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** (*Environmental sensing for Human-Robot collaborative spaces from Electromagnetic signals*)

**SOON** (*Social Network of Machines*) - smart maintenance for Industry 4.0

**BIG-SMART-LOG**: The Use of Big Data Analytics for Process Modelling in Smart Logistics Operations

**SPuMoNI**: Smart Pharmaceutical Manufacturing

Unifying elements- Keywords:

- Anomaly detection
- Process optimization
- Process mining
- Big Data processing
- Process monitoring
- Industrial process related
- Industrial connectivity
Major Achievements and Outputs

Publications:
  Int. peer reviewed venues and journals: >100

Demonstrators:
  - Environmental Radio-sensing (localization, gesture, ranging)
  - Failure monitoring and detection prototypes
  - Smart Logistics (Long-term multi-factor forecasting)
  - ALCOA Tool for Pharma quality

Open source software tools:
  - ML/AI Algorithms
  - Open data sets, Open source code (github), and containers (docker)
  - Blockchain (ethereum)

Areas of impact
  - Industrial process monitoring, optimization, validation
  - Smart/Predictive maintenance
  - Environmental protection & safety
  - Pharma industry
  - (Big) data processing
  - Beyond-5G industrial connectivity enablers
  - Products quality improvement
  - Environment monitoring and assessment
Upcoming Challenges and Needs

Long-term vision:
- Keep the human in the loop (HMI-HRI)
- Advances in next generation networks
- Smart Logistics
- New sustainable way of manufacturing through digitization
- Smart manufacturing

Research methods:
- Limitations of current AI methods
- Convergence in cyber-physical system modeling approaches
- Experimental validation of theoretical achievements
- Long-term multi-factor forecasting

Interdisciplinarity:
- Exchange of knowledge among cross-disciplinary teams
- Interplay between industry and academia
- Expertise sharing at a transnational level

Implementation:
- Implementation in the wild
- Experiments at real-world industrial infrastructures

Results exploitation:
- Collaboration in times of social distancing
Possible Roadmap

How to tackle the challenges within the topic:
- Using new technologies from A.I., blockchain, data analytics, deep learning, BigData and beyond 5G communication

How to achieve the expected impact:
- Active participation in standardization efforts
- Interaction with industrial stakeholders
- Industry collaboration partner 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
- Improvement the efficiency of industrial processes
- Fast responses by Chist Era personnel on requests
- Provided valuable time extensions in response to COVID-19 challenges
- Expanding the scientific research at the European level
- Involvement of academia in topic definition

Creating added value of implementing the project
- Video contest
- Future Tech Week Open research seminars

Satisfaction with the international and national implementation
- It is nice that the project size is virtually not limited by max budget since each partner is funded by national agency
- Yearly Chist Era project meetings are nice for exchange and informing of other’s process
- The two topical calls have been scientifically interesting (at least one of them) for a number of years in a row now.

Possible improvements
- National Agencies coordination for misalignment of national funding/start of the projects, or even possible lack of communication with CHIST-ERA
- Some countries missing (e.g. Germany, Italy not every year)
- Industrial partners are not funded most of the times
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.

**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

- Data processing and interpretation of the process under study may be non-trivial, especially when the RI targets are towards increased TRL levels.
- 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.
Open Science practices:

- **DMP:** implemented by majority of projects
- **OA publications:** green open access at public, University maintained database
- **Open data sharing:** github, IEEE DataPort, OpenAIRE
- **Data repositories used:** github, arxiv, Zenodo, TechRxiv

Obstacles to cope with good Open Science practices:

- Industry reticence to share data
- Rules for Open Access funds vary across national funding agencies
- No all publishers allow gold/green open access
- Data sets from industrial partners can sometimes not be shared

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 expensive technologies and challenging after the completion of the projects
Challenges linked to exploitation, and IPR:
- Software can’t be patented
- Problems in case of disseminating details before patenting

Steps taken towards commercialization:

**IP Registration:**
- invention disclosures formulated
- invention disclosure accepted for filing

**Spin-offs:**
- Towards commercialisation of machine learning-based research via a thorough customer discovery and value proposition.

Tension felt between technology transfer and Open Science
- Industrial, sensitive data
Questions