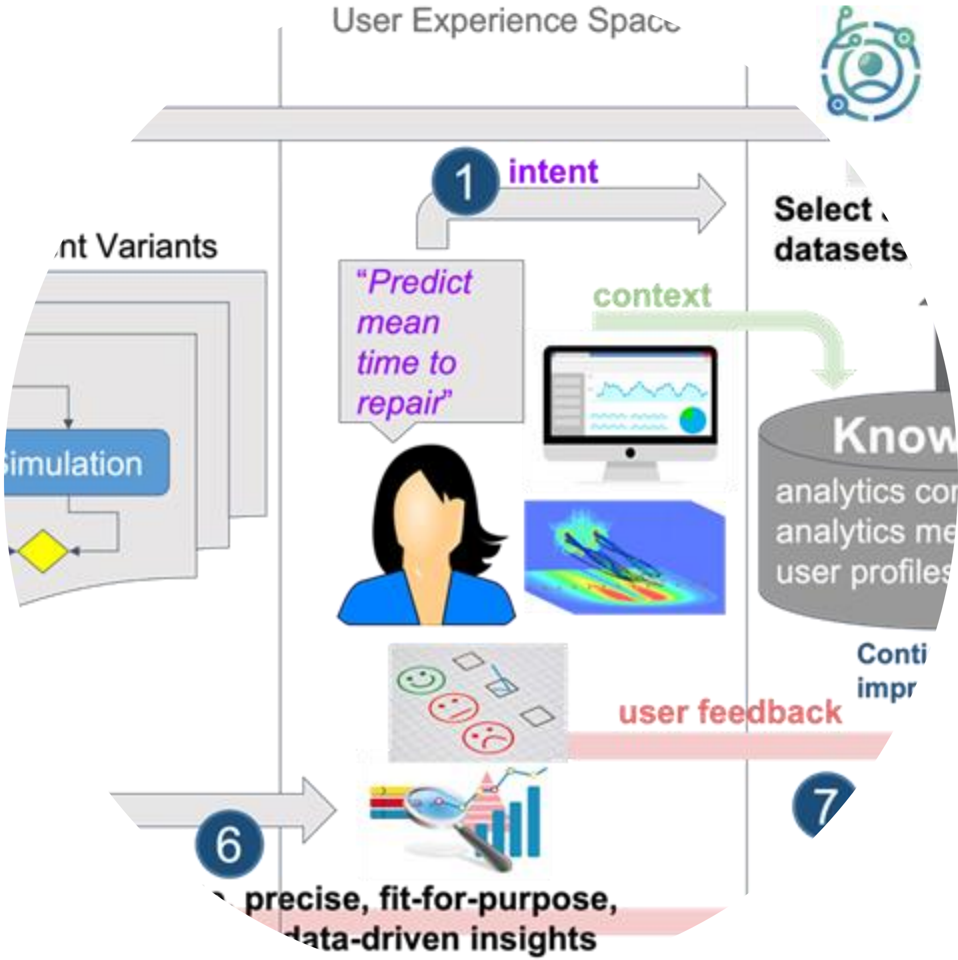
Methodology:

- Analyze linear (backward-forward) movements of the machine axes

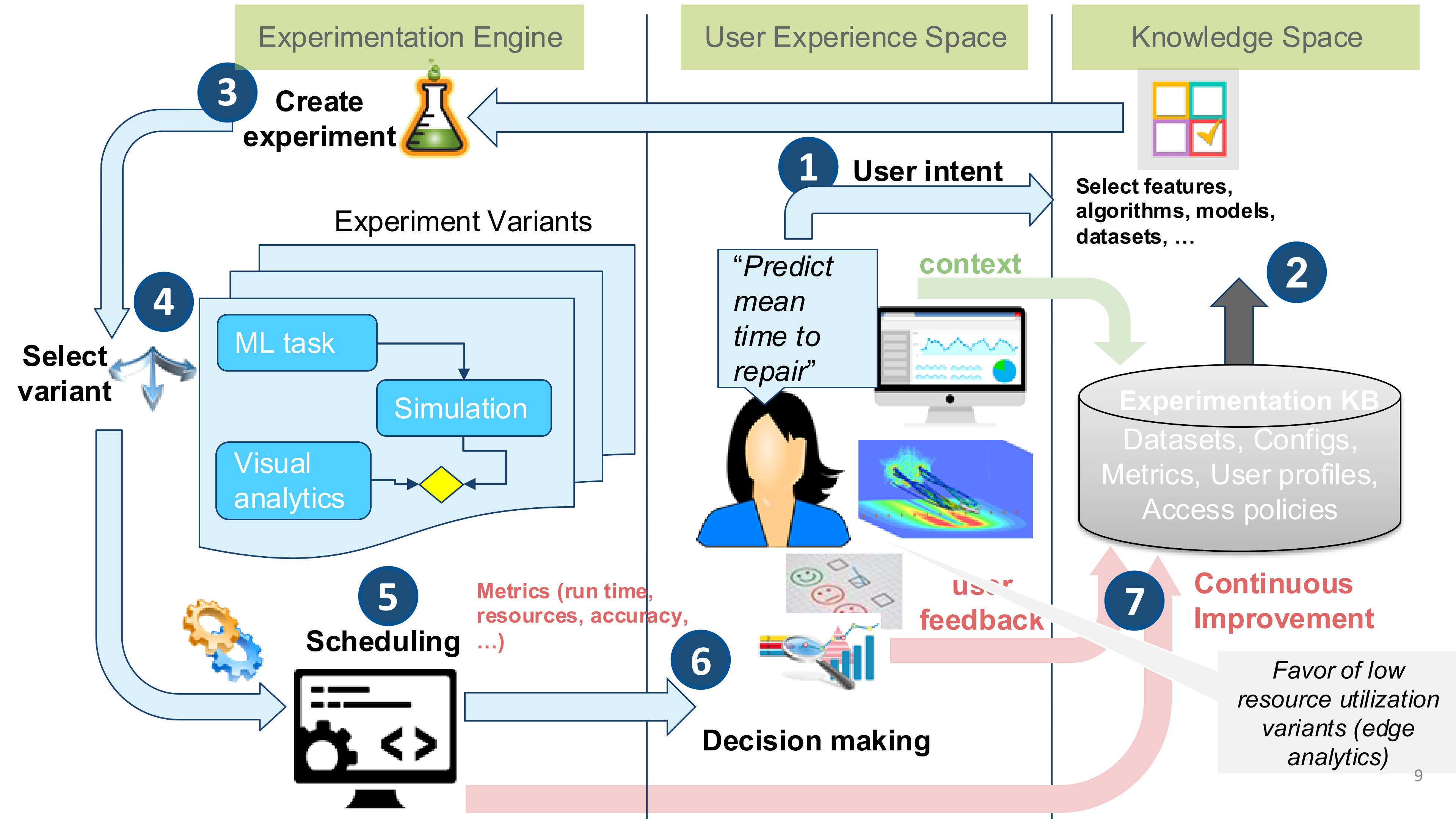
Funded by the European Union

ExtremeXP Concept

- A **human-centered approach** to AI and data driven analytics
- **Experimentation** is the core concept for generating extremely accurate analytics
- **Optimise the properties of a complex analytic process** (e.g., accuracy, time-to-answer, specificity, recall, precision, resource utilization) **by associating different user profiles to computation variants**.



Funded by the European Union

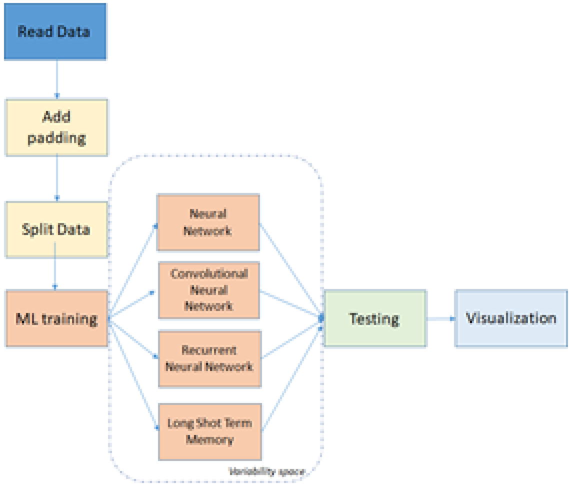


Experimentation Engine
User Experience Space
Knowledge Space
3
Create experiment
1
User intent
Select features, algorithms, models, datasets, …
Experiment Variants
"Predict mean time to repair"
context
2
4
ML task
Simulation
Visual analytics
Select variant
Experimentation KB
Datasets, Configs, Metrics, User profiles, Access policies
5
Scheduling
Metrics (run time, resources, accuracy, …)
user feedback
7
Continuous Improvement
6
Decision making
Favor of low resource utilization variants (edge analytics)

Experimentation approach to optimize complex analytics– Main Concepts

User Intent: Classify linear (backward-forward) movements of the machine axes as Electrical or Mechanical failures under the **constraint of >95%** accuracy

Experiment Design (Workflow)



Experiment Variants

- Task Variants
  - Neural Network (NN)
  - Convolutional Neural Network (CNN)
  - Recurrent Neural Network (RNN)
  - LSTM (Long Short Term Memory)
- Experimentation space per task
  - Model parameters
    - Number of layers
    - Number of nodes or units of each layer
    - Activation function
  - Training parameters
    - Number of epochs
    - Batch size

Generate and execute different workflow variants, based on the user intent (classify) and constraints, visualize data and metrics and gather user feedback

Funded by the European Union

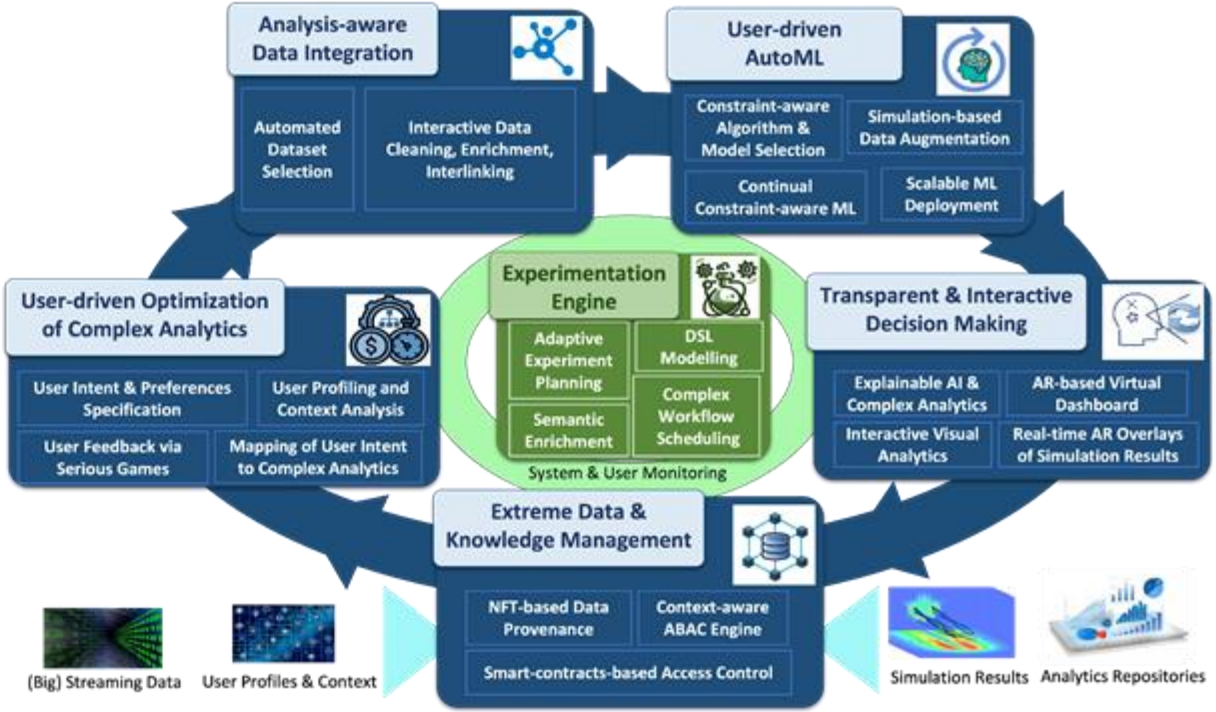
Yet another Auto-ML framework?

Three key characteristics

- **User involvement** at different phases of the experimentation –
  - **Human Tasks VS Automated Tasks**
    - At the beginning (specify intents, user contraints, access control policies)
    - During the execution of an experiment (view results, cancel/prioritize/specify new workflows, but also provide expert feedback on the results of a workflow run)
    - At the end (help encode experiment results in knowledge base for future use)
- Experimentation considers **AI training pipelines**, but **also other types of workflows** (data analytics, simulation, visualization), including hybrid ones
- Our framework comprises strategies for automating the execution and evaluation of workflows, but also **manages the knowledge created by experiments** (to better inform future ones)

Funded by
the European Union

ExtremeXP Framework



Funded by the European Union

ExtremeXP Added Value

fit-for-purpose

trustworthy

accurate

precise

ICT-related industry, SMEs, Open-source communities

Researchers and academia (Cloud, SW, AI, Data)

Public bodies and policy makers, european stakeholders

Citizens, End users of AI Tools

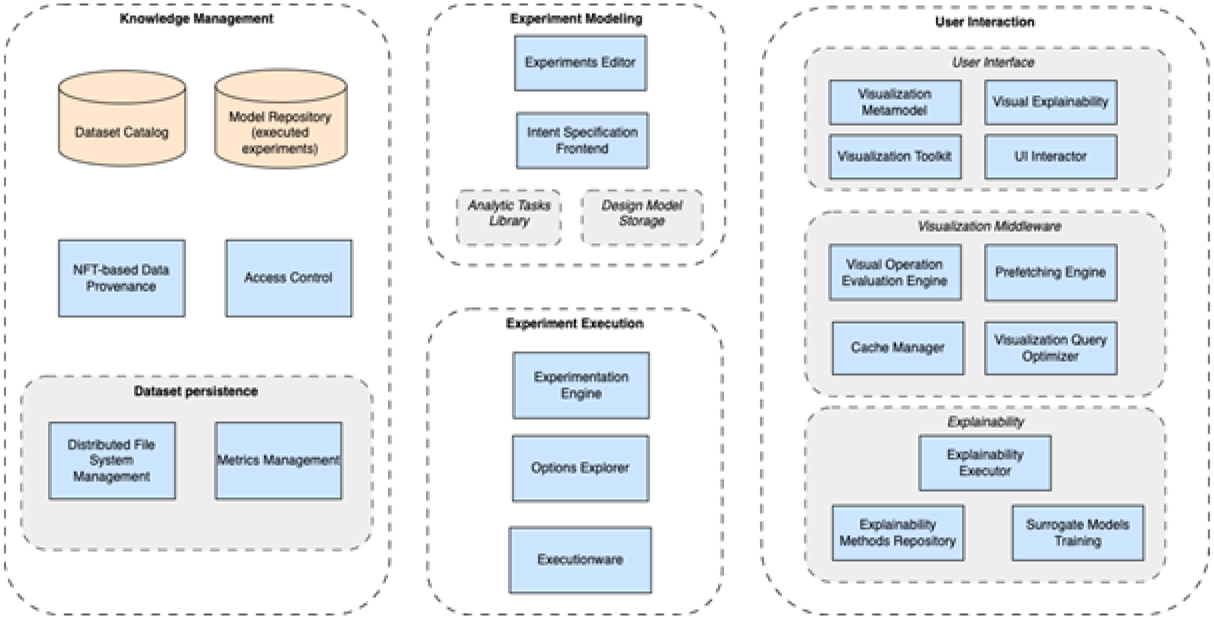
13

Funded by the European Union

ExtremeXP Details

O1: Specification and semantics for modelling complex user-experience-driven analytics

- **ExtremeXP Framework Architecture**



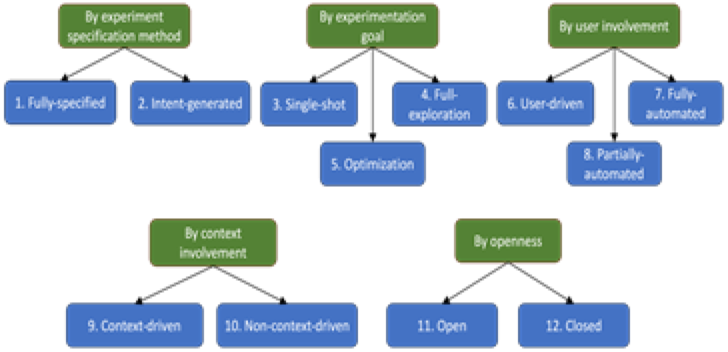
- **Data engineer**
- **Domain expert**
- **Data creator**

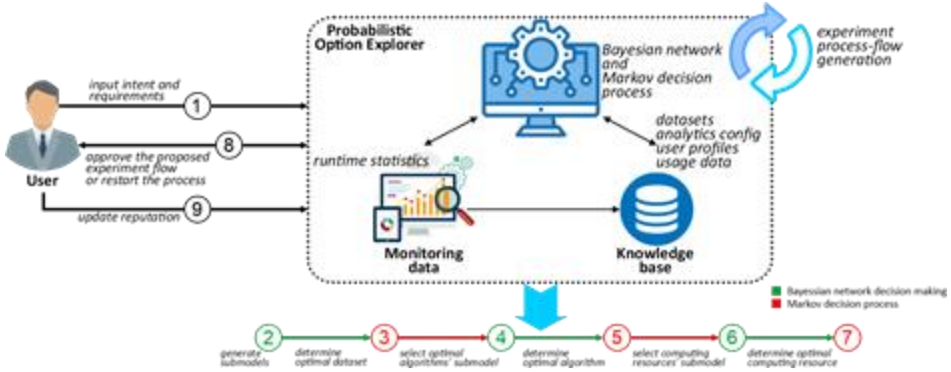
Funded by the European Union

ExtremeXP Details

O1: Specification and semantics for modelling complex user-experience-driven analytics

- **Foundational Concepts of ExtremeXP – Types of experiments**
- **Initial modelling and language foundations for experiment-driven analytics**
- **Traceability and trustworthiness for experiment-driven analytics**
- **Probabilistic Options Explorer for user centric experiment optimization**

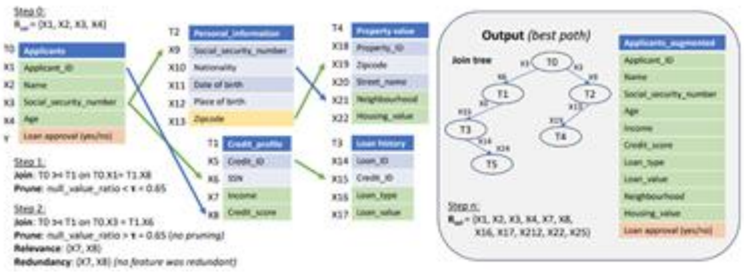


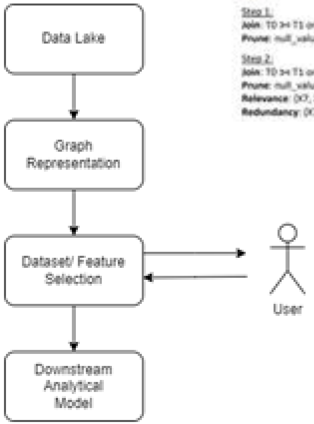


Funded by the European Union

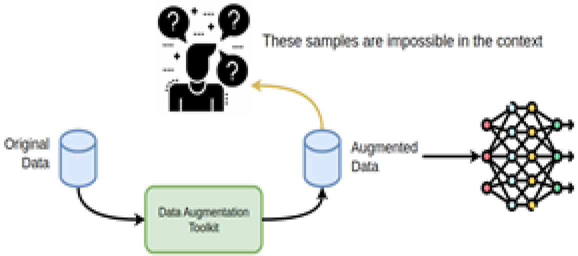
ExtremeXP details

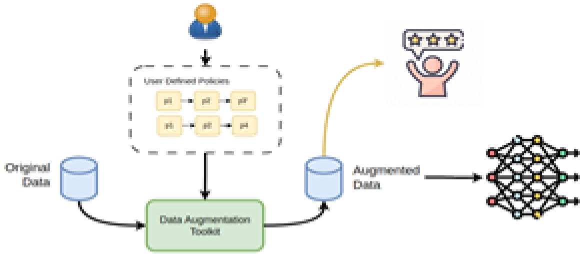
Automated and scalable data management for complex analytics workflows





- **Dataset selection and** feature discovery, which enhances machine learning models by identifying and augmenting base tables with relevant, high-predictive power features
- **Data Integration** on GPUs. Use of BERT models for entity resolution and data interlinking during user analysis
- **Data Augmentation** toolkit designed to generate an augmented dataset derived from an original dataset

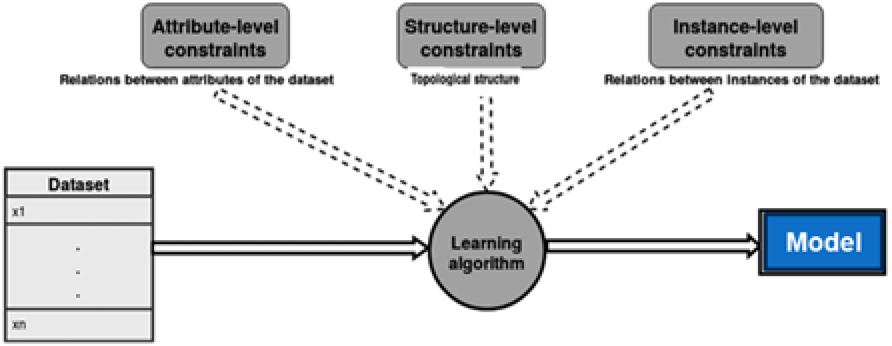




Funded by the European Union

ExtremeXP details

Scenario-driven and opportunistic machine learning

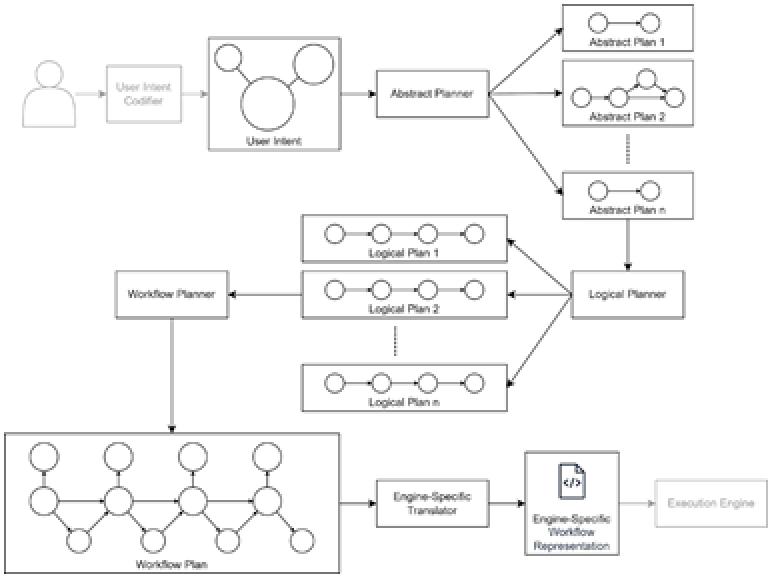


- **Clustering constraints** resulting in a bi-objective loss function to produce a better clustering assignment
- **Task-agnostic continual learning** (TACL) to accommodate the incremental nature of the experimental problems covering the needs of the use cases.

Funded by the European Union

ExtremeXP Details

O4: User-experience- and experiment-driven optimization of complex analytics



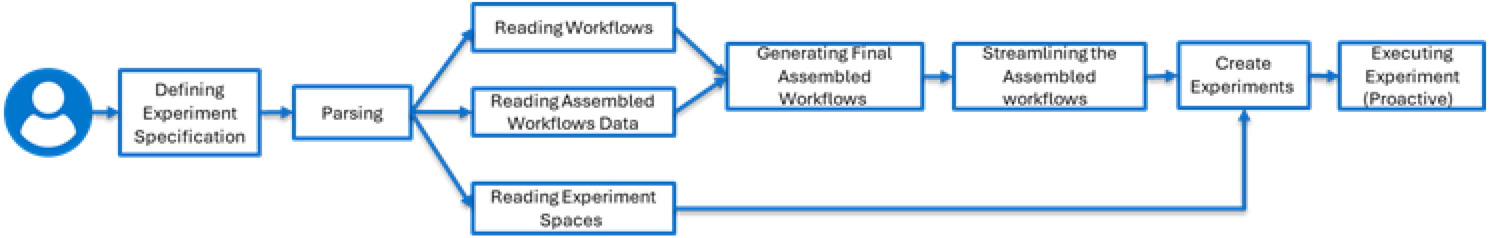
- Capturing user intents
  - **user intents, preferences and constraints**, powered by a knowledge graph.
- Mapping user intents to analytic workflows
  - **generates all possible workflows** based on a basic user intent

Funded by the European Union

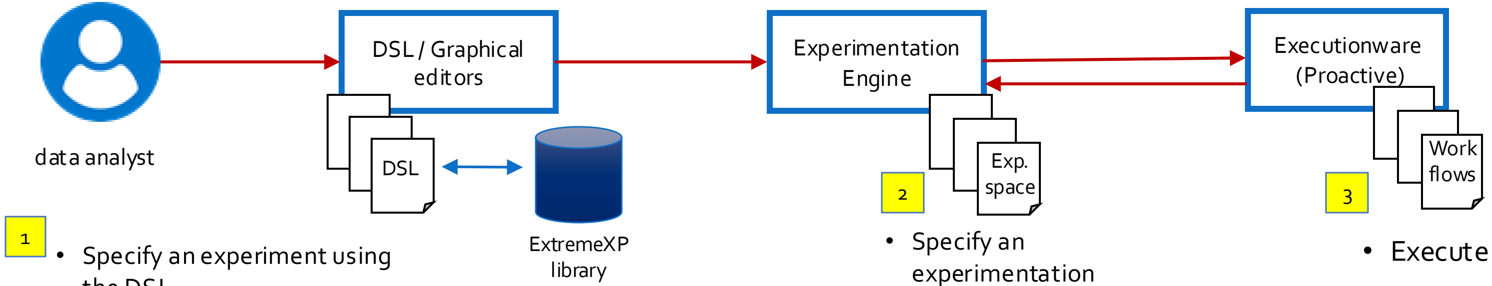
ExtremeXP Details

User-experience- and experiment-driven optimization of complex analytics

- Experimentation Engine
  - **experiment planning**
  - Tool for data **monitoring** of ExtremeXP engine metrics



Workflo
w



Experiment

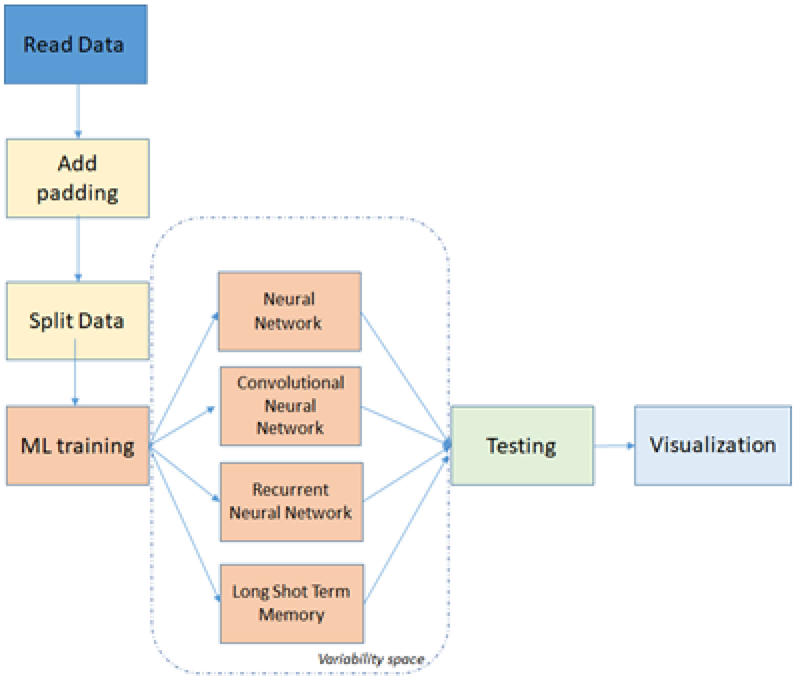
- Specify an experiment using the DSL
- Either from scratch or via reuse from the library

- Specify an experimentation space
- Generate workflows based on the space

- Execute workflows

Funded by the European Union

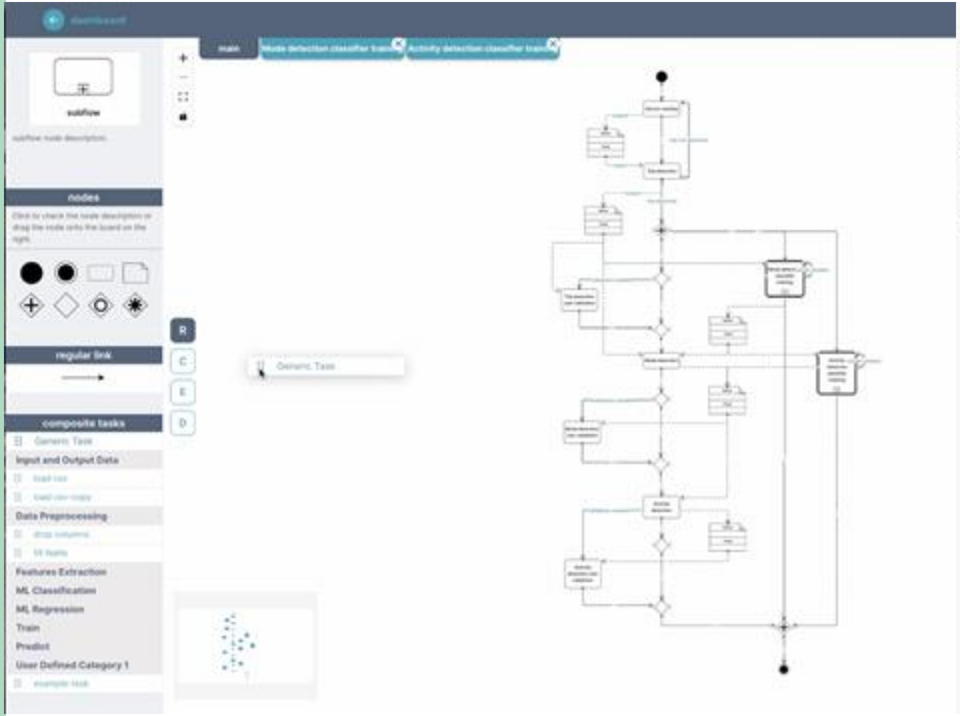
Defining Experiments using Graphical Editor & Domain Specific Language

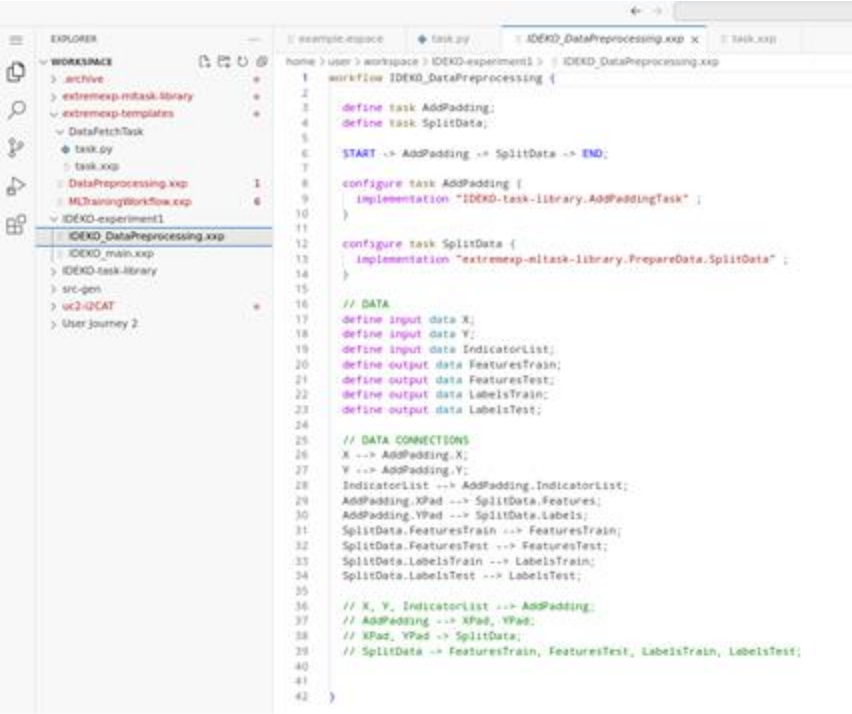



```
workflow IDEKO_V1 {

  define task ReadData;
  define task AddPadding;
  define task SplitData;
  define task TrainModel;

  START -> ReadData -> AddPadding -> SplitData -> TrainModel  -> END;

  configure task ReadData {
    implementation "tasks/IDEKO/read_data.py";
    dependency "tasks/IDEKO/src/**";
  }

  configure task AddPadding {
      implementation "tasks/IDEKO/add_padding.py";
      dependency "tasks/IDEKO/src/**";
    }

  configure task SplitData {
      implementation "tasks/IDEKO/split_data.py";
      dependency "tasks/IDEKO/src/**";
  }

  configure task TrainModel {
      dependency "tasks/IDEKO/src/**";
  }
```

Funded by the European Union

Designing tools

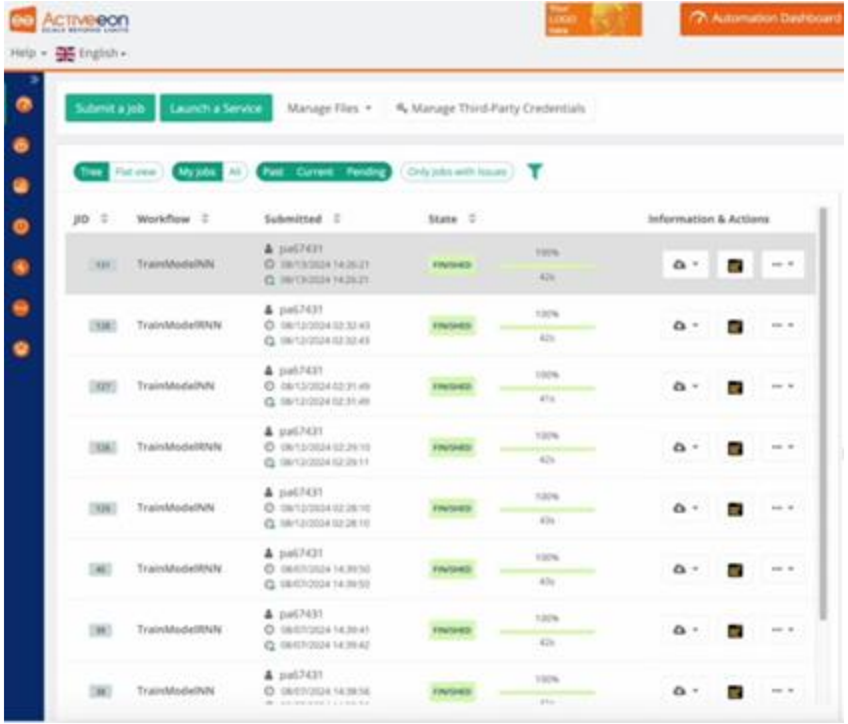




Funded by the European Union

Execution in ProActive

- Job Scheduling & Workload Automation
- Scalability
- Easy integration



Funded by the European Union

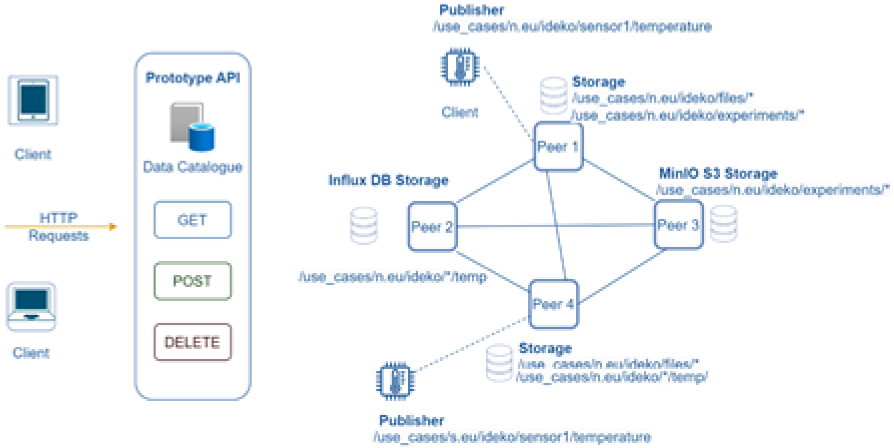
ExtremeXP details

Transparent decision making with interactive visualisation methods

- **SW components**
  - **Visualization** of Experiment workflow
  - **Visualization** of Experiment results
  - **Visualization** of Explanations
  - **Explainability methods** for pipeline variant, model and hyperparameter tuning
    - ALE, PDP plots
    - Counterfactuals, Influence Functions, Prototypes
- Gamification
- AR/VR Visualization
- Automatic specification of visualizations within workflows based on user intent

ExtremeXP Details

Extreme data access control and knowledge management



- Extend the Meta Data Schema (MDS) for capturing contextual information and creating meaningful **access control policies**
- Work on **context handlers** that will be developed as microservices for gathering contextual information necessary for evaluating incoming access requests.
- Integration of **decentralized data management tools** (like zenoh).
- Work on the access control mechanism using **smart contract and token standards**.
- Report in D5.2

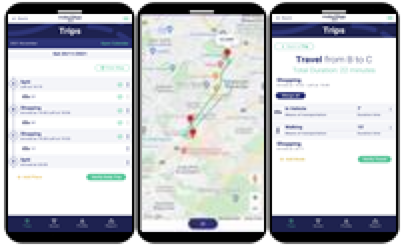
Funded by the European Union

ExtremeXP's Use Cases

Improvement of flash flood forecasting via AI

Increased Cybersecurity situation awareness
for efficient **threat mitigation**

Failure prevention
for manufacturing industry



Flexible **transportation analysis**
and visualization

Situational intelligence and decision making
for Public Protection and Disaster Relief

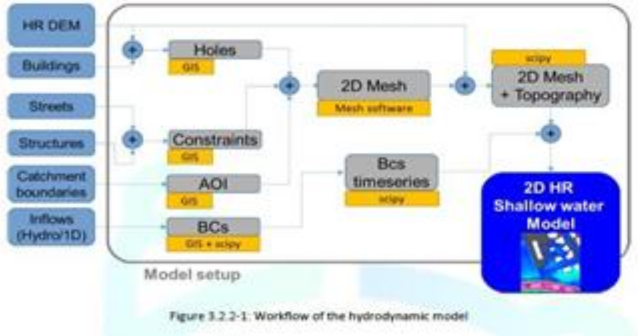
Funded by
the European Union

Urban flash flood prediction

Overall Goal: Improve accuracy of flood prediction

How ?
- #CAW1: Using AI to facilitate definition of input data.
  - #Exp1: Improve buildings' definition
  - #Exp2: Introduction of structures (streets, sidewalks, …)
  - #Exp3: Improve DEM resolution
- #CAW2: Using a surrogate model to predict flood.

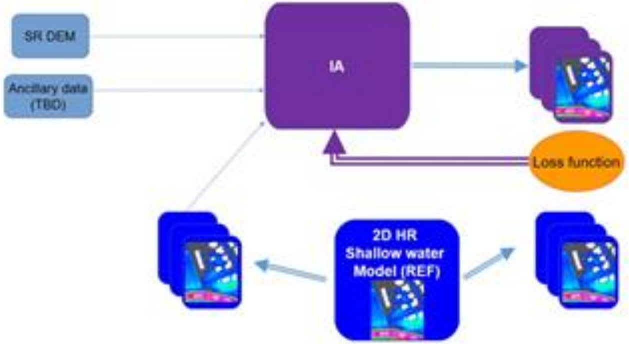
Challenges:
- Rapid events
- Urban areas with dense population and Infrastructures



CAW1: Hydrodynamic model workflow with inputs improved by AI

Flash flood events in Nîmes (3rd October 1988)

Area of Nîmes
- Particular location (near Meditarranean sea and backed by mountains blocking the warm water coming from the sea)
- Has experienced severe flash flood events (ex: 1988, 2002, 2005, 2014) caused by heavy rainfall



CAW2: Surrogate model

Increased Cybersecurity Situation Awareness for Efficient Threat Mitigation

Maxime Compastié

Senior Researcher

maxime.compastie@i2cat.net

Sophia Antipolis Sep. 19th 2025

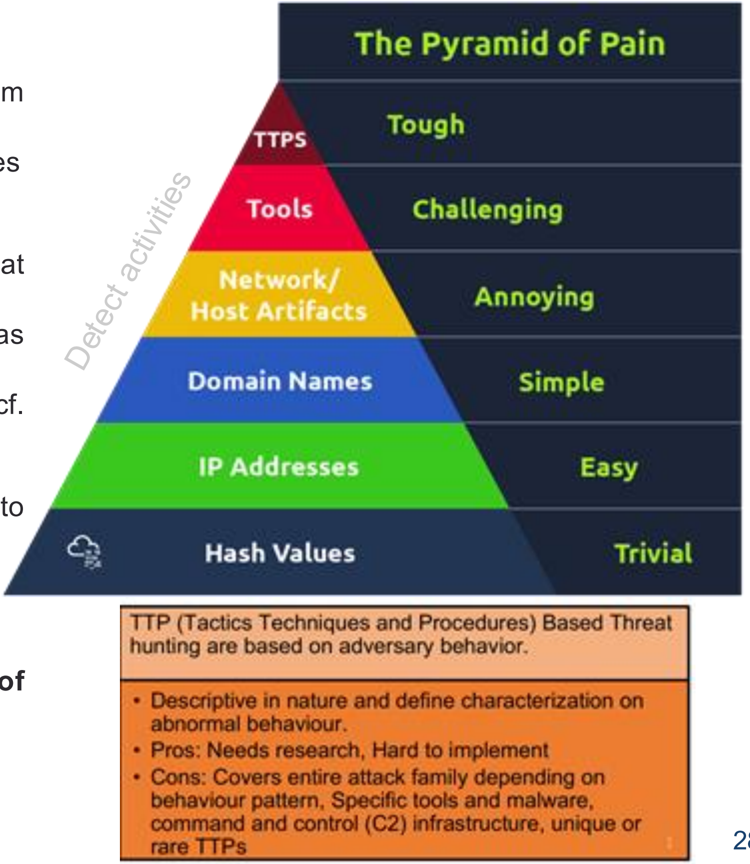
Motivation & Objective

- **Blossoming landscape** of cyber-threats, more complex and analyse from defenders' perspective (e.g. SOC operators)
  - Constraints: Multi-modality & high-volumetry from typical data sources
  - Extensive reliance on human expertise & insights => **Fatigue**

- Current efforts for Cybersec community to instrument AI-techniques for threat detection
  - Focused on practical indicators (hash-value) => **Lowly valuable** as easily changeable by attacker
  - **Our interest:** identifying and detecting attacker behavior ... **Hard!** (cf. Pyramid of pain)

- **Objective:** Exploit ExtremeXP experiment-driven analytics capacity to identify threat actors' behavior
  - Valuable as difficult to changes => Improved technique identification
  - Opportunities for anlysing attack scenarios => Detection of incentive

=> How to convert the process of threat behavior elicitation into a set of data analytics experiments?



SIEM - Security Information and Event Management

Funded by the European Union

The ExtremeXP Framework

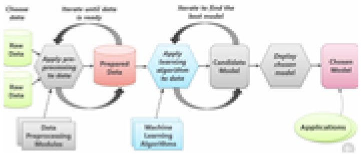
Extreme XP a Sota framework to learn rules by performing experiments that evaluates different analytics variants based on on an intent (in NLP), preference and constraint.

Intent

Preferences and Constraints

Raw Data

ExtremeXP



Proposed pipeline

ExtremeXP Components

We model variability using:

- **Intent:** The need of the end-user

  Train a model to detect the Midori threat

- **Preferences:** The desiderata of the previous intent

  An accuracy > .8 and reduce the number of FP

- **Constraints:** A limitation or a restriction

  The complexity needs to be bounded to log(n) and n^2

Funded by the European Union

Challenges and expected contributions

Approach: Leverage behavior-based Security Analysis

- Absence of sufficient dataset for specific attack techniques
  - ExtremeXP solution: Experiment generating a from supervised threat simulation dataset used for data augmentation

- Identification of adequate features relevant for threat detection
  - ExtremeXP solution: Training classification model as an evaluation, evaluation of their performance as an experiment

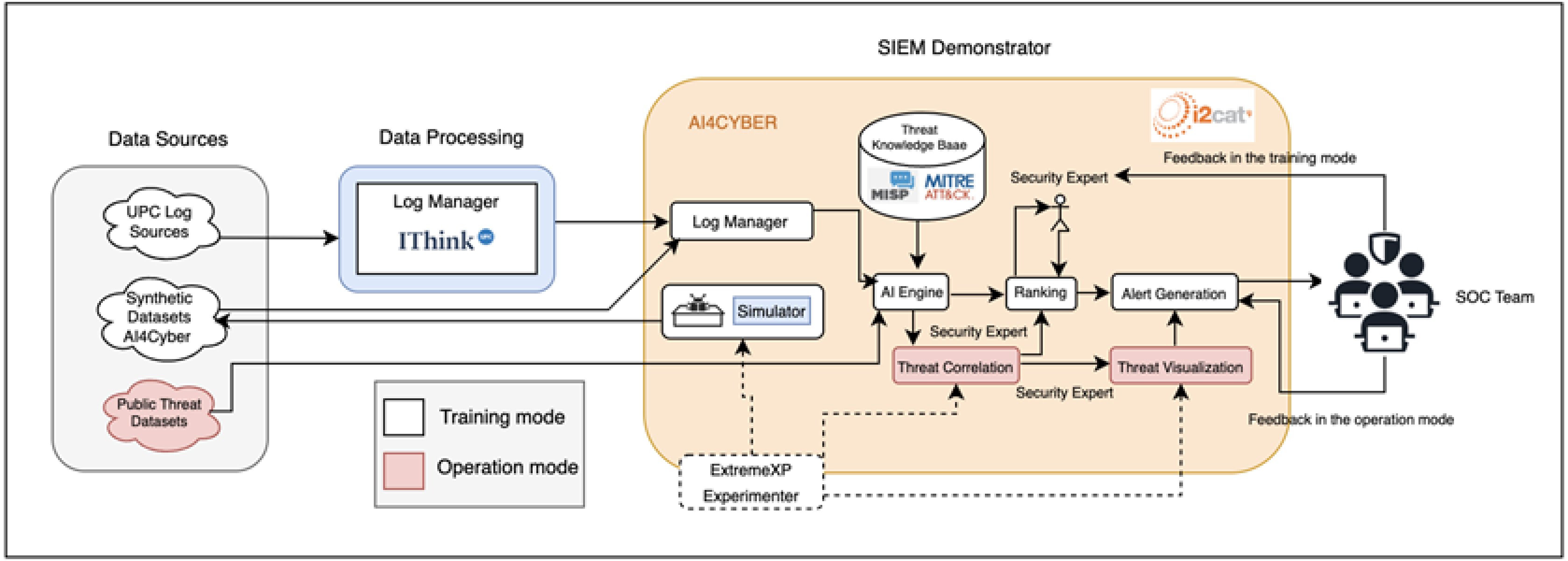
- Fit-for-purposeness of features to describe threat behaviors
  - ExtremeXP solution: Manned assessment of the quality of the explainations as experiment. Human-in-the-loop becomes Human-on-the-loop.

=> Implementation of different complex analytical workflows, run in an instance of ExtremeXP framework.

Funded by the European Union

*Adapted from "An Approach for Intelligent Behaviour-Based Threat Modelling with Explanations" Preetam et Al 2023

Use Case Architecture



UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH

i2cat

Funded by the European Union

Extreme Data from UPC

Network Firewalls
+5 FW security zones
~50GB/day

HTTP Load Balancers
~8GB/day

Servers
~20GB/day

+250 Applications
168,937,836 daily logs
~30GB/day

- Data Indexing
- Data Warehouse
- Data Presentation

- Security Information and Event Management (SIEM)
- Indexing and search
- Alerting & Monitoring
- Visualization & Reporting
- ML & Predictive Analytics

Funded by the European Union

The Extreme-XP Approach

Expected Results

KPI	Target
Detection of multimodal threat techniques referenced by the MITRE Att&ck framework	10
False positives and negatives on threat techniques classification	< 10 %
Mean time to classification compared to traditional human techniques	< 30 seconds

Funded by the European Union

ExtremeXP Project Factsheet

ExtremeXP (GA ID:101093164)

Call Topic: **HORIZON-CL4-2022-DATA-01-01**

Budget € 10.359.472 (EU contribution: **€ 10.011.820**)

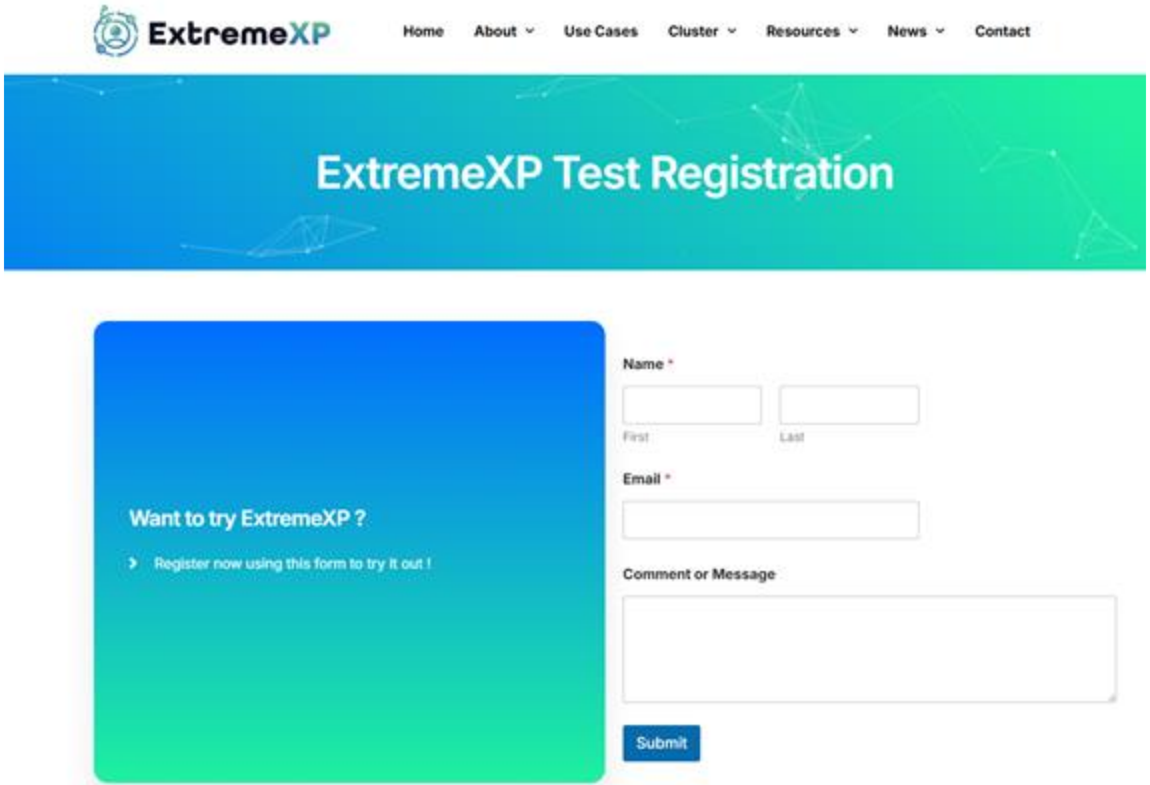
Started in **Jan 2023** – Duration: **36M**

20 partners: 12 academic, 5 SMEs, 3 industrial

1	**Athena Research Center (coordinator)**	RTD
2	**Activeeon**	SME
3	**Airbus Defense and Spaces SLC**	LG
4	**BitSparkles**	SME
5	**Bournemouth University**	U
6	**CS-Group**	LG
7	**Charles University of Prague**	U
8	**Deutsches Forschungszentrum für Künstliche Intelligenz**	RTD
9	**Fundacio Privada I2cat, Internet I Innovacio Digital A Catalunya**	RTD
10	**Institute of Communications and Computer Systems**	RTD
11	**IDEKO**	RTD
12	**INTERACTIVE4D**	SME
13	**INTRACOM TELECOM**	LG
14	**IThinkUPC**	SME
15	**MOBY X**	SME
16	**SINTEF**	RTD
17	**Technical University of Delft**	U
18	**University of Ljubljana**	U
19	**Universitat Politècnica De Catalunya**	U
20	**Vrije Universiteit Amsterdam**	U

Funded by the European Union

Try ExtremeXP : https://extremexp.eu/extremexp-test-registration/



THANK YOU