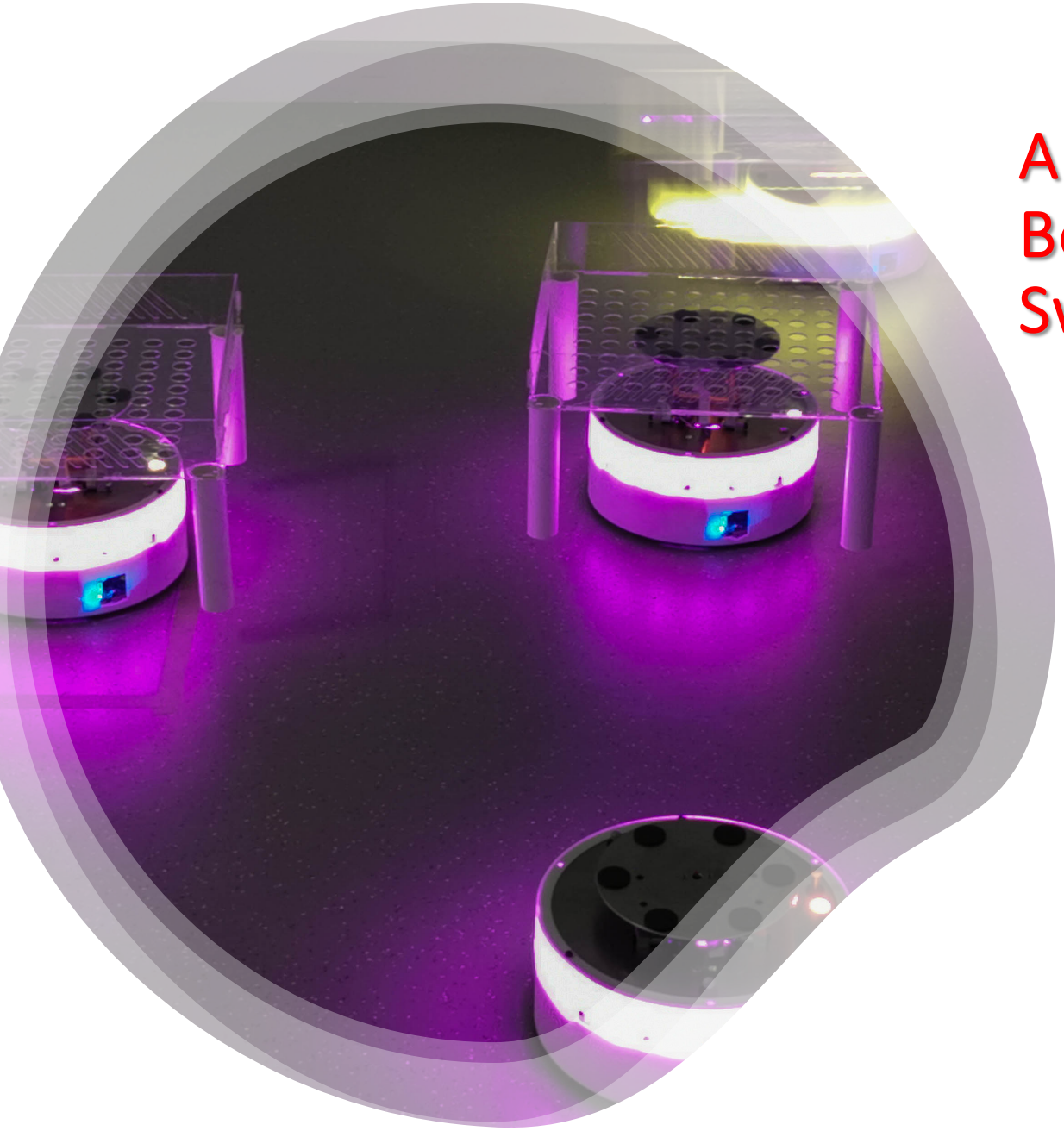

AERoS: Assurance of Emergent Behaviour in Autonomous Robotic Swarms

Dhaminda B. Abeywickrama[1], James Wilson[1], Suet Lee[1], Greg Chance[1], Peter D. Winter[1], Arianna Manzini[1], Ibrahim Habli[2], Shane Windsor[1], Sabine Hauert[1], Kerstin Eder[1]

[1]*University of Bristol*

[2]*University of York*

WAISE'2023 Workshop @ SafeComp 2023 Conference
Toulouse, France

19 Sep. 2023

What is a Swarm?

Autonomous

Large Number of Agents (10+)

Restrained Homogeneity

Local Sensing & Communication

Emergent Behaviour (EB) & Assurance

Simple algorithms are executed by individual agents

EB arises from the interactions of the agents with each other and the environment

EB can be difficult to model & can pose a critical challenge for assurance

"How do you ensure safety of a swarm where the swarm's behaviour is an emergent consequence of the interaction of individual agents with each other and their environment?"

Case Study: Cloakroom

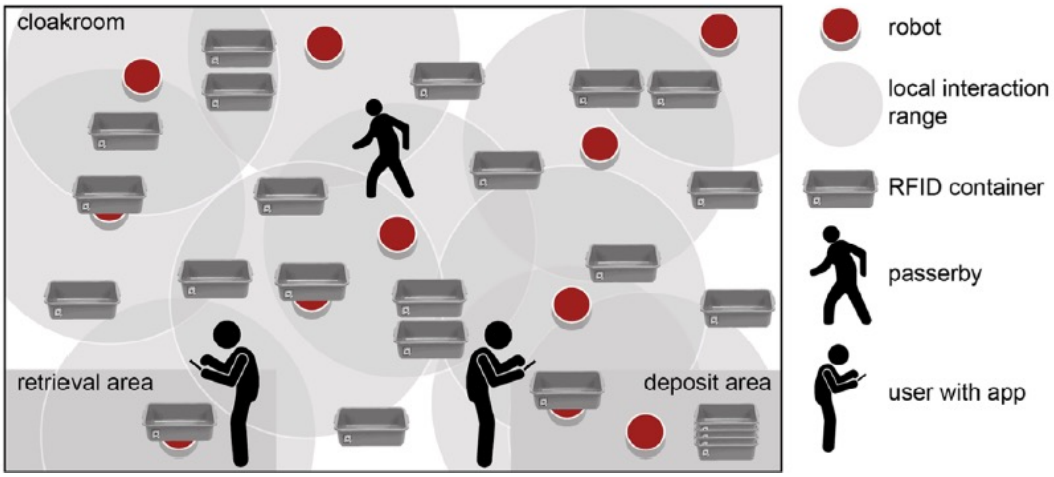

Fig. 1: Pop-up cloakroom powered by a swarm of robots using distributed situational awareness [Jones et al., 2020].

Fig. 2: Swarm of DOTS moving boxes in our lab environment [Jones et al., 2022]

Standards and Guidance Reviewed

- **Standards applicable in the current study:**
  - **IEEE P7001 – Transparency of Autonomous Systems**
  - **ISO 13482:2014**, Robots and robotic devices — Safety requirements for personal care robots
  - **ISO/TR 23482-2:2019**, Robotics — Application of ISO 13482 — Part 2: Application guidelines
  - **ISO/TR 23482-1:2020**, Robotics — Application of ISO 13482 — Part 1: Safety-related test methods
- **Other standards reviewed:**
  - **ISO 10218-1:2011**, Robots and robotic devices — Safety requirements for industrial robots — Part 1: Robots
  - **ISO 10218-2:2011**, Robots and robotic devices — Safety requirements for industrial robots — Part 2: Robot systems and integration
  - **ISO/TS 15066:2016**, Robots and robotic devices — Collaborative robots
  - **ISO 18646-1:2016**, Robotics — Performance criteria and related test methods for service robots — Part 1: Locomotion for wheeled robots
  - **ISO 18646-2:2019**, Robotics — Performance criteria and related test methods for service robots — Part 2: Navigation
  - **ISO 18646-3:2021**, Robotics — Performance criteria and related test methods for service robots — Part 3: Manipulation

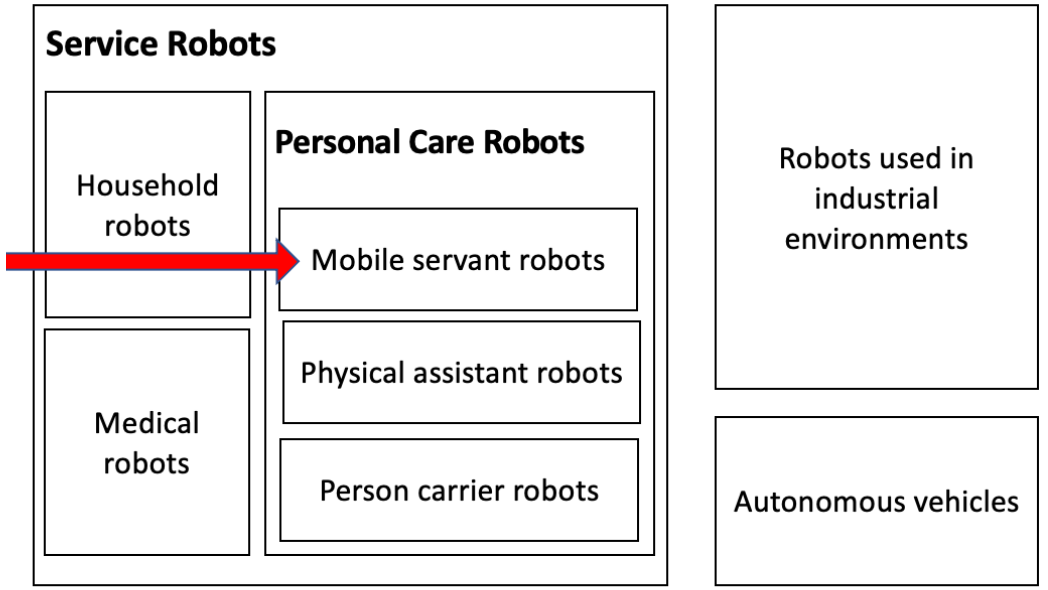

Fig. 3: Categorization of mobile service robots and relation with other relevant areas [adapted from ISO/TR 23482-2:2019]

AMLAS Process

- Assurance of Machine Learning for use in Autonomous Systems (AMLAS):
  - Provides **guidance** on how to systematically integrate **safety assurance** into the development of ML components [Hawkins et al., 2021]
- Assurance activities performed in **parallel** to the development of ML component [Hawkins et al., 2021]
- **Iterative** process
- Explicit and structured **safety case:**
  - Set of **argument patterns** (in goal structuring notation), and the underlying **assurance activities** instantiated to develop ML **safety cases**

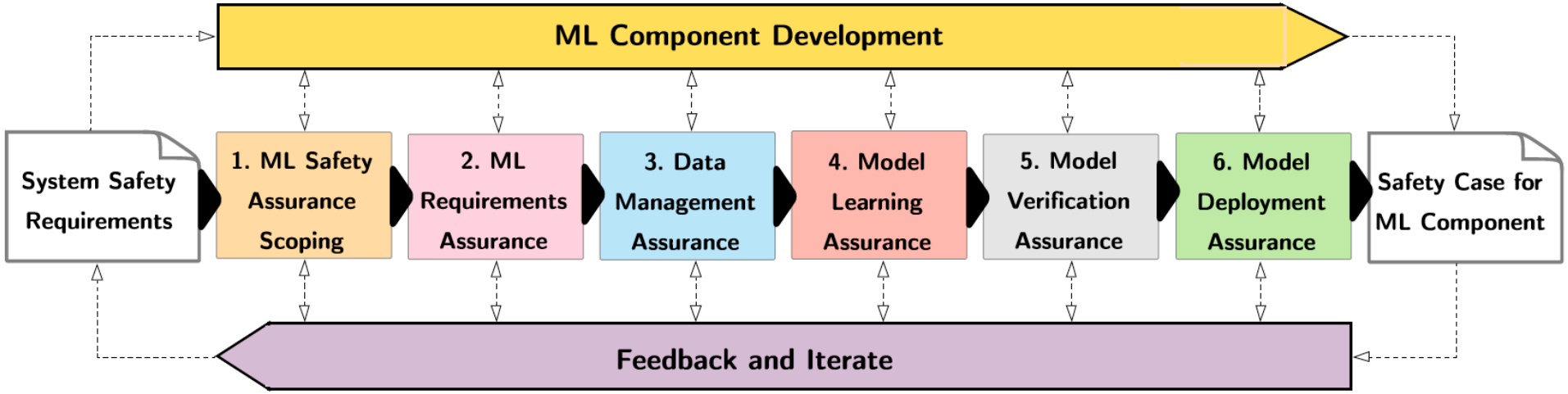

Fig. 4: AMLAS process [Hawkins et al., 2021].

AERoS Process

- **Safety assurance process** based on AMLAS targeting robotic swarms:
  - **Six** main stages
- **Iterative** by design
- Assurance activities performed **parallel** to EB development
- Each stage describes its **inputs**, **outputs**, **assurance activities** & produced **artefacts**:
  - **EB safety case** for swarms

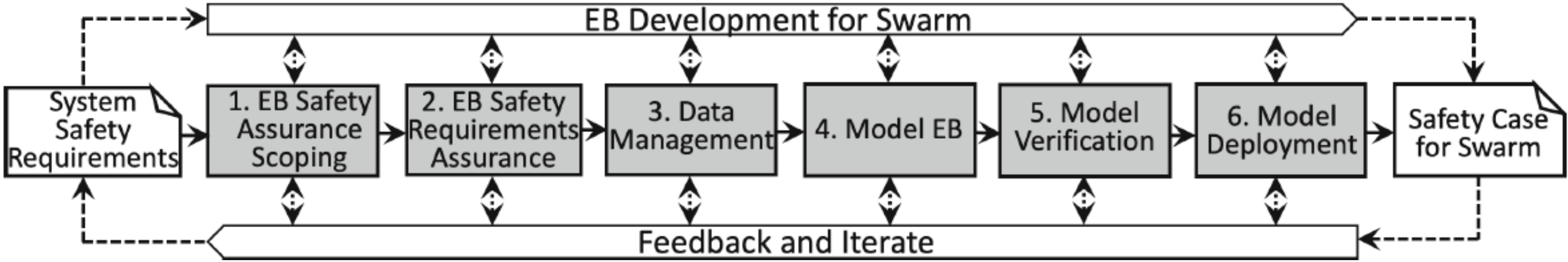

Fig. 5: The AERoS process with its six stages adapted from AMLAS.

Stage 2: EB Safety Requirements Assurance

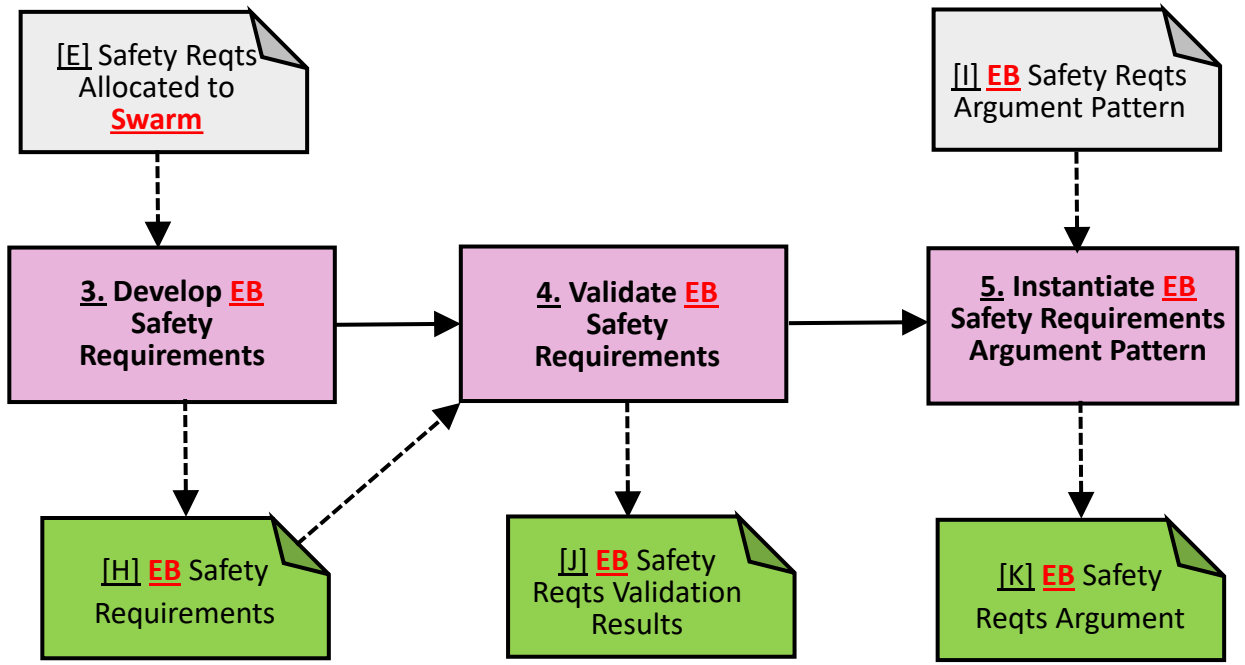

Figure 10: AERoS EB safety requirements assurance process.

Stage 2: Safety Requirements

Table 1. Examples of performance, adaptability, environmental, and human-safety safety requirements for the cloakroom scenario.

RQ	Performance Requirements
1.1	The swarm *shall* experience $<$ **1 high-impact (V $>$ 0.5 m/s)** collisions across **a day** of **faultless** operation
1.2	The swarm *shall* experience $<$ **0.1%** increase in **high-impact** collisions across **a day's** operation with **full communication faults** occurring in **10%** of the swarm
1.3	The swarm *shall* experience $<$ **0.1%** increase in **high-impact** collisions across **a day's** operation with **half-of-wheels motor faults** occurring in **50%** of the swarm
1.4	The swarm *shall* experience $<$ **2 high-impact (V $>$ 0.5 m/s)** collisions across **a day** of **faulty** operation
1.5	The swarm agents *shall* **weigh $<$ 3 kg** and shall have **acceleration $<$ 4 m/s** so that the **maximum collision force** in the swarm is within acceptable bounds
1.6	The swarm agents *shall* only carry objects of **weight $<$ 2 kg**

Faultless Operations

Failure Modes

Worst case

Stage 2: Safety Requirements

Adaptability Requirements	
2.1	The swarm *shall* have < **10%** of its agents **stationary*** outside of the **delivery site** at a given time. *Agents are considered stationary once they have not moved for > **10 s**
2.2	All agents of the swarm *shall* move at least every **100 s** if outside of the **delivery site**
2.3	The swarm *shall* experience < **10%** increase in the **number of stationary agents** at any time with **half-of-wheels motor faults** occurring in **50%** of the swarm
2.4	The swarm agents *shall* experience < **10%** increase in **stationary time** with **half-of-wheels motor faults** occurring in **50%** of the swarm
2.5	The swarm *shall* experience < **10%** increase in **number of stationary agents** at any given time with **full communication faults** occurring in **10%** of the swarm
2.6	The swarm agents *shall* experience < **10%** increase in **stationary time** with **full communication faults** occurring in **10%** of the swarm

Stage 2: Safety Requirements

Environmental Requirements	
3.1	The swarm *shall* perform as required in environmental density levels **0–4 p_o of objects** (sum of boxes and agents per m^2) in the environment
3.2	The swarm *shall* perform as required when **floor incline** is **0–20°**
3.3	The swarm *shall* perform as required in a **dry environment**
Human-Safety Requirements	
4.1	The swarm agents *shall* travel at speeds of less than **0.5 m/s** when within **2 m** distance of a **trained human** (a worker who has received relevant training)
4.2	The swarm agents *shall* travel at speeds of less than **0.25 m/s** when within **3 m** distance of an **attendee**
4.3	The swarm agents *shall* only come within **2 m** distance of a **human** $<$ **10** times collectively across **1000 s** of **faultless** operations
4.4	The swarm *shall* experience $<$ **10%** increase in **human encounters** across **1000 s** of operation with **full communication faults** occurring in **10%** of the swarm
4.5	The swarm *shall* experience $<$ **10%** increase in **human encounters** across **1000 s** of operation with **half-of-wheels motor faults** occurring in **50%** of the swarm
4.6	The swarm agents *shall* only come within **2 m** distance of a **human** $<$ **20** times collectively across **1000 s** of **faulty** operations.

AERoS Process

- **Safety assurance process** based on AMLAS targeting robotic swarms:
  - **Six** main stages
- **Iterative** by design
- Assurance activities performed **parallel** to EB development
- Each stage describes its **inputs**, **outputs**, **assurance activities** & produced **artefacts**:
  - **EB safety case** for swarms

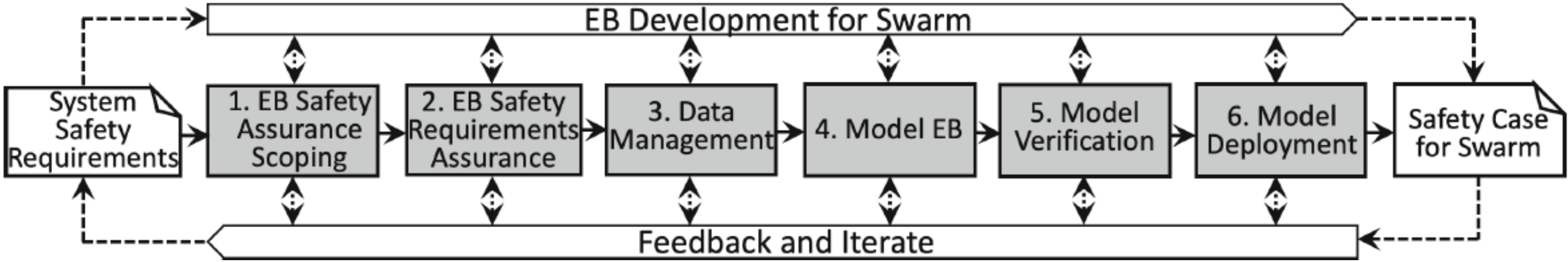

Fig. 5: The AERoS process with its six stages adapted from AMLAS.

Stage 3: Data Management

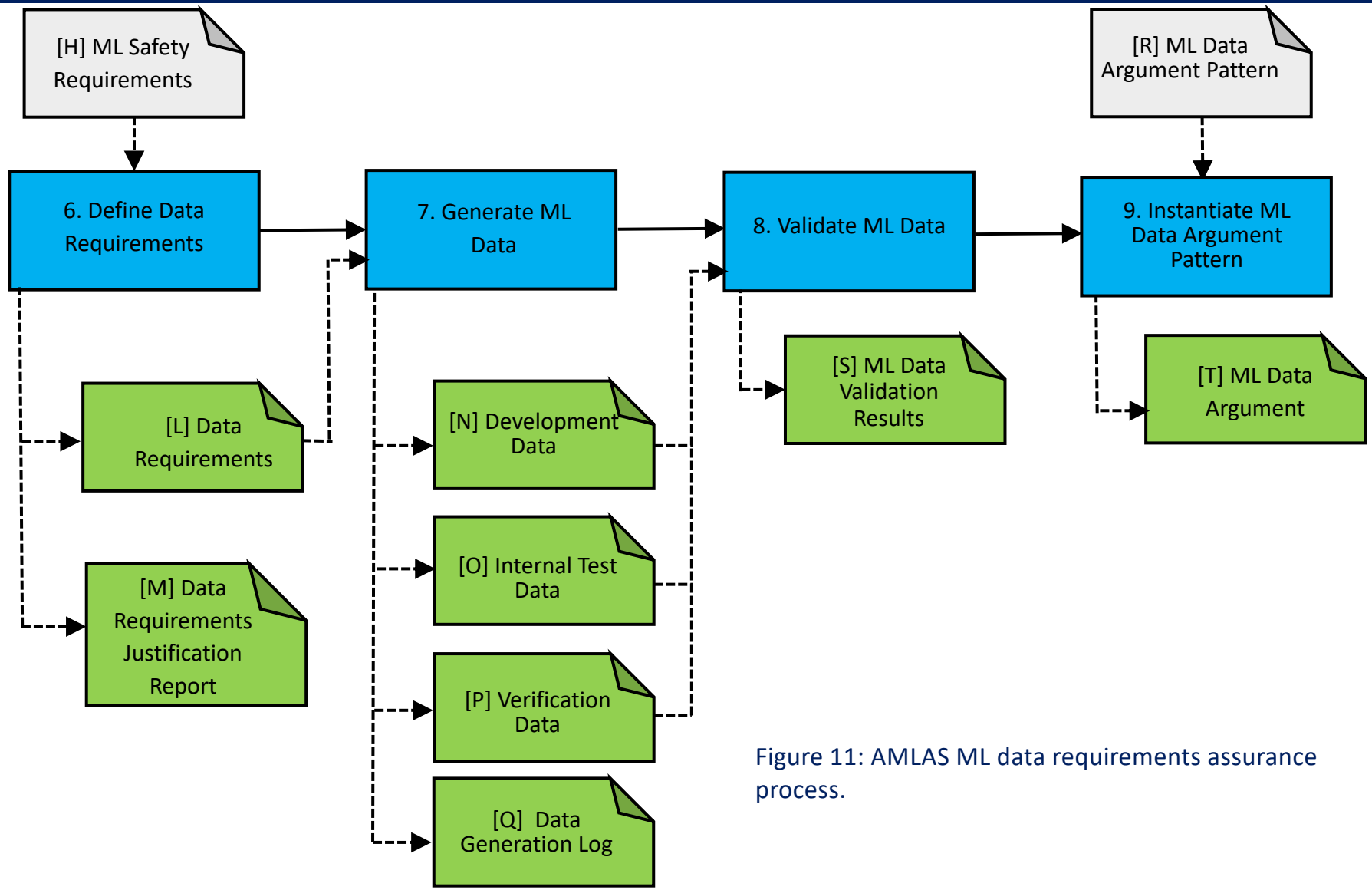

Figure 11: AMLAS ML data requirements assurance process.

Stage 3: Data Management

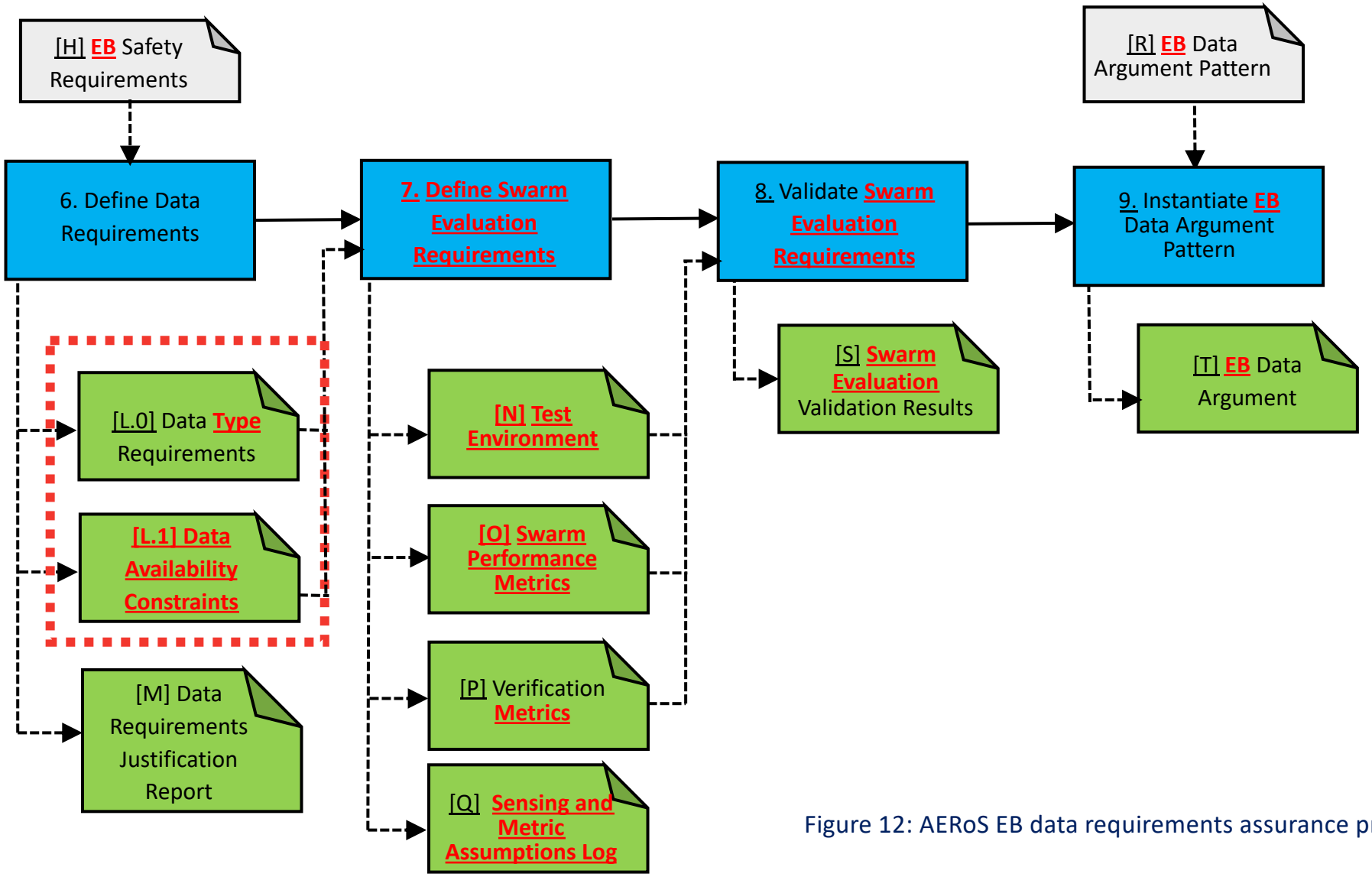

Figure 12: AERoS EB data requirements assurance process.

Conclusion

- Safety Assurance Process for autonomous robotic swarms
- **Limitations**:
  - Individual robots' adaptation
  - Additional swarm use cases
- **Future work**:
  - Other assurance properties like regulation & ethics

References

[Jones et al., 2020]. S. Jones, E. Milner, M. Sooriyabandara, and S. Hauert, “Distributed Situational Awareness in Robot Swarms”, in Adv. Intell. Syst, 2: 2000110. https://doi.org/10.1002/aisy.202000110.

[Hawkins et al., 2021] Hawkins, R., Paterson, C., Picardi, C., Jia, Y., Calinescu, R., and Habli, I., “Guidance on the Assurance of Machine Learning in Autonomous Systems (AMLAS)”, arXiv e-prints, 2021.

[Jones et al., 2022] S. Jones, E. Milner, M. Sooriyabandara, and S. Hauert, “DOTS: An Open Testbed for Industrial Swarm Robotic Solutions”, arXiv. 2022. https://arxiv.org/abs/2203.13809.

Thank you!

dhaminda.abeywickrama@bristol.ac.uk

suet.lee@bristol.ac.uk

Full Paper