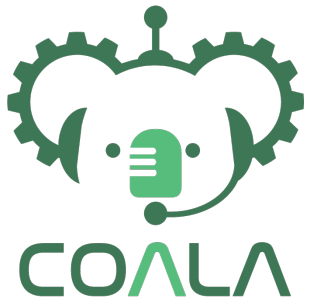


COgnitive Assisted
agile manufacturing for a LAbor force
supported by trustworthy
Artificial Intelligence

Legal Issues and EU AI Act Implications on COALA Project

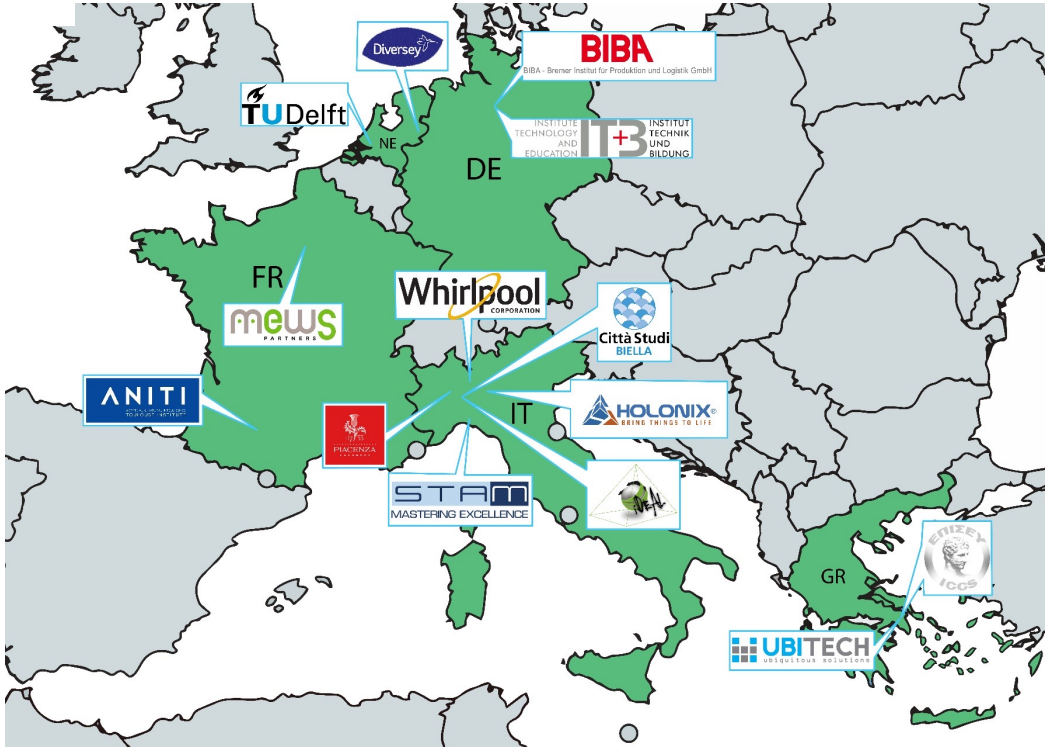
H2020 COALA Project

Céline Castets-Renard, Law Professor, ANITI

Horizon 2020
European Union Funding
for Research & Innovation

COgnitive Assisted
agile manufacturing for a LAbor force
supported by trustworthy
Artificial Intelligence

- Objective: AI for manufacturing
- Topic: ICT-38-2020
- Call: H2020-ICT-2018-20
- Lead: BIBA – Bremer Institut für Produktion und Logistik GmbH
- Duration: 36 Months
- Start: 2020/10



VISION: develop a human-centered AI-based digital assistant to optimize manufacturing quality and training

- Provide a more proactive approach to support situations characterized by cognitive load, time pressure, and little or zero tolerance for quality issues.
- Support **workers that need to use analytics tools** and **new workers that perform on-the-job training.**

Vision

The COALA vision for AI in manufacturing is the **development of human-centered digital assistant** that provides a more proactive and pragmatic approach to support operative situations characterized by cognitive load, time pressure, and little or zero tolerance for quality issues.

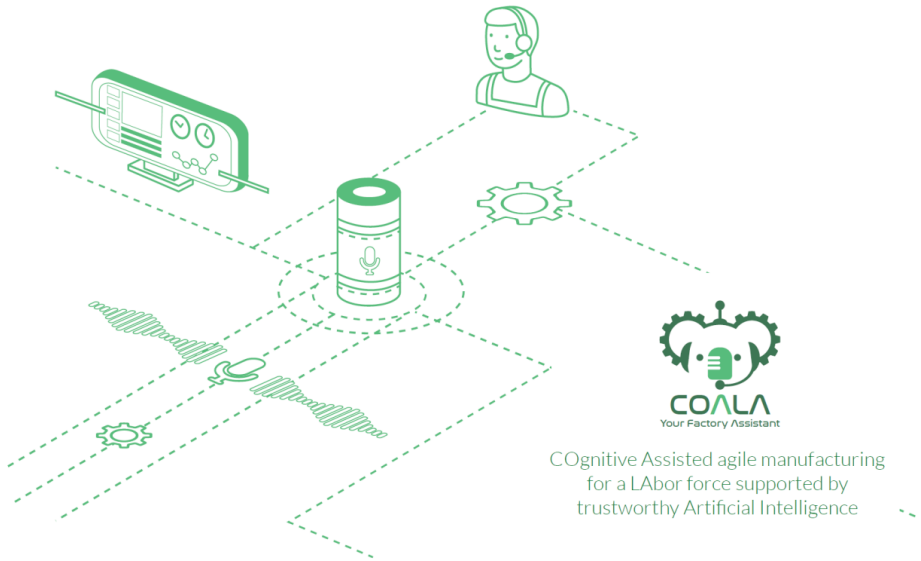
COALA will help shaping the complementarity in the collaboration between the AI-based assistant and the human so that:

- the AI will take over time consuming and stressful tasks reliably and credibly
- the human will focus on understanding and problem-solving in complex & knowledge-intensive situations.

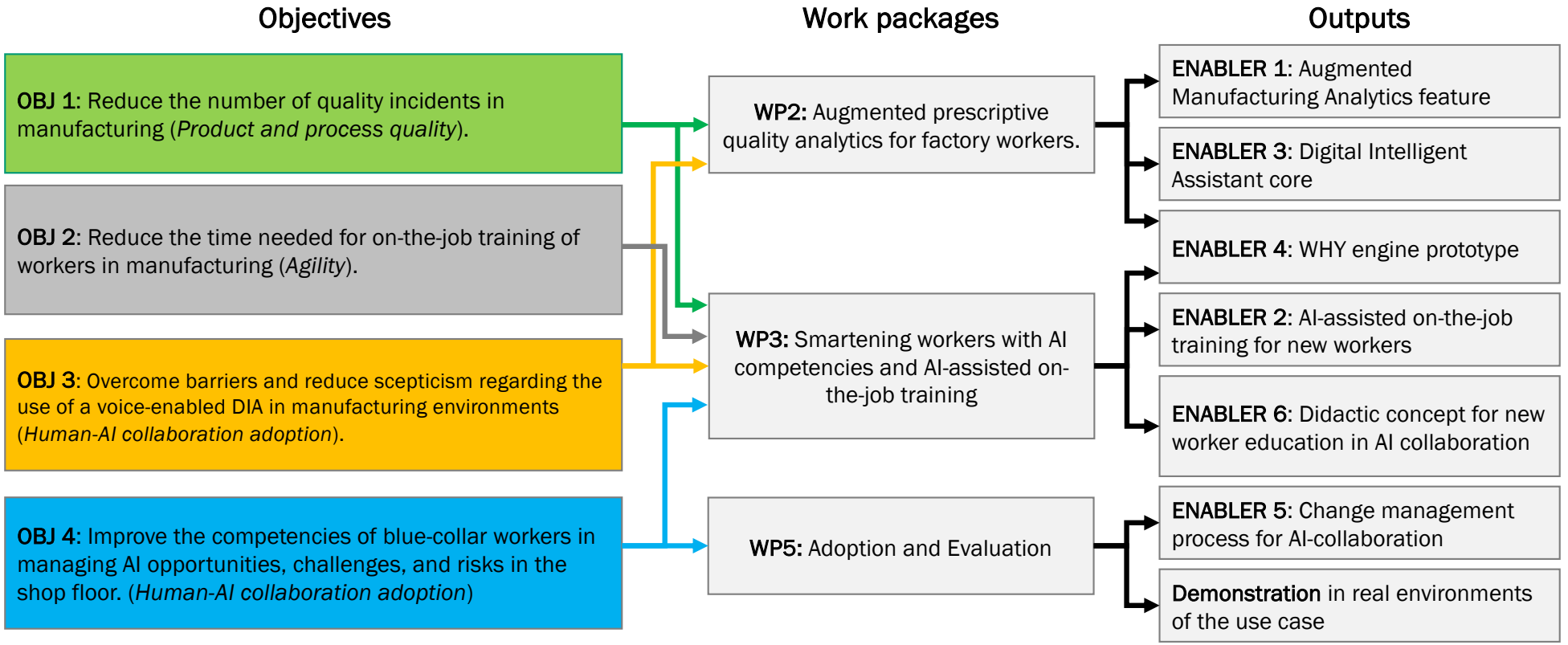
COALA's AI-focused education and training concept will prepare the human-side of the collaboration by offering concept for teaching professionals systematically and, in the language of the workers, about the capabilities, risks, and limitations of AI in manufacturing.

Solution

- COALA will provide a solution for **cognitive assistance** that consists of a composition of **trustworthy AI components** with **a voice-enabled digital intelligent assistant** as an interface.
- The solution will support **workers that need to use analytics tools** and **new workers that perform on-the-job training.**
- Complementary to the technology, **an education and training concept** that focuses **on building blue-collar worker competencies in human-AI collaboration** will be developed.
- The COALA solution will transform how workers perform their jobs and it allows companies to maintain or increase the quality of their production processes and their products.



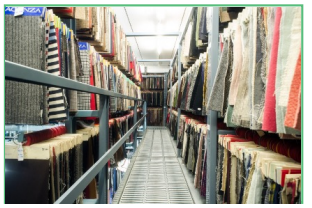
Objectives & Enablers

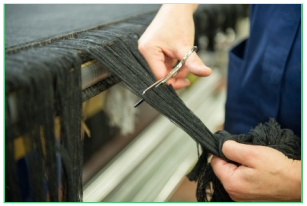


Business Cases Overview

Textile Production

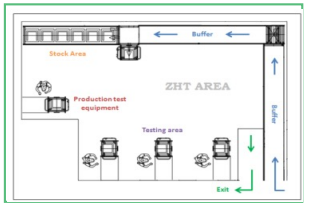
- reduce the time and cost of machine operator trainings
- maintain the worker's autonomy vs. unquestioned execution of instructions
- exploit and spread the concept and solution





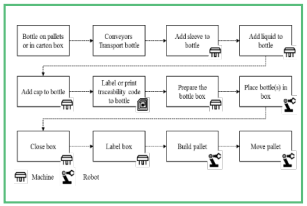
White Goods Production

- adopt a predictive quality strategy
- link the quality control of the finished product with design stage and shop floor
- reconfigure its production line and facilitate root cause analysis



Detergent Production

- increase efficiency of production and reduce change over time
- capturing the best practices and transferring them between manufacturing lines
- tacit line-manager knowledge capturing, formalizing, and transfer to co-workers



Ethical & Legal Issues: Human-AI collaboration

- Gender Bias
- Trust
- Balance of Power
- Privacy (excluded in this presentation – GDPR applies / DPO)

Gender bias: Human-AI collaboration

Most of the workers are men *(only men are working in 5-10L in production line*)

However: No impact

- User profiles should not be classified in terms of gender aspects
- Recommendations should not be adapted to the tone and other aspects of the voice settings

Trust: Human-AI collaboration

- Trustworthy interactions are essential and relevant as soon as a human operator/worker is involved.
- **All scenarios are therefore concerned**, and it will be a critical indicator in evaluating all scenarios.

- **Attention points**
- The **feeling of continuous recording** of voice or video could affect trust in COALA
- A **lack of fluidity** in the conversation could affect trust in COALA (e.g. not understandable, vague or latency)
- A **lack of effectiveness** in the conversation could affect trust in COALA (e.g. inadequate recommendations or poor-quality risk assessment)
- A **lack of emotion and authenticity** could affect trust in COALA (e.g. NLU performance, perceived bot friendliness, confirmation request, paraphrasing responses)
- A **lack of mutual trust** could affect trust in COALA (e.g. too many confirmation requests giving users the impression that they are not trusted)

Balance of power and others : Human-AI collaboration

Balance of power

- **All scenarios are concerned** but no major issues have been raised regarding balance of power.

- Only the following points have been listed :
- Recommendation management options could be available to **set the level of autonomy and freedom** of the operator/worker in following or not the recommendations
- Recommendation management options could be available to provide the operator with the **ability to define a time period during which he/she doesn't want to be interrupted** by the COALA assistant

- **Others: could be an accuracy issue**
- **Age and social background** may affect how users talk to the assistant, leading to biases similar to those of gender. In addition, a change in these workforce characteristics might affect the Dialog Management training data, leading to a **lower accuracy of the dialog model**.

European Commission Proposal on AI (AI Act) - April 21, 2021
Legal Issues: Risks-Based Approach

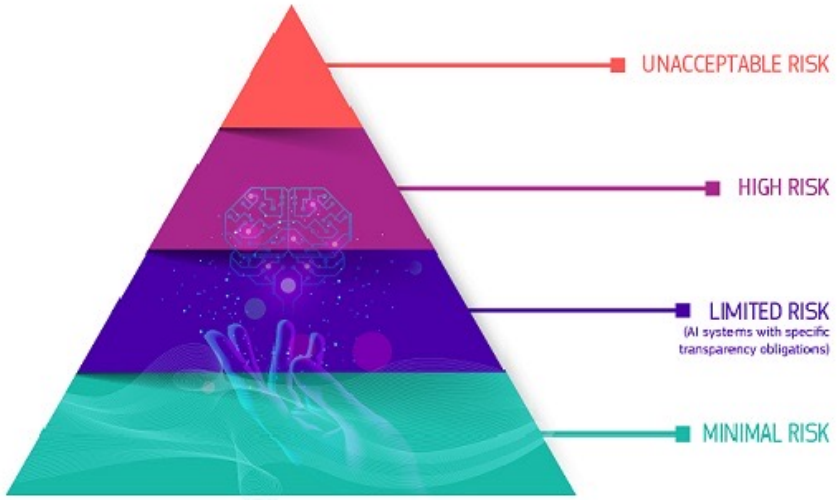
- **Material**
  - AI Systems which generate outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with (art. 3(1))

Annex 1: referred to in Article 3(1)

(a) Machine learning approaches, including supervised, unsupervised and reinforcement learning, using a wide variety of methods including deep learning;

(b) Logic- and knowledge-based approaches, including knowledge representation, inductive (logic) programming, knowledge bases, inference and deductive engines, (symbolic) reasoning and expert systems;

(c) Statistical approaches, Bayesian estimation, search and optimization methods.



Unacceptable Risks – AIS Prohibited (Article 5)

COALA
Your Factory Assistant

X **Subliminal manipulation** resulting in physical/ psychological harm

X **Exploitation vulnerable groups such as children or mentally disabled persons** resulting in physical/psychological harm

X **General purpose social scoring by public authorities**

X **"real Time" Remote biometric identification** in publicly accessible spaces for law enforcement purposes (with 3 exceptions)

High-risk AI Systems (Title III) - Specific Use Cases (Annexes III)

- **Certain AI systems in the following fields:**
  - Biometric identification and categorisation of natural persons
  - Management and operation of critical infrastructure
  - Education and vocational training (purpose of determining access to educational training institutions or assessing students)
  - Employment and workers management, access to self-employment (for recruitment, selection or promotion)
  - Access to and enjoyment of essential private services and public services and benefits
  - Law enforcement
  - Migration, asylum and border control management
  - Administration of justice and democratic processes

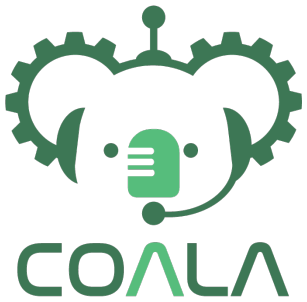
Limited Risks

Article 52 - Transparency obligations for certain AI system

Information : Providers shall ensure that **AI systems intended to interact** with natural persons are designed and developed in such a way that natural persons are informed that they are interacting with an AI system, **unless this is obvious from the circumstances and the context of use.**

Conclusion

- Only one simple legal requirement
- But ethical issues which deserve attention
  - Human-Ai interaction: Need for Human Control
  - Otherwhise: Risk of System Unaccuracy in the iterative process = Need for Data Management System



COgnitive Assisted
agile manufacturing for a LAbor force
supported by trustworthy
Artificial Intelligence

Thank you for your attention!

Horizon 2020
European Union Funding
for Research & Innovation