SAFETY OF ARTIFICIAL INTELLIGENCE EMBEDDED IN MACHINERY

ALLOWING SAFER MACHINES ON THE EUROPEAN MARKET
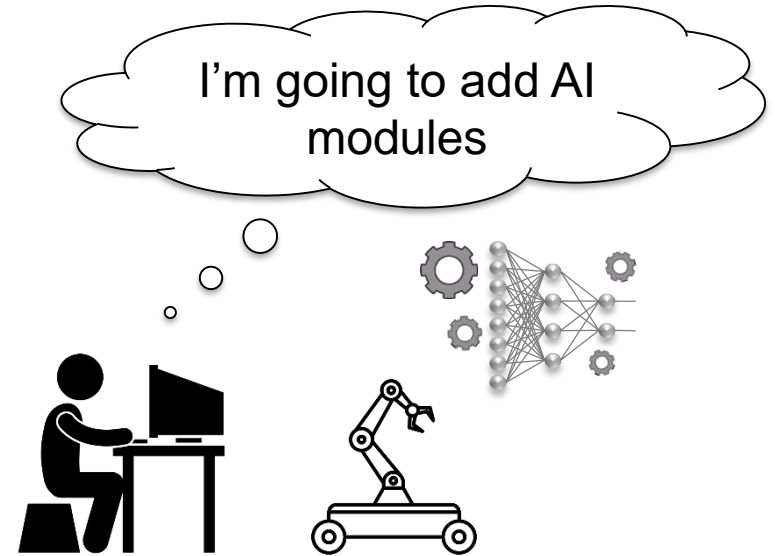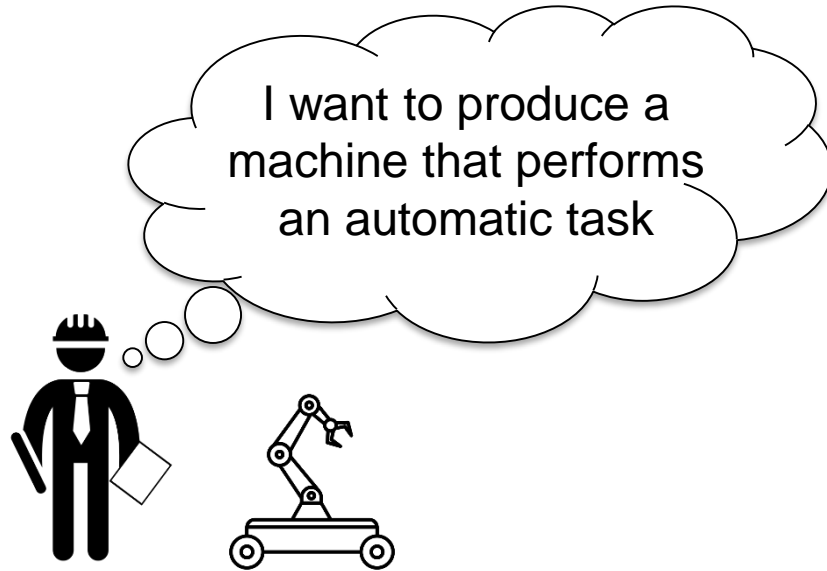
DR. AGNES DELABORDE - agnes.delaborde@lne.fr

LNE – FRENCH NATIONAL LABORATORY FOR METROLOGY AND TESTING

LABORATOIRE
NATIONAL
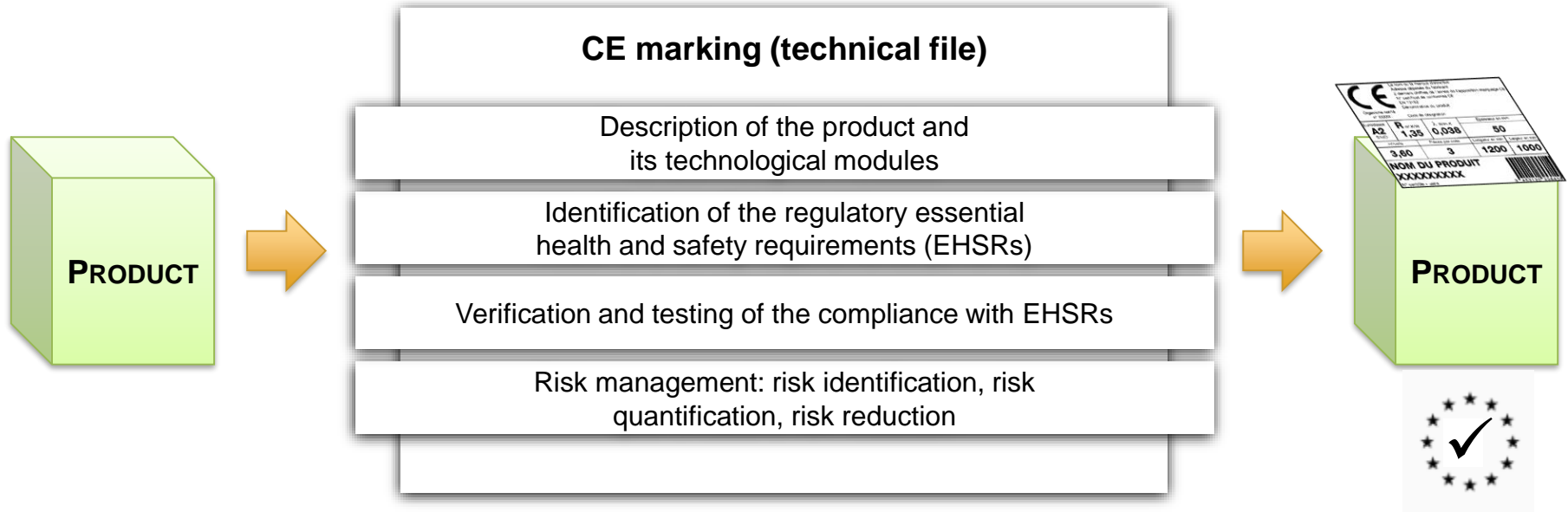DE MÉTROLOGIE
ET D'ESSAIS

LNE

From the manufacturer's perspective

AI AND AUTONOMOUS MACHINES – THE CONTEXT

Regulatory requirements

- All machines (thus including robots) must receive CE marking before being put on the European market

PRODUCT

CE marking (technical file)

Description of the product and its technological modules

Identification of the regulatory essential health and safety requirements (EHSRs)

Verification and testing of the compliance with EHSRs

Risk management: risk identification, risk quantification, risk reduction

PRODUCT

Agnes Delaborde - Safety of Artificial Intelligence embedded in machinery

AI AND AUTONOMOUS MACHINES – THE CONTEXT

Main issues

Both small start-ups and large groups wonder whether their autonomous products are compliant.
What verification methods?

Public authorities must have some control over these products (adapted regulatory requirements), without hampering innovation and competitiveness.

The mediatic coverage of AI relays many scientific imprecisions (exaggeration of the performance, overdramatization of risks), which affects all the stakeholders.

Agnes Delaborde - Safety of Artificial Intelligence embedded in machinery

FOCUS: AI IN RISK ASSESSMENT OF A MACHINE

Many research issues

Risk assessment: principles

- Identification of the risks:
  - Source
  - Nature
  - Criticity
- Identification of the causal chain
(interactions between the modules)
- Identification of the failure modes
(causes of failure)
- Quantification of the risks
(based on the probability of occurrence of the danger)

Reference texts mainly deal with software functional performance
(the system crashes or not)

Can we predict and model the failure modes?

Can we predict **all** the "wrong" decisions of a complex system?

Are standard risk quantification methods adapted to AI?

Will a given ''wrong'' decision really cause a damage?

No reference "acceptable" error rates

No approved methods for quantifying error rates

Not enough hindsight on the users' "natural" behaviours
in front of AI-driven machines

Agnes Delaborde - Safety of Artificial Intelligence embedded in machinery

ARTIFICIAL INTELLIGENCE AND MACHINERY

Leverage research advances to empower industry with tools for safety compliance

Provide a
reliable performance assessment
of autonomous machines

Equip the industry with
**tools for safety
verification and validation**

→ Identify and understand the causes of underperformance
→ Provide benchmarking methods for the quantification of errors

→ Accompany safe developments through testing
→ Develop methods and testbeds for safety V&V
→ Recommendations to regulatory instances (EC, ISO...)

Agnes Delaborde - Safety of Artificial Intelligence embedded in machinery

LABORATOIRE
NATIONAL
DE MÉTROLOGIE
ET D'ESSAIS
LNE

SAFETY OF ARTIFICIAL INTELLIGENCE EMBEDDED IN MACHINERY

Thank you!

Feel free to contact me: agnes.delaborde@lne.fr

LNE – French national laboratory for metrology and testing

LABORATOIRE
NATIONAL
DE MÉTROLOGIE
ET D'ESSAIS

LNE