# AI REGIO TEchnical and REgulatory SAndboxes (TERESA)

Marina Cugurra, R&I Lawyer and Ethics Expert (ETA)

EBDVF 2022, 22 November 2022 (Prague)

AI REGIO

eta consulting

# The AI REGIO Innovation Action

**AI REGIO** - Regions and DIHs alliance for AI-driven digital transformation of European Manufacturing SMEs

**GRANT AGREEMENT**: 952003 (Innovation Action)

**START DATE**: 1 OCTOBER 2020

**DURATION**: 36 MONTHS

**TOTAL FUNDING**: 8 Millions EUR

**CONSORTIUM**: 36 FULL BENEFICIARIES

**COORDINATOR**: Politecnico di Milano (Sergio Gusmeroli)

AI REGIO

# Human centricity at the core of AI REGIO research

- Three core elements of **Industry 5.0**: human-centricity, sustainability, and resilience

- **Pivotal nature** of work in most adult lives.

- In the **Workplaces of the future**, humans and machines will share physical spaces working not only sequentially but even with close, physical real-time responses from machines/robots to users

- Putting the cutting-edge CI-driven **technologic advances** **at the service** of **human needs and interests** in view of adapting the production process to the needs of the worker

- Several **legal and ethical aspects** to investigate

AI REGIO

# AI REGIO contribution towards a human-centric, legal compliant and ethically-sound human-machine co-working environment

**Survey** on the main ethical, legal, regulatory, psychological and societal impacts of Industry 5.0 and CI solutions, focusing on related challenges and expectations

**AI REGIO EL toolkit** (Ethical Strategy, EDPIA, Ethics, Fairness, Privacy and Security by Design Approach, EL requirements, FRIA, Consultation with stakeholders…)
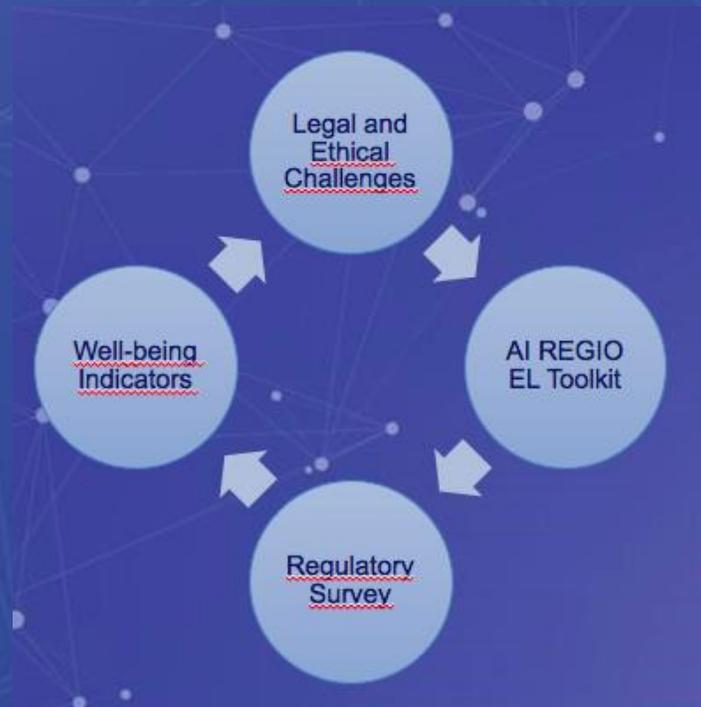
**Well-being indicators**

**TERESA** Experimentations, some AI REGIO experiments



AI REGIO

# TERESA's role in AI REGIO

Addressing the **regulatory and ethical issues** derived from the full adoption of **human-AI interaction** in Production Systems, via the development of **technology-regulatory sandboxes** in the different regions through the **DFs' facilities**



**Human-centric and trustworthy human-machine CI-empowered co-working environment**

# What is a TERESA?

"**TEchnical and REgulatory SAndbox**" for AI, enabling a <u>testing environment</u> for innovative CI-empowered products and services with human-machine interaction, where the <u>ethical challenges</u> and shortcomings of the <u>regulatory framework</u> concerning such products and services can be addressed

**Experiments running on a limited scale and in a secure and controlled way**, according to the "test before invest" paradigm (**Technical Sandbox**), with the involvement of **volunteers** (representing the Civil Society dimension within the sandbox) to test such solutions in **real regulatory conditions** in a gradual and **controlled way** before going to the Industrial plants, pursuant to a specific testing plan agreed and monitored by the competent authority (**Regulatory Sandboxes**)

Implementation in DFs some TERESA for experimenting ethics-related human-machine interaction models

AI REGIO

# TERESA in AI REGIO: key features

| |
|---|
| "hands-on" bottom-up **approach** tailored to the specific needs of manufacturing |
| Innovative **AI applications/tools/services** for human-machine interaction according to the CI methodology |
| Relaxed concept of Regulatory Sandboxes with the adequate level of <u>flexibility</u> |
| Enhancement of the regulatory response to **innovation** in the manufacturing domain |
| Engagement and knowledge-exchange between the **CA** and the **DFs** |
| Safeguards for **voluntary participants** |
| **Precursor initiative** aligned with the AI Act and linked to the AI TEF for Manufacturing |

# The DF's role

The **Didactic Factories** in the AI REGIO ecosystem **conceive, design and implement AI REGIO TERESAs**, if opportune with the help of <u>supporting technological partners</u>, and through the involvement of both the <u>Competent Authority/ies</u> and <u>participants</u> who will voluntarily take part to

these testing activities

DFs as **ideal place within AI REGIO to implement the TERESAs**
-Already **functional experimental facilities with ties** to industry problems, SMEs and manufacturing employees (closeness to **real factories)** where **constraints can be relaxed**, while guaranteeing the **safety, dignity and wellbeing** of the participants involved in the experiment
-Crucial role in addressing **ethical and regulatory issues** arising from adoption of AI and human-machine interaction

AI REGIO

# The CA's role

## Enhanced Knowledge

- Better **understanding** of human-machine collaboration and related new challenges and risks

- **Monitoring** the actual development and testing

- **Dialogue and experience sharing** with DFs and, eventually, with volunteers (key barriers faced during experiment's running)

## Contribution to regulatory/ethical development

- Timely **update of regulatory policies** addressing barriers to innovation for **bridging the gap** between innovation and regulation
- Inform **policy making and development** (identification of areas for improvement)
- **Recommendations and Proposals** of regulatory reform
- Test a **certain customised regulatory/ethical inspired approach**
- **Operationalization** of the ethical principles/codes of conduct

## Reducing regulatory and ethical uncertainty

- Early-stage **liaison** and proactively **engagement** with DFs (and volunteers): frank and fast **feedback, guidance and indications** on regulatory and ethical requirements and risks

- **Trust building** and strong **signal** to innovators about the propensity of the CA to support innovation

AI REGIO

# AI REGIO TERESA proposed model

Innovation ecosystem based on the synergy between DFs (drivers, innovation facilitator), Innovators (AI REGIO Consortium) and Standardisation Bodies

- **One SB for each (or some of) AI REGIO TERESA**
  - Network of 34 National Standardisation Bodies
  - UNI

- **Introduce TERESA DFs/SB Model in European CT** (CEN-CENELEC Focus Group on AI, STAIR, …)
  - Possible replicability of AI REGIO model
  - Alignement with the new regulatory reforms, including the AI Act (role of DIH, regulatory sandboxes, support to SME)
  - operationalization of the requirements of the ethics guidelines for trustworthy AI (EC's priority)

# WISE TERESA: EL issues within AI REGIO



**W**ell-being, Comfort and Acceptance

**I**nclusion and special categories of workers

**S**afety of the worker

**E**rgonomics and improving working conditions

AI REGIO **well-being indicators** for the workplace

- **Well-being, Comfort and Acceptance**. Impact on mental well-being and self-esteem, frustration, feeling of usefulness, emotional dependence and overconfidence on the machine, human dignity, autonomy and oversight, concerns/willingness in collaborating with a machine

- **Inclusion and special categories of workers**: effects on older workers, effects on novices, effects on workers with cognitive or physical disabilities/impairment, social isolation, risk of discrimination/bias

- **Safety of the worker**: Health and Safety of the workers, risks of harm, personal data

- **Ergomics and improving working conditions**: ergonomics, stress reduction, fatigue reduction, effects on workers' skills

EL Category **Mapping** for AI REGIO TERESA

# Human well-being indicators for AI-empowered workplaces and human-machine collaboration in the Manufacturing domain

Metrics covering **multiple dimensions of wellbeing,** capturing the factors which allow the **comprehensive assessment** of the benefits and possible challenges of CI artefacts to test and monitor that such artefacts contribute to the **operator's flourishing** and do not bring **unintended negative consequences** that could diminish human comfort, as well as that **new routes to a human-centric AI in the workplace** are identified.

Identify **risks** for the workers and challenges from an ethical and societal viewpoint, and take appropriate **mitigating actions** when necessary

Workers's **positive attitude** in engaging with cutting-edge CI-driven technologies in their workplace (**Satisfaction & Trust building**)

Harnessing AI potential, its **acceptance and adoption**

AI REGIO

# AI REGIO well-being indicators

Safety and Physical Health

Community

Self-esteem and Sense of worthiness

Self-actualization

Psychological Stress

Time-saving

Fatigue reduction

Social inclusion

Productivity/Performance

Sense of autonomy

Skills

Job satisfaction

Psychological Well -being



Common format (to be adapted to the specific context) for assessing WISE implications in AI REGIO TERESA
Insights and lessons learnt useful for the CA and future regulatory advancements

AI REGIO

# Examples of AI REGIO TERESA

### «Mini Factory» TERESA

- Switzerland
- Human-robot collaboration in the SUPSI Mini Factory through different small experiments dedicated to Collaborative Robotics and Human-centred Production Systems, with different scenarios where a cobot and humans work together in various tasks (assembly, screwdriving) and with varying degrees of collaboration (separated and independent, sequential, synchronous, etc.)

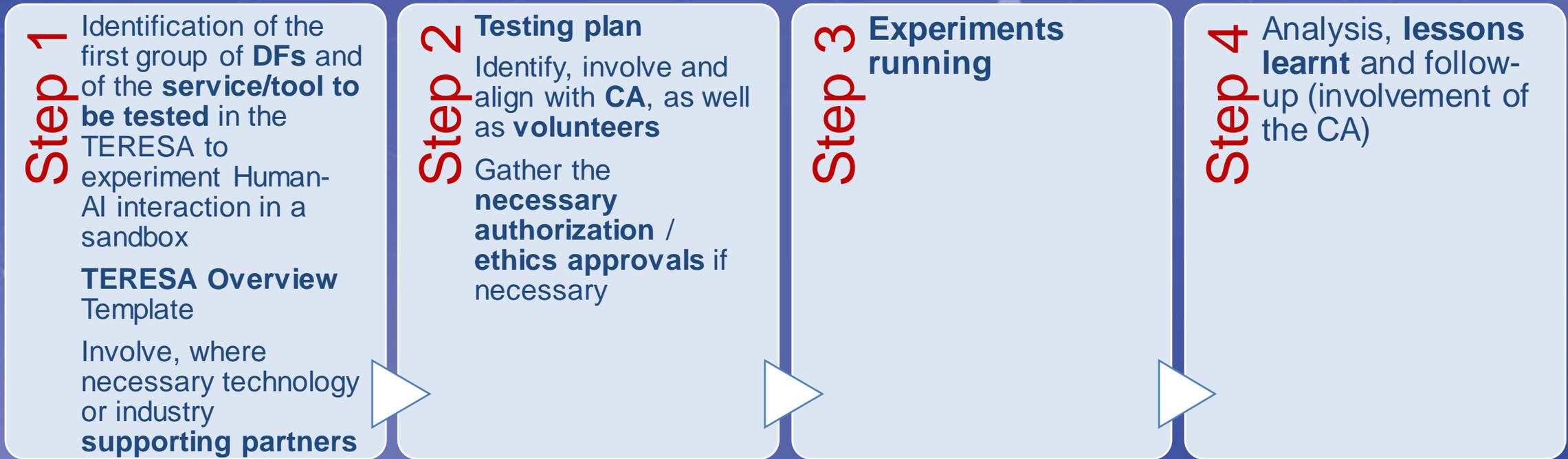### «BIC – Factory of the Future Experience Center» TERESA

- The Netherlands
- Fast, flexible and faultless assembly of different products, with multiple experiments such as operator support system in a manual assembly workplace and handling machine data, production processes and information exchange along the chain
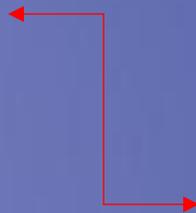
### «MADE Competence Center» TERESA

- Italy
- Human-centered Capsule Picking Vision System including a vision station with 3D inspection for recognizing stacks of capsules located inside a box and identifying their optimum gripping point (so that they can be easily picked up by the robot)

AI REGIO

# TERESA overall timeline

**Step 1**
Identification of the first group of **DFs** and of the **service/tool to be tested** in the TERESA to experiment Human-AI interaction in a sandbox

**TERESA Overview** Template

Involve, where necessary technology or industry **supporting partners**

**Step 2**
**Testing plan**
Identify, involve and align with **CA**, as well as **volunteers**

Gather the **necessary authorization / ethics approvals** if necessary

**Step 3**
**Experiments running**

**Step 4**
Analysis, **lessons learnt** and follow-up (involvement of the CA)

Enlargement of the DFs' core group hosting a TERESA

AI REGIO

# Workshop with the AI REGIO Regional Champion DFs

**Usefulness of the TERESA** as a practical tool for testing cutting-edge AI products and services for human-machine interaction,

**Challenges/barriers** in conducting a TERESA (cost constraints, difficulty to engage SMEs, difficulty to involve volunteers,…)

Interest in **enriching some running experiments** with legal and ethical considerations

**Further EL interesting issues**: data privacy, cooperation and relations within company teams, proper inclusion of workers during the development and implementation phases of an innovation.
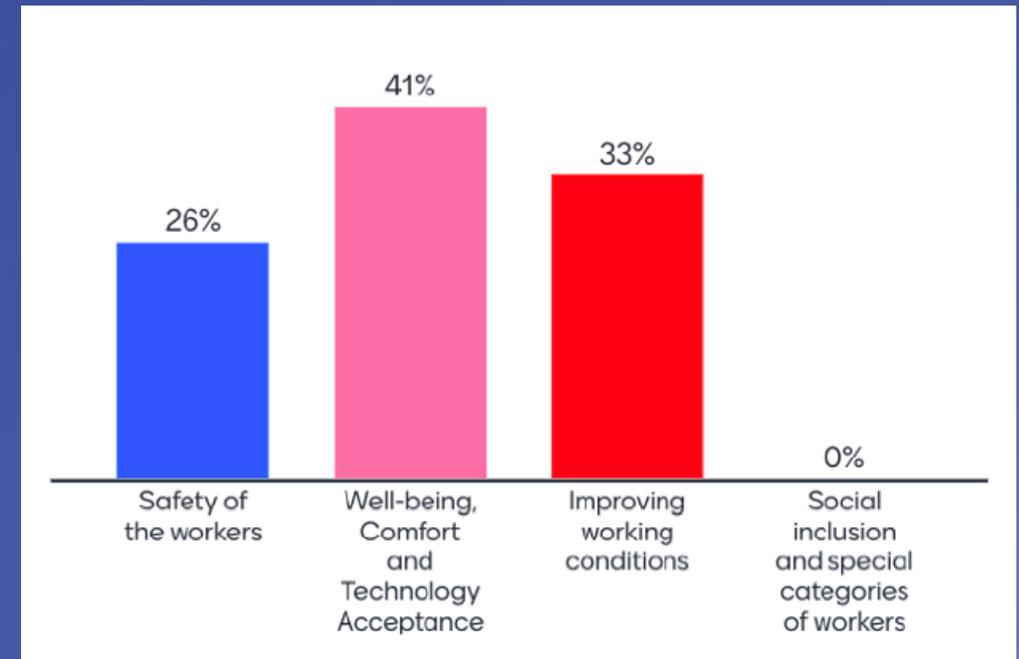
**Type of Competent Authority**: not only Standardisation Bodies, but also other options (Regional Authorities or Local Bodies,…)

**17 February** 2022
**Objective**: investigate expectations regarding TERESA, current ideas on possible experimentations and challenges and deepening next steps
**Participatory session** conducted through Mentimeter

Categories of the EL issues



41% — Well-being, Comfort and Technology Acceptance
33% — Improving working conditions
26% — Safety of the workers
0% — Social inclusion and special categories of workers

AI REGIO

# Thank you very much

**Marina Cugurra** (R&I Lawyer and Ethics Expert, ETA)
**marina.cugurra@eta-one.com**