CHIST-ERA Projects Seminar 2022

Big data and process modelling for smart industry (BDSI)

Stefano Carrino (SOON)
Vicente Rodríguez Montequín (SOON)
March 30, 2022
Introduction: Projects of the Topic

Big data and process modelling for smart industry (BDSI)

- **FIREMAN** *Predictive maintenance in industrial processes empowered by IoT connectivity and Machine Learning*
- **PACMEL** *Process-aware Analytics Support based on Conceptual Models for Event Logs*: process mining, time series analysis, integration of heterogeneous data sources, integration of domain knowledge
- **RadioSense** *Passive sensing for robotic-assisted collaborative industrial spaces*: EM modelling, device-free radio sensing, human-machine interfaces, distributed and federated machine learning
- **SOON** *Social Network of Machines* - smart maintenance for Industry 4.0, multi-agent solutions

*Most of the projects of the Call 2017 are already finished.*
Major Achievements and Outputs - I

Publications
❖ Int. peer reviewed venues and journals: >120

Demonstrators
❖ Environmental radio-sensing in an industrial environment (localization, gesture, ranging)
❖ Failure & asset condition monitoring and detection prototypes (e.g., for grid-forming converters)
❖ Process simulators and optimization algorithms
❖ Event-driven sampling in industrial environments (e.g., microgrids) and 5G test networks (5GTN)
Major Achievements and Outputs - II

Open source software tools
❖ Novel ML/AI algorithms (centralized, federated, distributed)
❖ Open data sets, Open source code (github), and containers (docker)
❖ Knowledge Augmented Clustering (KnAC)

Areas of impact
❖ Industrial process monitoring, optimization, validation
❖ Smart/Predictive maintenance
❖ Environmental monitoring and functional safety
❖ (Big) data processing
❖ Beyond-5G industrial connectivity enablers
❖ Products quality improvement
Upcoming Challenges and Needs - I

Long-term vision

❖ Keep the human in the loop (HMI-HRI)
❖ Advances in next-generation networks & industrial connectivity
❖ New sustainable way of manufacturing through digitization
❖ Smart manufacturing (robotic assisted) and multi-agent systems

Research methods

❖ Address limitations of current AI methods (centralized vs. distributed)
❖ Interpretable/explainable modules for actionable insights and decision-making
❖ Convergence in cyber-physical system modeling approaches
❖ Experimental validation of theoretical achievements
❖ Stream mining with industrial data
❖ Data and model-driven analysis (statistical, physical/electromagnetic modelling)
Interdisciplinarity

❖ Exchange of knowledge among cross-disciplinary teams
❖ Interplay between industry, academia and business
❖ Expertise sharing at a transnational level

Results exploitation

❖ Collaboration in times of social distancing
❖ Pandemcy-related delays impact the schedule of exploitation
❖ Implementation in the wild
❖ Experiments at real-world industrial infrastructures (high TRL)
❖ Follow-up proposals (Horizon EU, EIC transition)
Possible Roadmap

How to achieve the expected impact
❖ Active participation in standardization efforts
❖ Interaction with industrial stakeholders
❖ Collaboration with industrial partners to implement the results

Where to make available the outputs after the project
❖ Github, Dataport IEEE, Green open access databases
❖ Project webpages and social media

Potential users of the results
❖ Maintenance teams, factory owners, network operators, service providers, equipment manufacturers, software developers

How potential users will be contacted
❖ Networks of project partners
❖ Dissemination channels
❖ Open workshops
Role of the CHIST-ERA Support

Reaching the main achievements of the project

❖ Expanding the scientific research at the European level
❖ CHIST-ERA personnel very helpful and fast
❖ Extensions in response to COVID-19 have been possible

Creating added value of implementing the project

❖ Follow up process provide value to our projects
❖ Video contest (participation in competitions of 2021 & 2022)
❖ Support for Future Tech Week Open research seminars (2021, 2020, 2019)

Satisfaction with the international and national implementation

❖ Yearly CHIST-ERA project meetings good to exchange and collaborate
❖ Nice that the project size is virtually not limited by single maximum budget
❖ Topic selection process: very valuable due to alignment with scientific community (academia, industry, funder organizations, …)

Possible improvements

❖ Coordination of national Agencies: misaligned funding or project start;
  ✔ Has been improved
❖ Some countries missing (e.g. Germany, Italy not every year)
❖ Industrial partners often not funded
Integration of RRI practices in the projects

Consortium agreements describe the treating of some aspects

- **Open Science**: open data, standardized evaluation platforms, scientific “challenges”
- **Science education**: Academic courses and theses (BSc/MSc/PhD) on the addressed topics. Organization of training events and special sessions in international conferences. Also, organization of cross-project* workshops.
- **Public engagement**: Raise public awareness via public webinars. Reach, stimulate and engage a critical mass of relevant stakeholders.
- **Ethics**: Responsibility towards environment, Discussions started with ethics experts on ethics in environmental sensing
- **Governance**: Discussion started with stakeholders in industry and TLC sectors

Major hurdles to RRI implementation in the projects

- Industry partners may be fearful of disseminating data and internal process description publicly
- Gender imbalance in ICT
- Limited access to public engagement and talks due to COVID-19 pandemic

*https://soon.umfst.ro/rationality.html
Open Science practices

❖ DMP: implemented by majority of projects
❖ OA publications: green open access at public, University maintained databases (e.g., aaltodoc.aalto.fi)
❖ Open data sharing: github, IEEE DataPort*, OpenAIRE, Zenodo***
❖ Data repositories used: github*, arxiv, Gitlab****, TechRxiv,

Obstacles to cope with good Open Science practices

❖ Industry reticence to share data
  ✔ Data sets from industrial partners can sometimes not be shared
❖ Rules for Open Access funds vary across national funding agencies
❖ Not all publishers allow gold/green open access
❖ Academic recognition of other sources of open access different to SCI journals

Costs of implementing the Open Science practices

❖ Gold Open Access fees can be high (e.g., Elsevier’s fee is 2,500 EUR)
❖ Supporting open data could be costly due to long term maintenance after project completion

* https://github.com/Superpalo/FIREMAN-project/blob/master/09_PowerConverter_dataset_preprocessing.ipynb
* https://github.com/labRadioVision, DOI: https://dx.doi.org/10.21227/0wmc-hq36
*** https://zenodo.org/record/4459969#.YeEv4IRByh4
**** https://version.aalto.fi/gitlab/salamid1/sidelinkchanneldataset
Challenges linked to exploitation, and IPR

- Difficulty transferring to other cases/scenarios
- (Often) Industry is not ready for our solutions! (No data available, no digitalization, etc.)
- Software can’t be patented
  - ✔ Problems in case of disseminating details before patenting

Steps taken towards technology transfer

- Spin-offs:
  - ✔ Towards commercialisation of machine learning-based research via a thorough customer discovery and value proposition.
- Follow up projects at higher TRL
  - ✔ Already started & Novel proposals in the next months
  - ✔ Involvement of industrial partners (SMEs) and policy makers

Tension felt between technology transfer and Open Science

- Industrial, sensitive data
Big data and process modelling for smart industry (BDSI)

Thank you very much

Questions?