ATLAS

Verification Models and Phantoms

for autonomous robotic systems

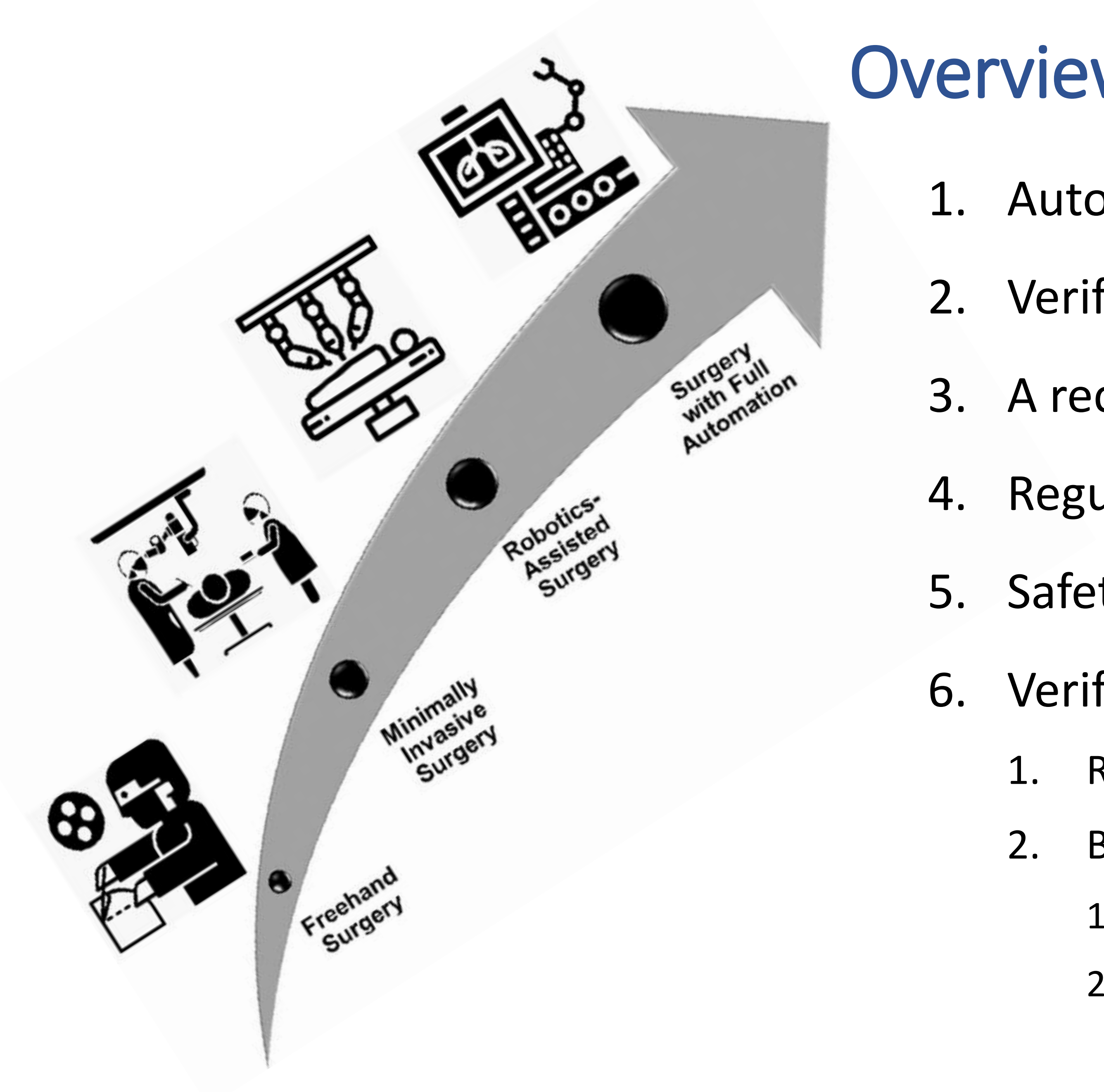
NTA1

Diego Dall'Alba
UNIVR - Altair Robotics Lab

NTA1 @ Virtual Event 22 – 24 March 2021

UNIVERSITÀ di VERONA
Dipartimento di INFORMATICA
ALTAIR robotics lab

Overview



1. Automation VS Autonomy
2. Verification model in automation
3. A recap of autonomy levels in surgical robotics
4. Regulations related to Autonomous systems
5. Safety verification in autonomous systems
6. Verification models
    1. Risk analysis based
    2. Based on simulation
        1. Virtual environment
        2. Synthetic environment (phantom)

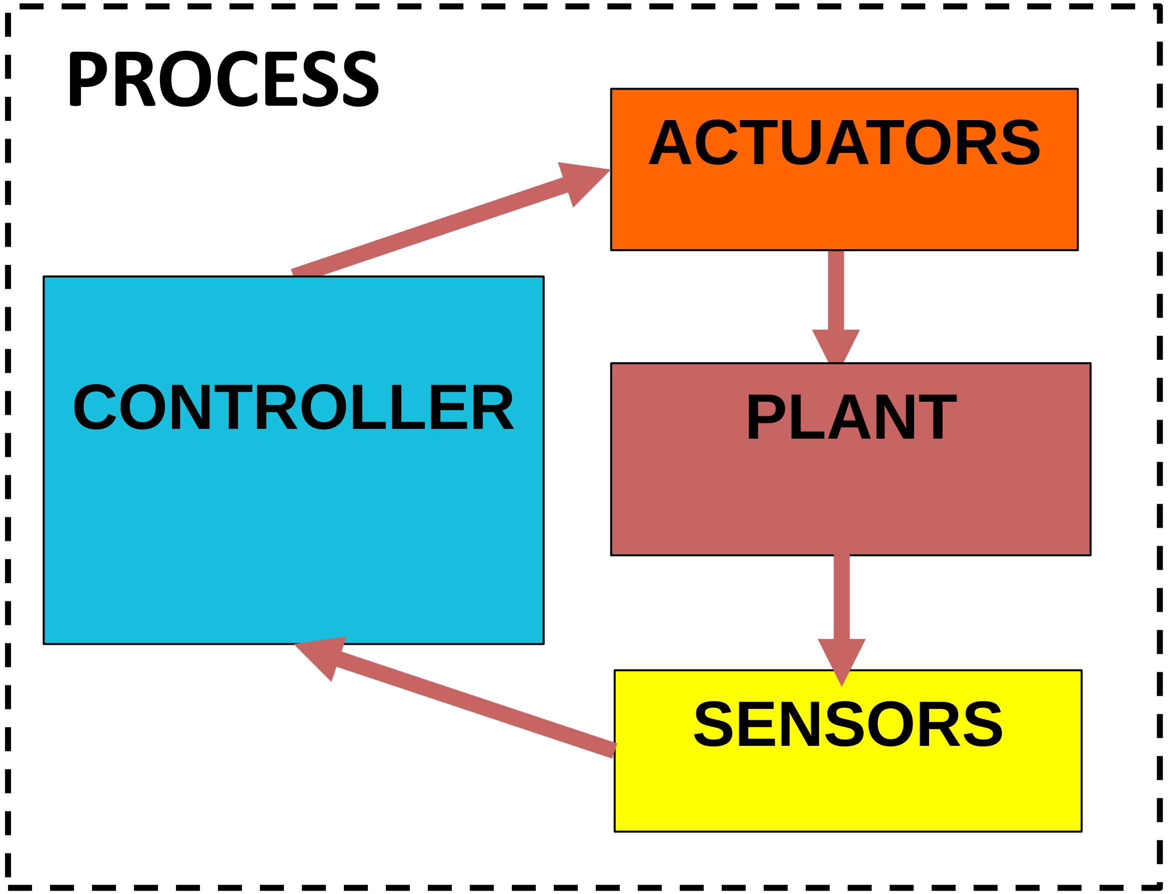
Automation VS Autonomy

- There is some ambiguity:
  - **What is "Autonomy" for some people is "Automation" for other**
- **Automation:**
  - **The technique of making an apparatus, a process, or a system operate with a self-acting or self-regulating mechanism (Merriam Webster)**
  - **It is the ability to carry out actions without Human interventions.**
  - These actions are well defined, can be described with precise rule, and they are done in a known and well-structured environment. Automation has a small and defined degree of adaptation.
- **Autonomy:**
  - **The quality or state of being "self-governing", (Merriam Webster)**
  - **It is the ability to carry out complex tasks and take cognitive decisions.**
  - These actions are defined in general terms, executed adapting previous knowledge, in an unknown and uncertain environment with adaptation learned by the system.

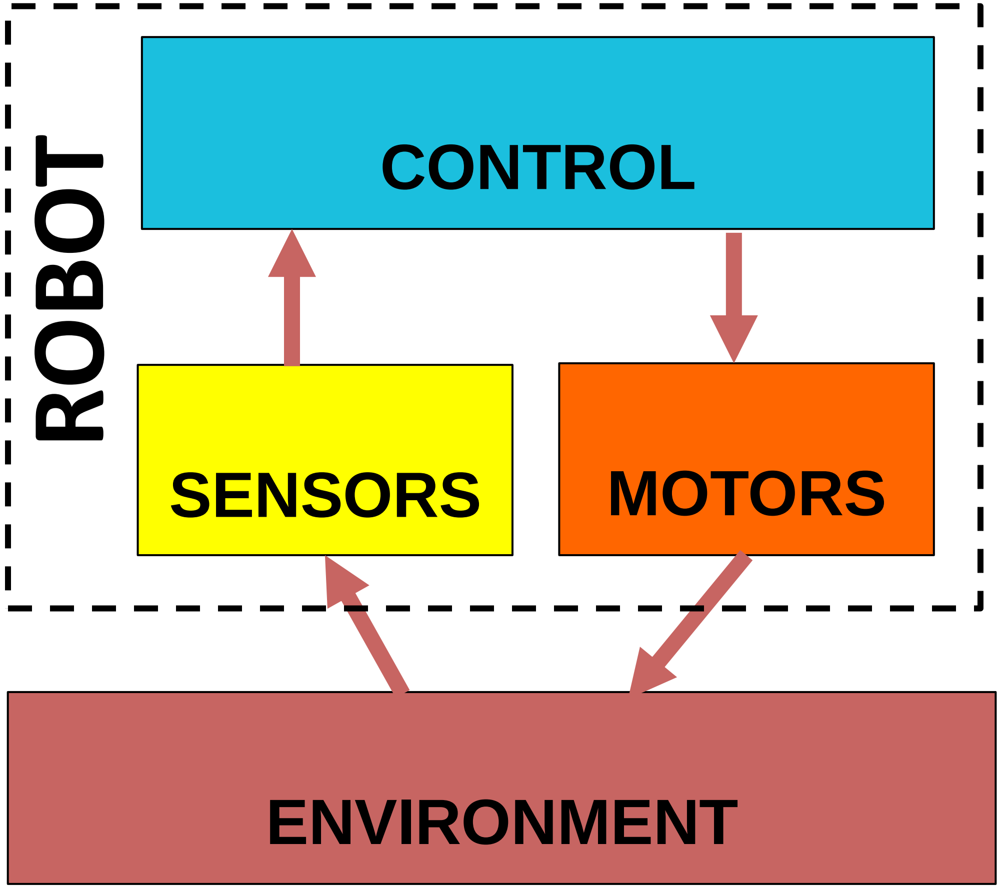
The robot makes the new plan

Automatic Control and Robotics

Automatic control is the application of **control** theory for regulation of processes without direct human intervention.

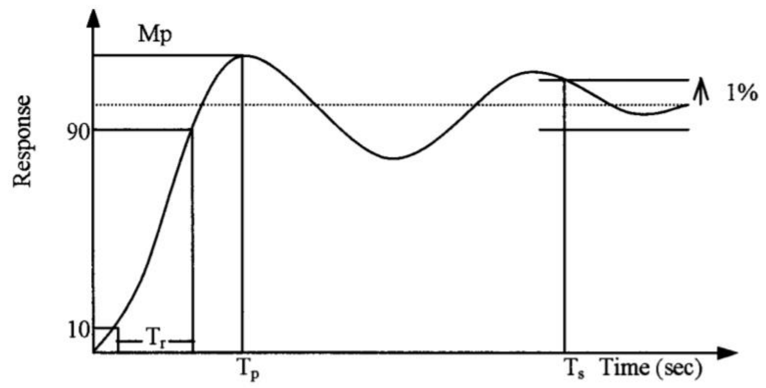


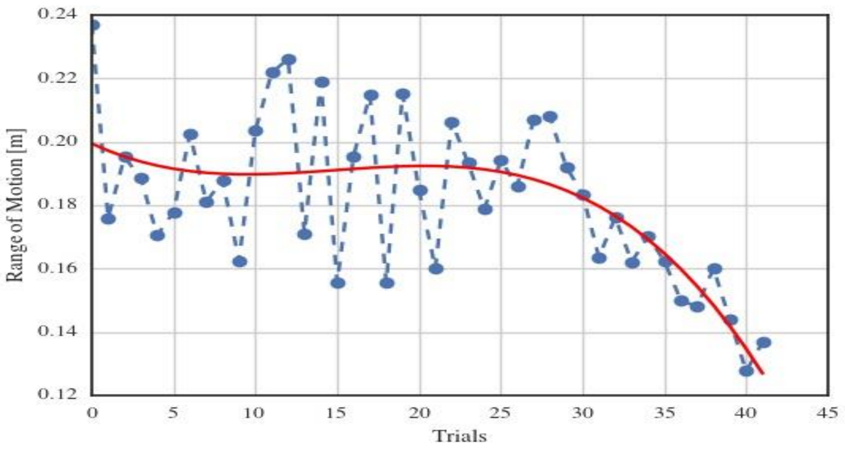
Robotics is the **Intelligent Connection between Perception and Action** to achieve a desired **Result**



Verification models in Automation and Robotics

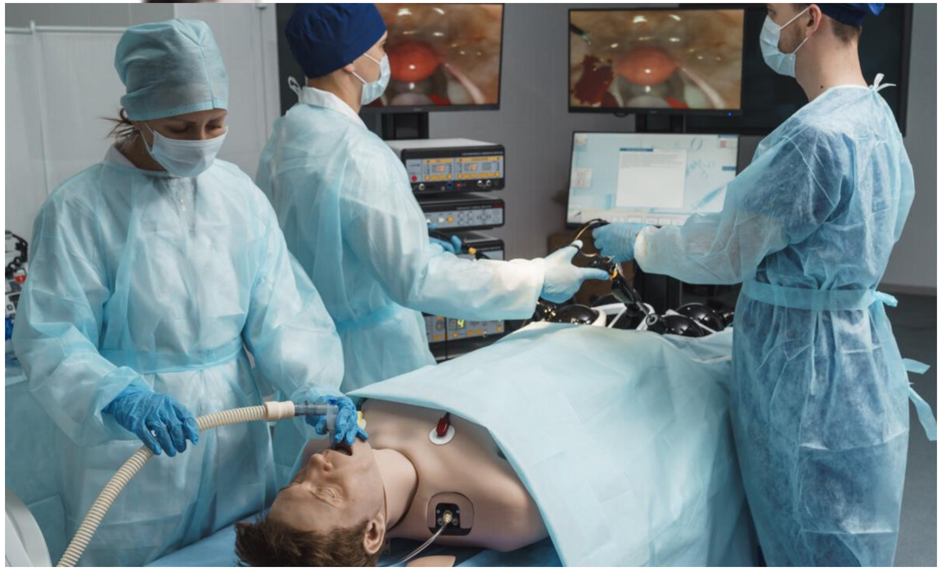
- System theory is providing a rigorous framework for the evaluation of automated (surgical) systems
- We could select a specific property and verifying the system according to a chosen metrics, for example we could evaluate the stability of a control system by measuring its tracking error.
- There are many approach for this verification process already applied in complex domains (for instance aeronautics or nuclear applications):
  - Theoretical demonstration
  - Simulation based
  - Formal methods



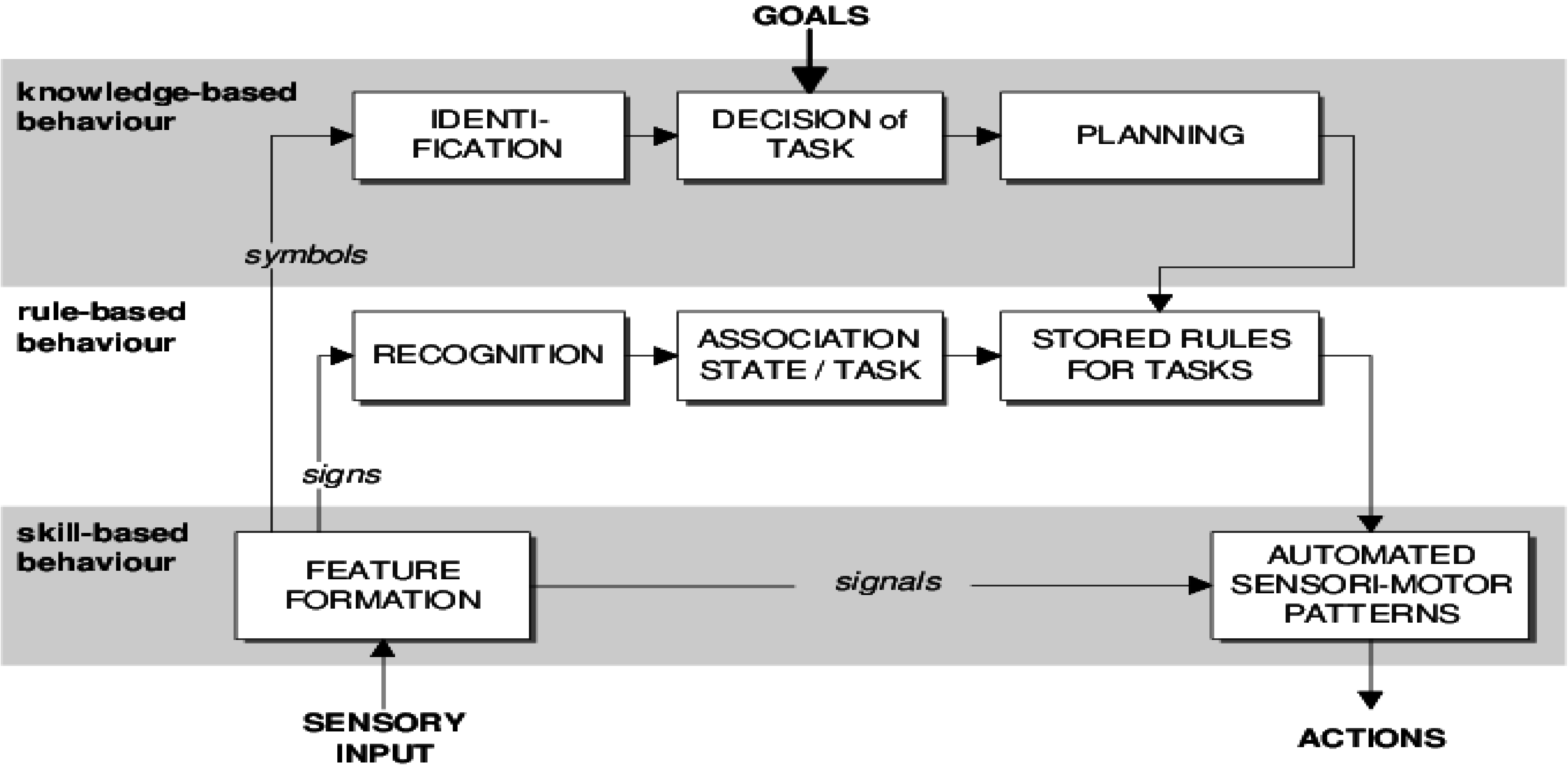


Verification models in Automation and Robotics

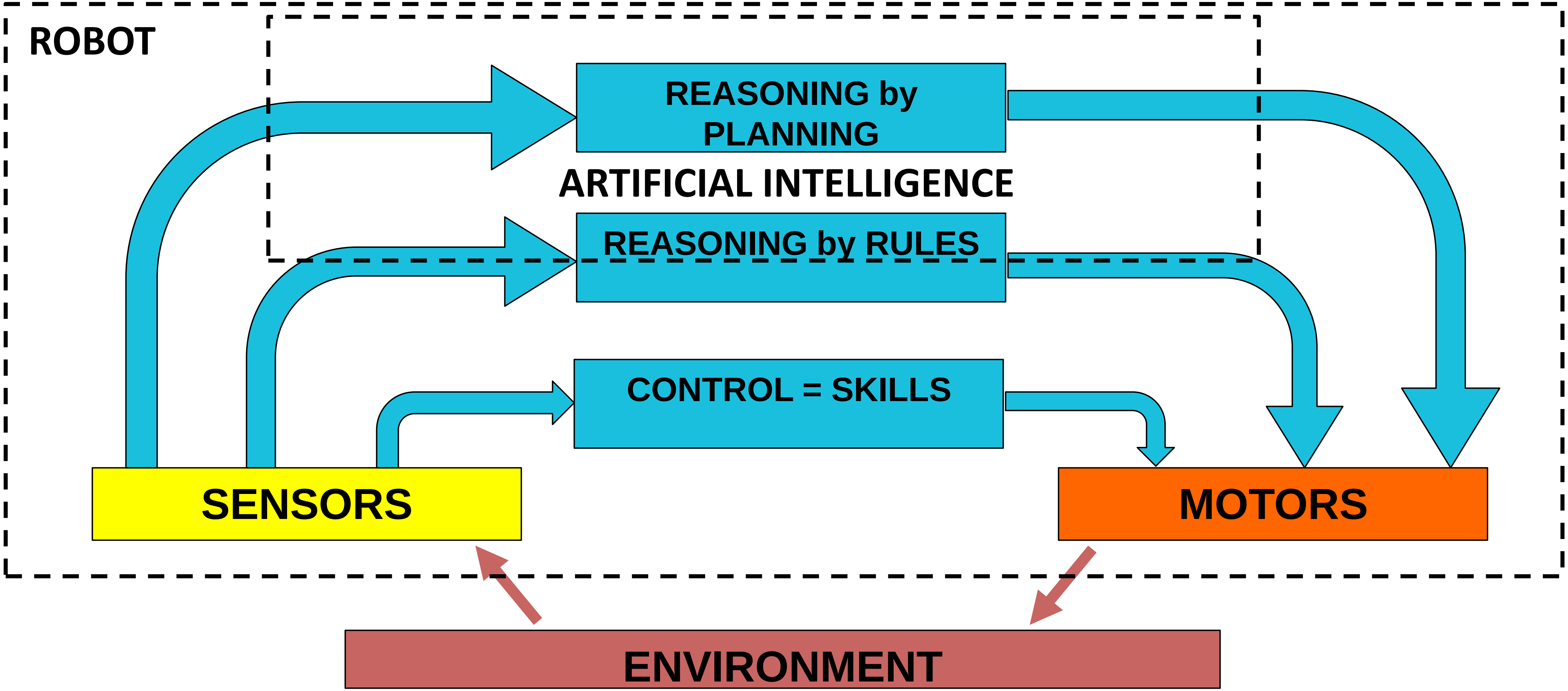
- In clinical practice, biomedical devices (included automatized system) can be tested in different scenarios that are progressively more realistic and complex.
- The main and most common categories are:
  1. virtual reality simulators;
  2. dry lab environment (with increasingly complexity and realism);
  3. wet lab environment (e.g., ex-vivo or in-vivo animal experiment);
  4. patient trials.



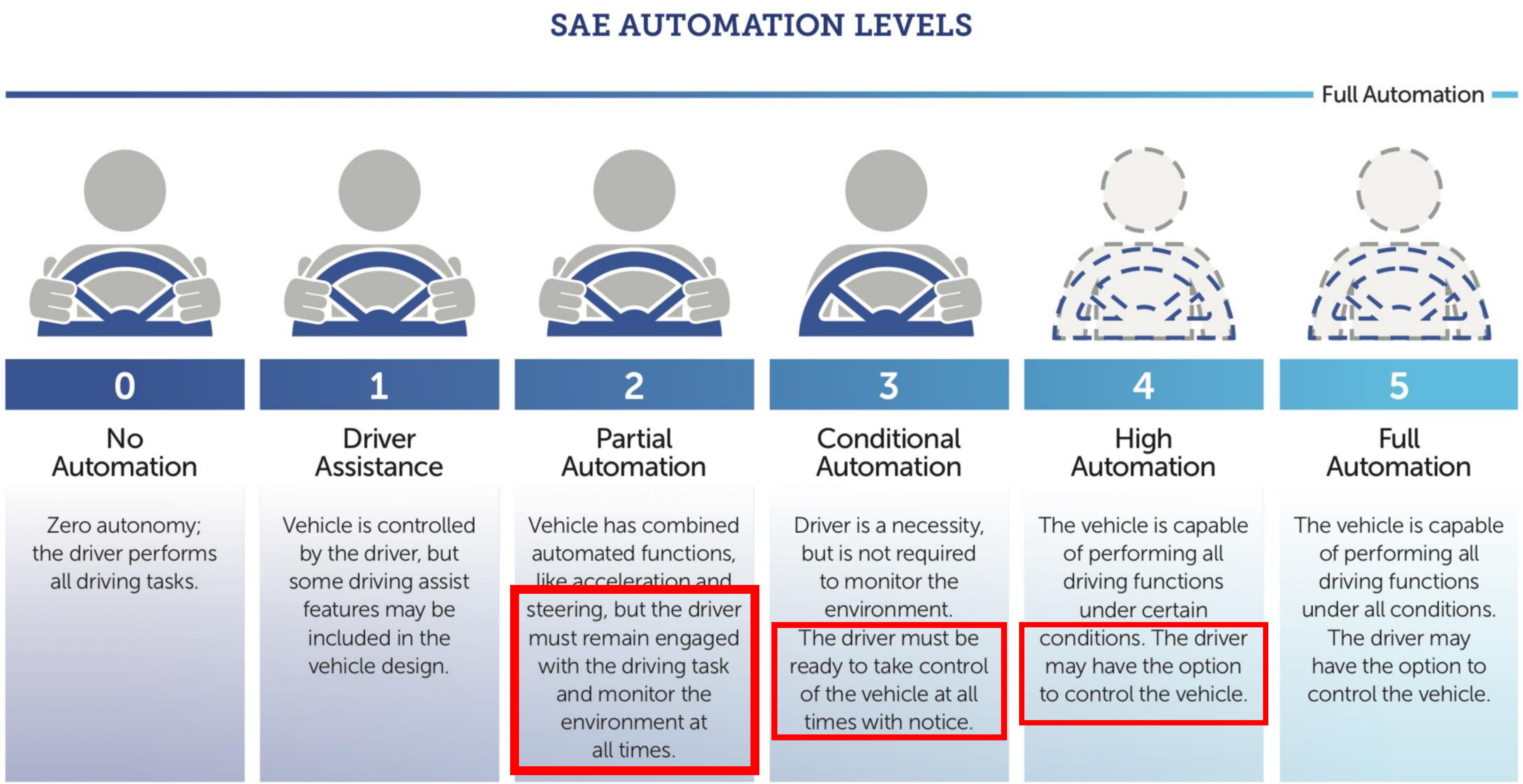
Model of Human Reasoning (Rasmussen 1985)



Autonomous Robotics

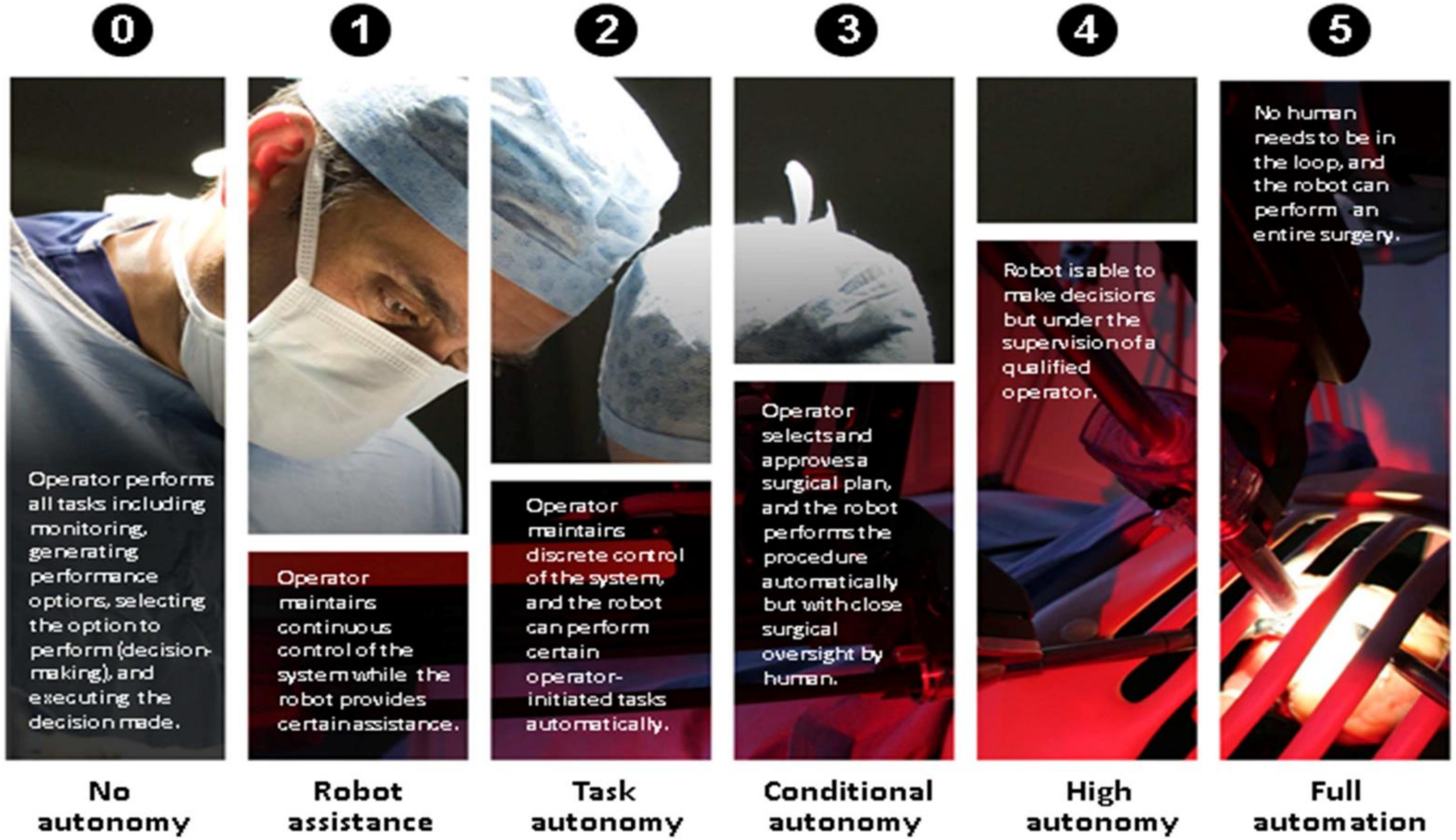


Classification of Automated Driving Systems



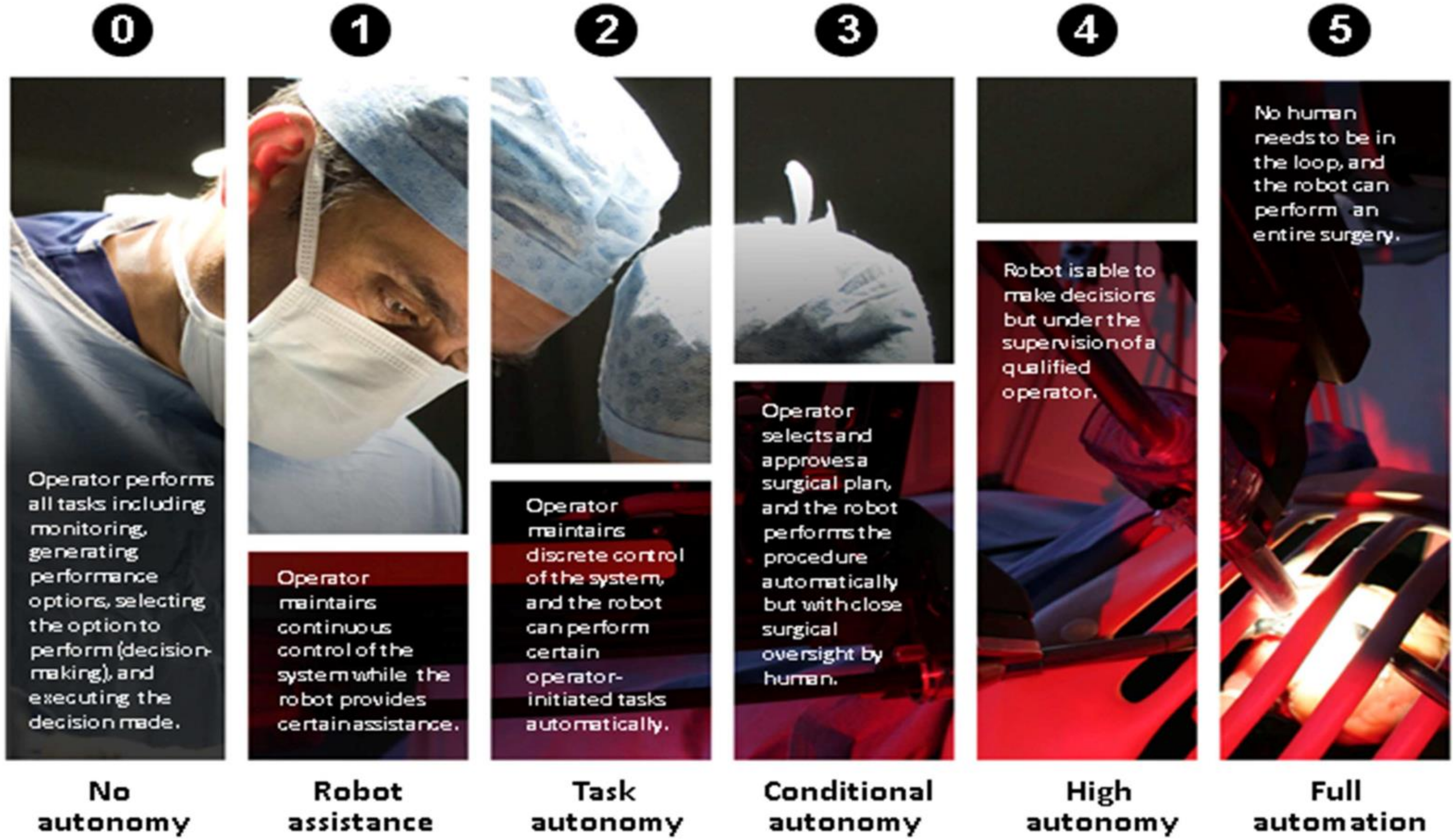
US, Department of Transportation Automated Driving Safety Guidelines, September 2017

Classification of Autonomous Robotic Surgery



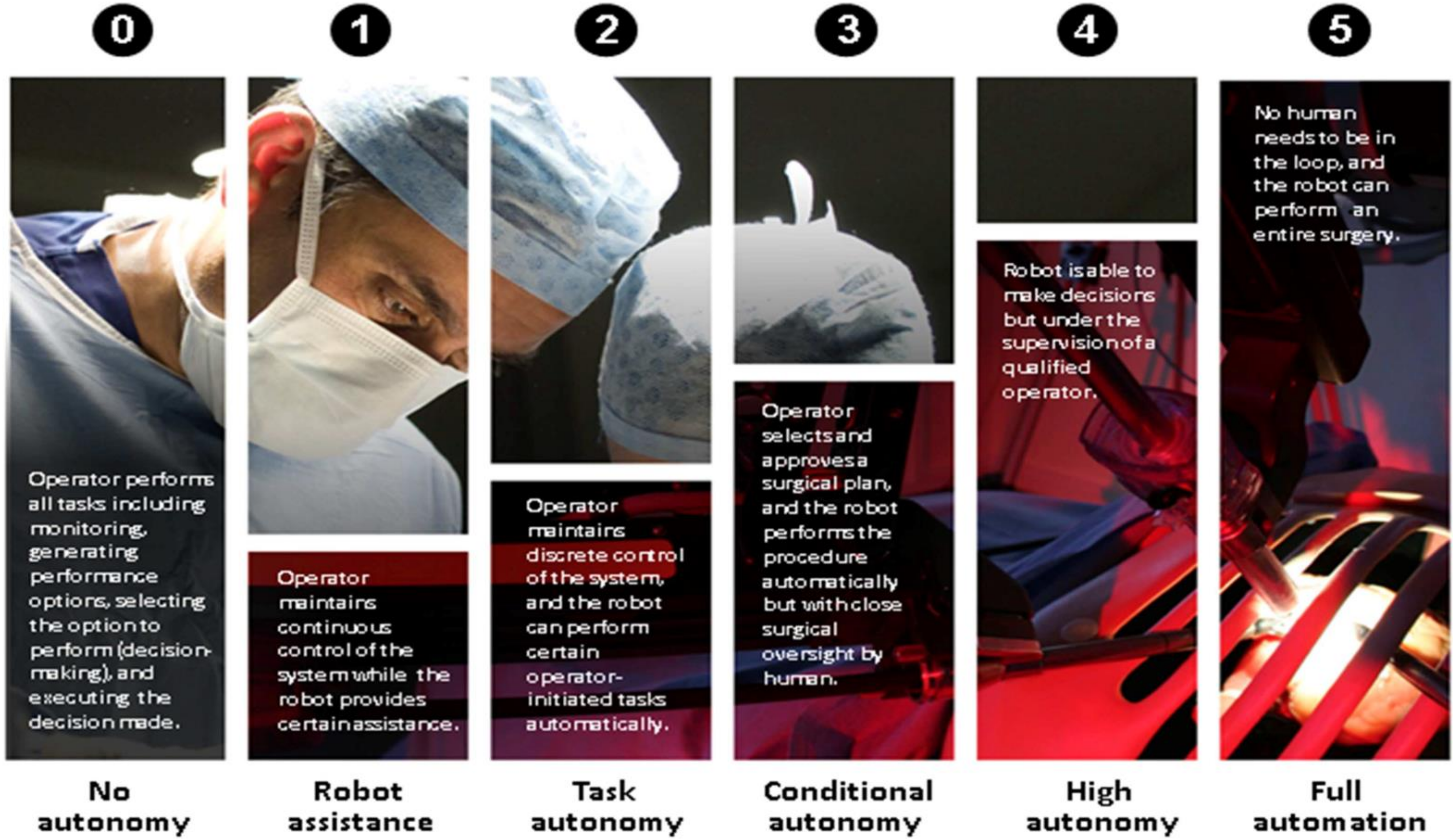
Level 0: No autonomy. This level includes tele-operated robots or prosthetic devices that respond to and follow the user’s command.

Classification of Autonomous Robotic Surgery



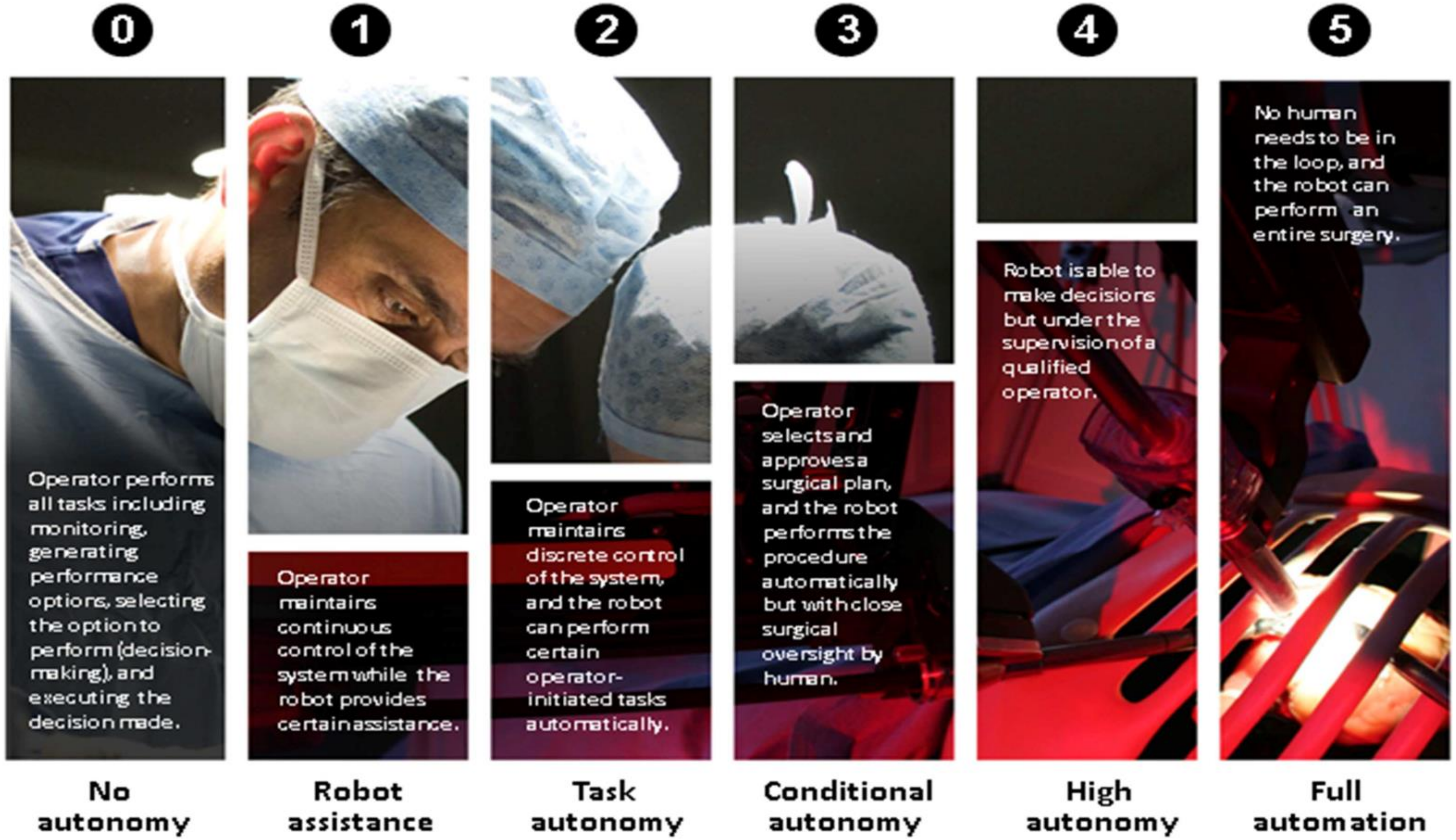
Level 1: Robot assistance. The robot provides some mechanical guidance or assistance during a task while the human has continuous control of the system (e.g. virtual fixtures).

Classification of Autonomous Robotic Surgery



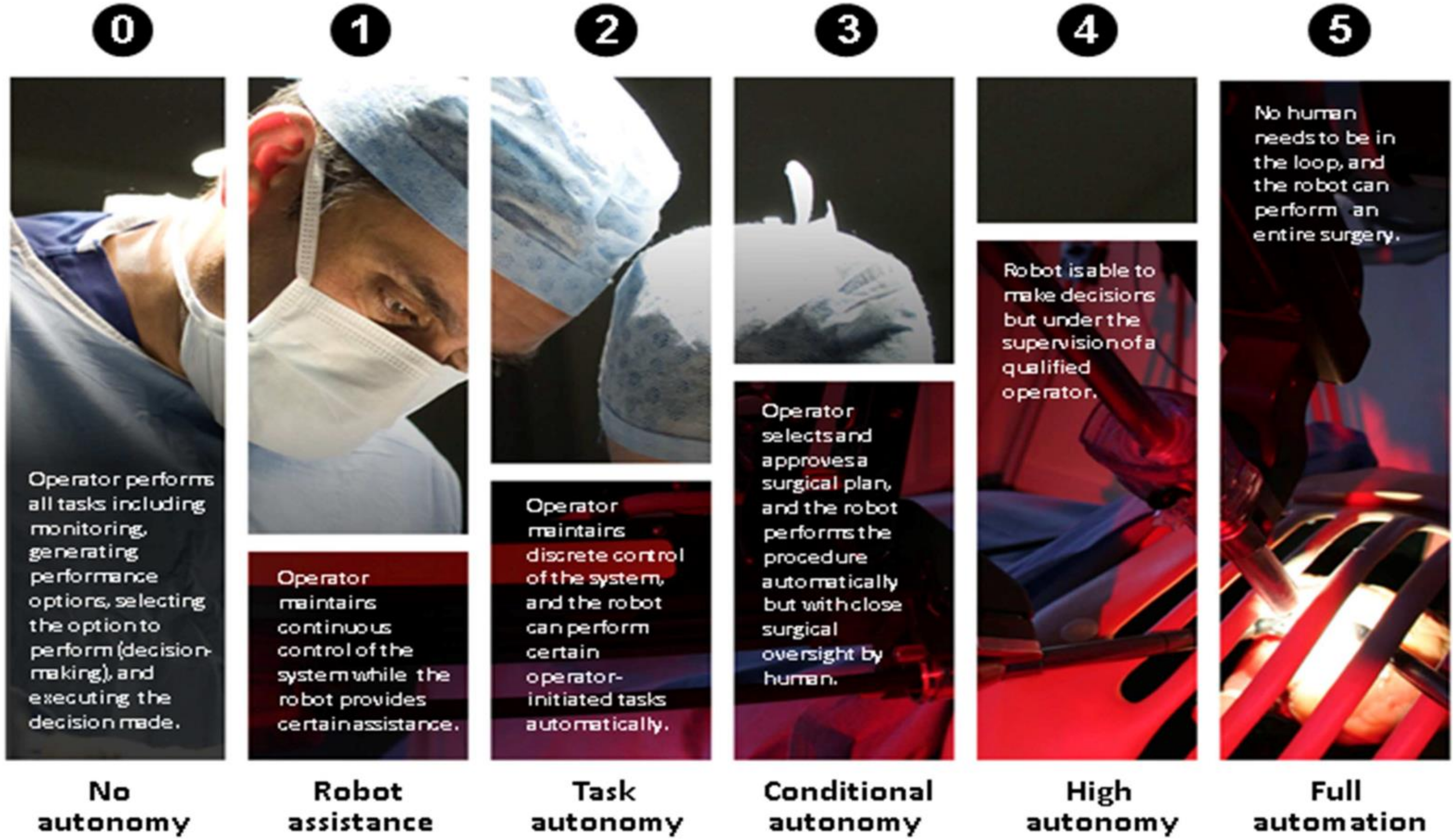
Level 2: Task autonomy. The robot is autonomous for specific tasks initiated by a human who has discrete, rather than continuous, control of the system (e.g. surgical suturing).

Classification of Autonomous Robotic Surgery



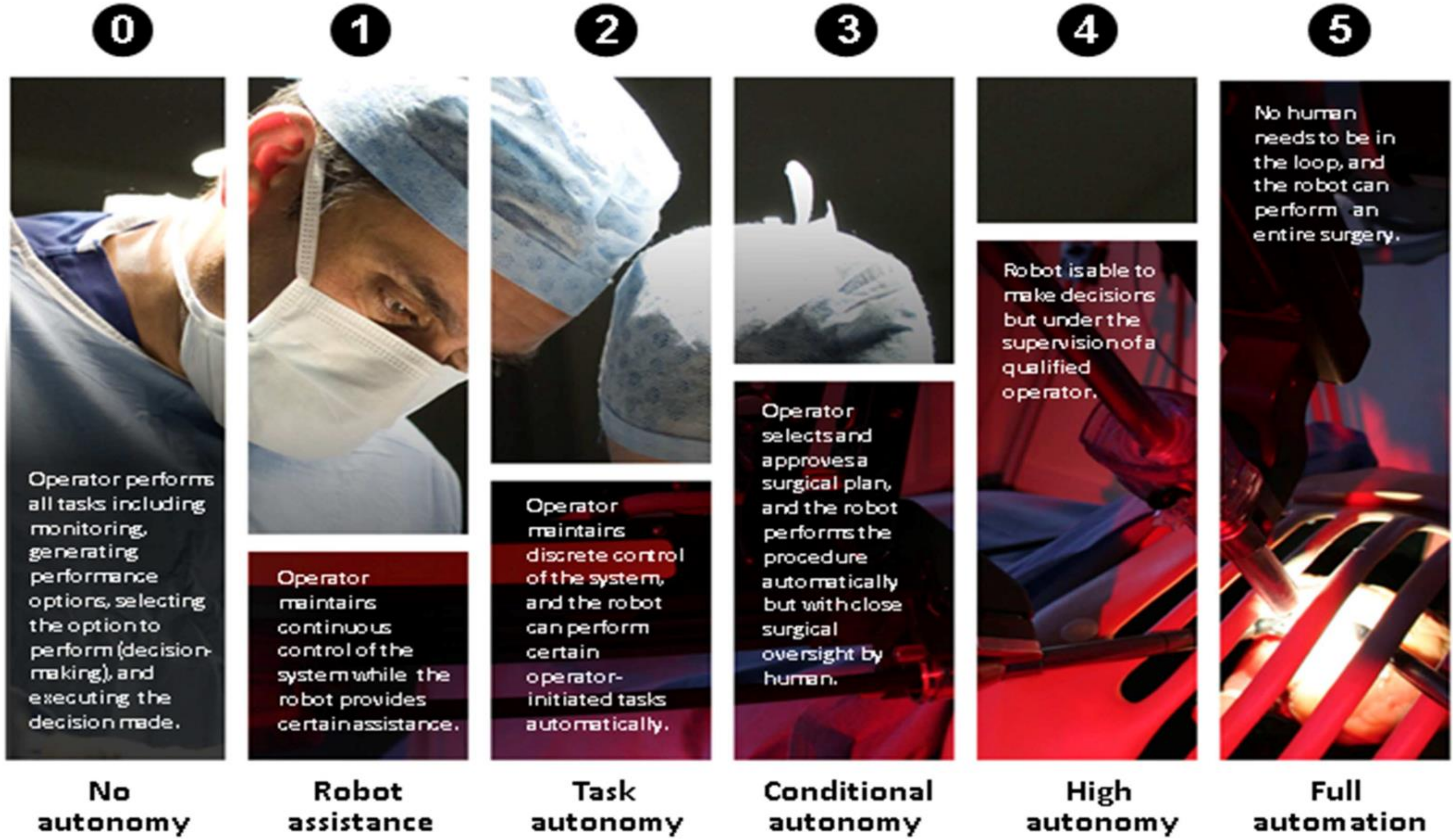
Level 3: Conditional autonomy. A system generates task strategies but relies on the human to select from among different strategies or to approve an autonomously selected strategy

Classification of Autonomous Robotic Surgery



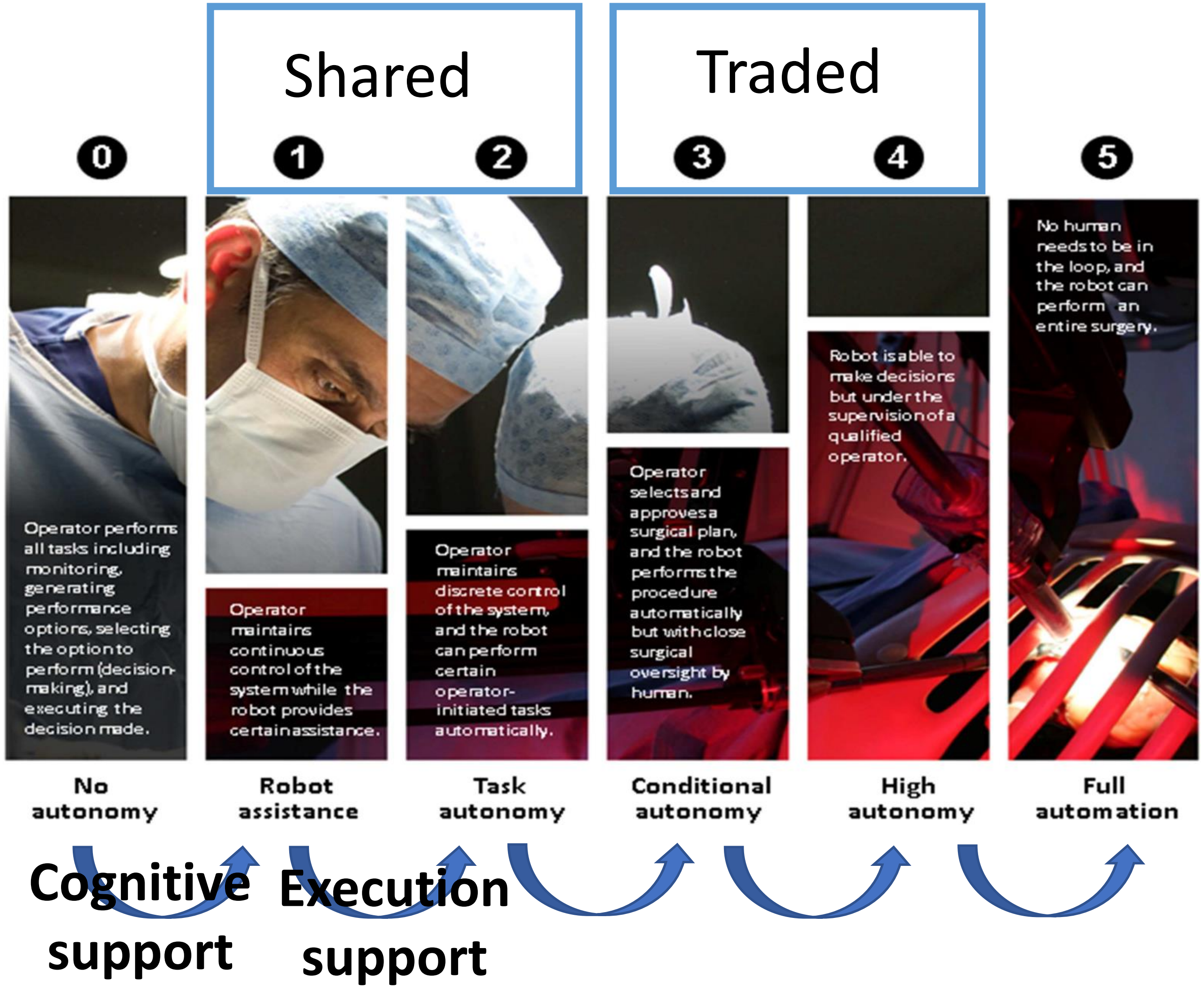
Level 4: High autonomy. The robot can make medical decisions but under the supervision of a qualified doctor (robotic resident).

Classification of Autonomous Robotic Surgery



Level 5: Full autonomy (no human needed). This is a “robotic surgeon” that can perform an entire surgery and is currently in the realm of science fiction.

Transition between levels



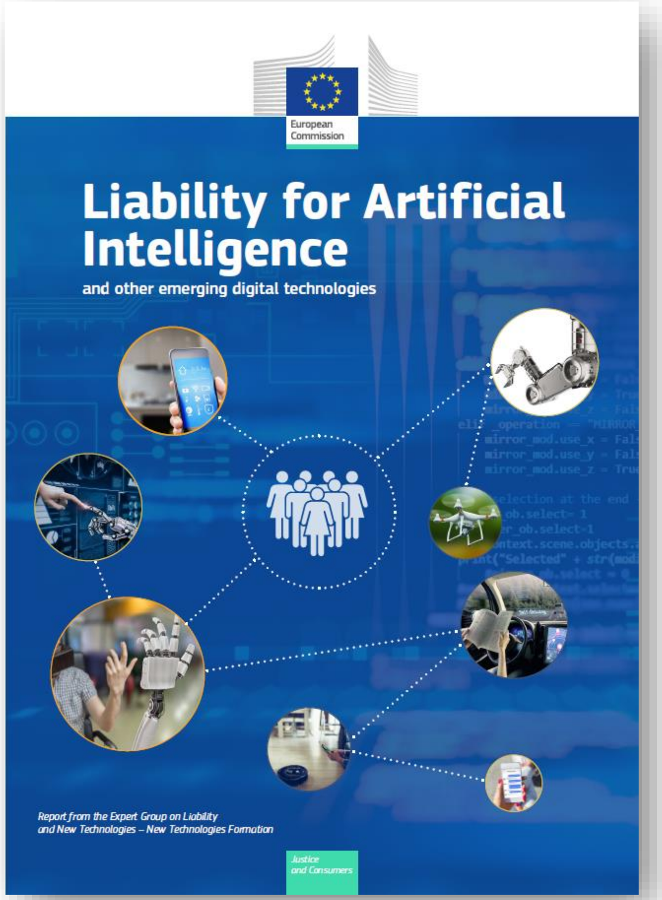
- We expect a sequential process in the transions amongs Autonomy levels
- Moving from shared control, to traded control and finally (Level 5) autonomous control.

Not all the transitions have the same difficult, for example:

- Level 0 →Level 1 requires the development of cognitive support. <u>Missing Reasoning components</u>
- Level 1 → Level 2 requires the development of execution support.

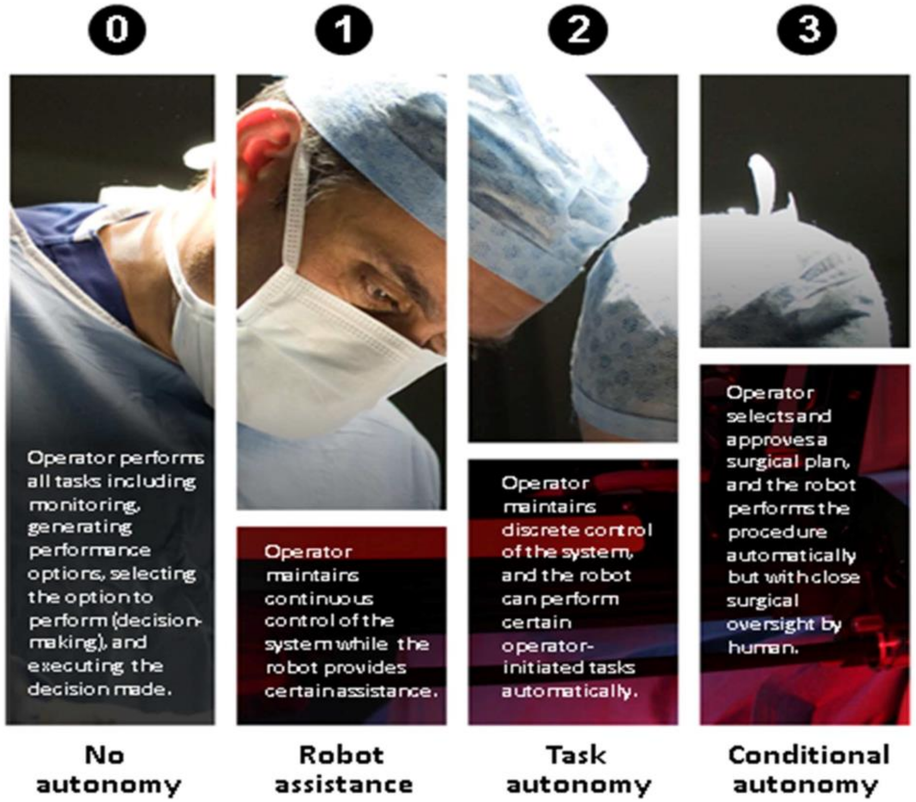
Regulations related to Autonomous systems

- Adoption of autonomous system must come with sufficient safeguards, to minimize the risk of harm these technologies may cause, such as bodily injury or other harm.
- **In the EU, product safety regulations ensure this** is the case. However, such regulations cannot completely exclude the possibility of damage resulting from the operation of these technologies
- **Surgical autonomous systems must meet all the certification criteria of human surgeons**: they practice medicine and therefore it is not just a "simple" system verification problem but much more → no actual regulations, guidelines and procedures exist.



https://ec.europa.eu/transparency/regexpert/index.cfm?do=groupDetail.groupMeetingDoc&docid=36608

Autonomous system verification: safety



For the lowest levels of autonomy (up to 3 and possibly 4), we can divide autonomous system into some key elements:

- reasoning,
- control,
- perception
- interaction between subsystems.

We target the verification of a specific property of the system: **safety**.

Safety is an emerging property of complex systems, which cannot be verified just by looking at the individual component, but it is necessary to examine the system as a whole.

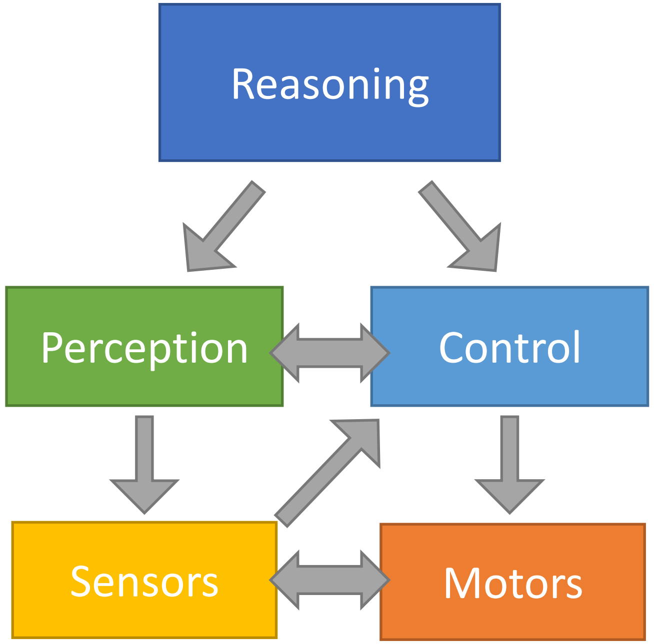
Safety verification in autonomous systems

A common approach is to verify each component of the system with exhaustive tests:

- **Reasoning**
- **Perception**

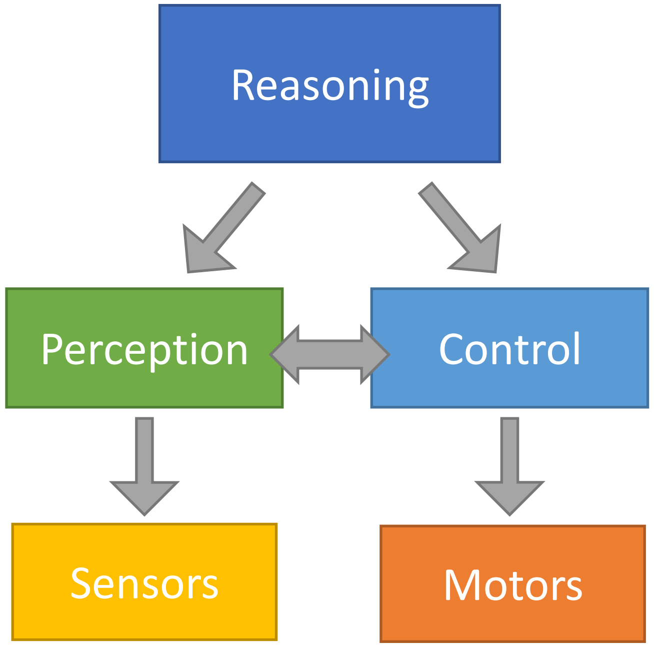
} Open problem for methods based on AI

- **Control**: this is probably the simplest problem, since exhaustive literature is available on control system verification.
- **Interconnections**: they must include all the feed forward and back connections that are necessary. <u>Very complex and still an open problem.</u>



Safety verification in autonomous systems

Following the reductionist approach mentioned above, safety of complex system has been focusing on assuring quality and reliability of the single components, assuming that when all components are working properly the complete system will be safe.



The final goal is to obtain the best overall safety of the whole system, but:

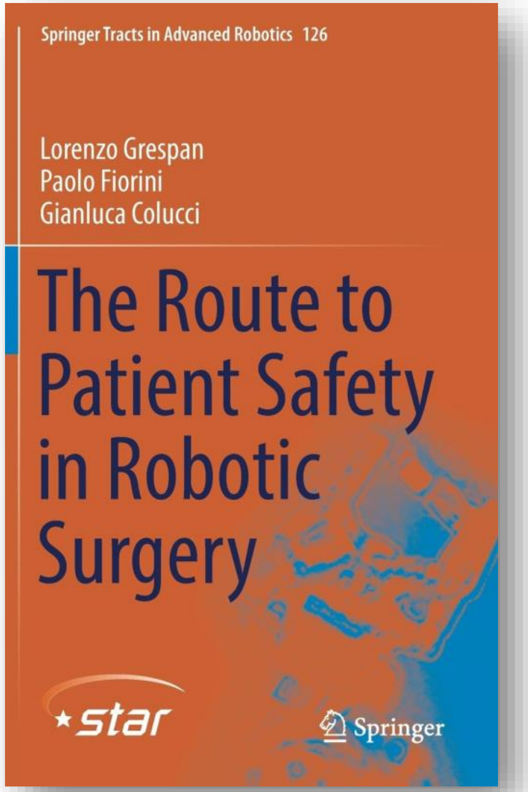
- The verification of each components could lead to the identification of issues in one of the component and then to a *local improvement* of safety
- This local improvement could lead to an overall *reduction of the global safety*

Safety verification as risk analysis

The widely enforced method for the certification of medical devices, either to obtain the European *CE* mark or the USA *FDA* approval, is based on risk analysis, whose results must be addressed to remove and/or minimize the identified risks.

The main risk analysis methods are:

- Preliminary Hazard Analysis (PHA).
- Fault Tree Analysis (FTA) → (IEC 61025)
- Failure Mode and Effects Analysis (FMEA).
- Hazard and Operabillty Study (HAZOP).
- Hazard Analysis and Critical Control Points (HACCP).



Grespan, Lorenzo, Paolo Fiorini, and Gianluca Colucci. *The route to patient safety in robotic surgery*. Cham: Springer International Publishing, 2019.

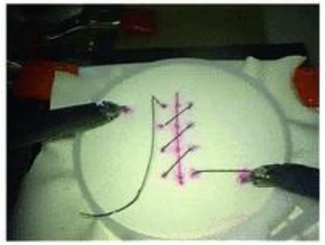
Systems-Theoretic Accident Model and Processes (STAMP)

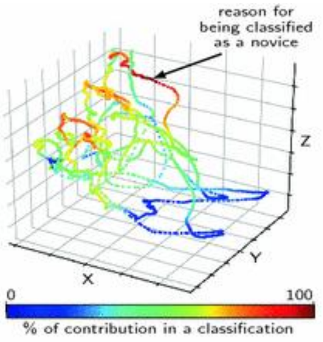
- STAMP models the concept that safety as an emergent property.
- Safety arises from the interactions of the components of a system, rather than from those individual components themselves.
- To this end, accidents are modelled as problems of control, where an accident occurs because the system controls were insufficient to constrain the behavior to a safe operating realm.
- STAMP is based primarily on three key concepts derived from systems theory:
  - safety constraints,
  - Hierarchical control structures,
  - process models

Leveson, Nancy G. *Engineering a safer world: Systems thinking applied to safety*. The MIT Press, 2016.

An alternative approach could safety by design, i.e., considering safety constraints from early stages of design.

Common problems in safety verification

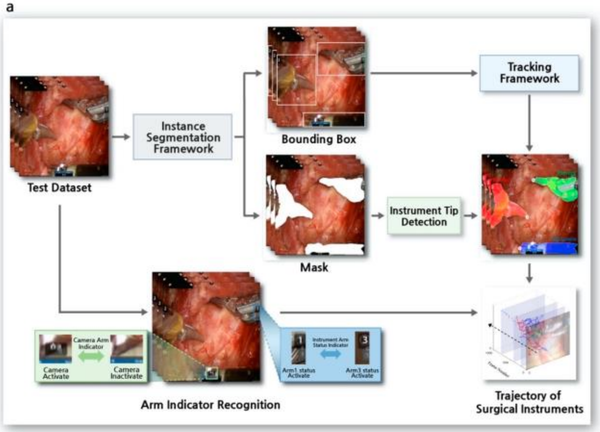




Despite the safety verification approach chosen, we need advanced testing approach since:

- We cannot manually perform all the required tests → **automatic evaluation of surgical skills**
- It is not possible to work in real environments for many reasons:
    - Ethical
    - Economical
    - Practical
    - Many others

  → **Simulation (virtual or synthetic) based evaluation**



<u>We can use all the background already developed for the evaluation of human surgeons.</u>

Surgical skills assessment: traditional approach

Traditionally, surgical skills assessment is based on the evaluation from an expert surgeon, following some standardized protocols:

- Objective Structured Assessment of Technical Skills (OSATS)
- Global Operative Assessment of Laparoscopic Skills (GOALS)
- Global Evaluative Assessment of Robotic Skills (GEARS)
- Many other protocol exists, which are more and more procedure and approach specific, for instance:
  - R-OSATS (Robot-Objective Structured Assessment of Technical Skills),
  - SARMS (Structured Assessment of Robotic Micro-surgical Skills)
  - GDCS (Generic Dedicated Scoring Criteria)
  - ARCS (Assessment of Robotic Console Skills)
- These procedure are time consuming (since an expert operator need to supervise the trials or to perform the evaluation offline based on data recordings)
- We are not limited to the assessment of technical skills, you could also evaluate cognitive or social skills (non- technical skills)

Modified objective structured assessment of technical skills (OSATS)

Gentleness Minimizing tissue injury	1 2 3 4 5	Rough, tears tissue and poor control Minor trauma with occasional breaks Appropriate tension with negligible injury
Time and motion Efficiency in movement	1 2 3 4 5	Uncertain, inefficient and lack of progress Slow, reasonable and organized Confident, efficient and fluid
Instrument handling Fluid use of instruments	1 2 3 4 5	Overshoots target, slow to correct Some overshooting, but quick to correct Accurate direction, correct plane, minimal readjustments
Flow of operation Smooth transitions between steps	1 2 3 4 5	Uncertain, constantly changing focus Slow, but planned and reasonably organized Safe, confident, maintains focus until time to move on
Tissue exposure Tissue retraction and camera visualization	1 2 3 4 5	Use of one hand and poor coordination Use of both hands, but with sub-optimal dexterity Expertly utilized both hands complementarily
Summary score Overall assessment of trainee's technical skill	1 2 3 4 5	Deficient Average Masterful

Automatic surgical skills assessment

During the surgery we can record the following information:

- **Robotic kinematic data**
  - Instrument and endoscope traveling distance
  - Motions velocity
  - Acceleration
  - Deceleration
- **System event data**
  - System controls and commands
  - Energy application
- **Surgical video Data**
  - Surgical footage annotation

The Automated Performance Metrics (APMs) derived from such data can be used for surgeon's skills assessment

Automated Performance Metrics (APMs)

Kinematic data →

- Task completion time
- Instrument travelling distance
- Instrument moving velocity
- Instrument moving curvature
- Instrument acceleration
- Instrument deceleration

Systems events data →

- Camera movement
- Master clutch use
- Third arm instrument swap
- Energy application

Instruments grip force →

Experienced robotic surgeons had lower instruments applied forces than novices

Association of APMs and Manual Assessment

Domain efficiency →

- Task completion time
- Instrument travelling distance
- Instrument moving velocity
- Camera traveling distance
- Frequency of camera adjustment

Depth perception →

- Instrument traveling distance
- Task completion time

Bimanual dexterity →

Instrument travelling distance
Ratio of dominant/non instrument use

- Excessive instrument force
- Number of instrument collisions

← **Force sensitivity**

- Instrument traveling distance

← **Movement fluidity**

Automated Performance Metrics (APMs)

Advantages

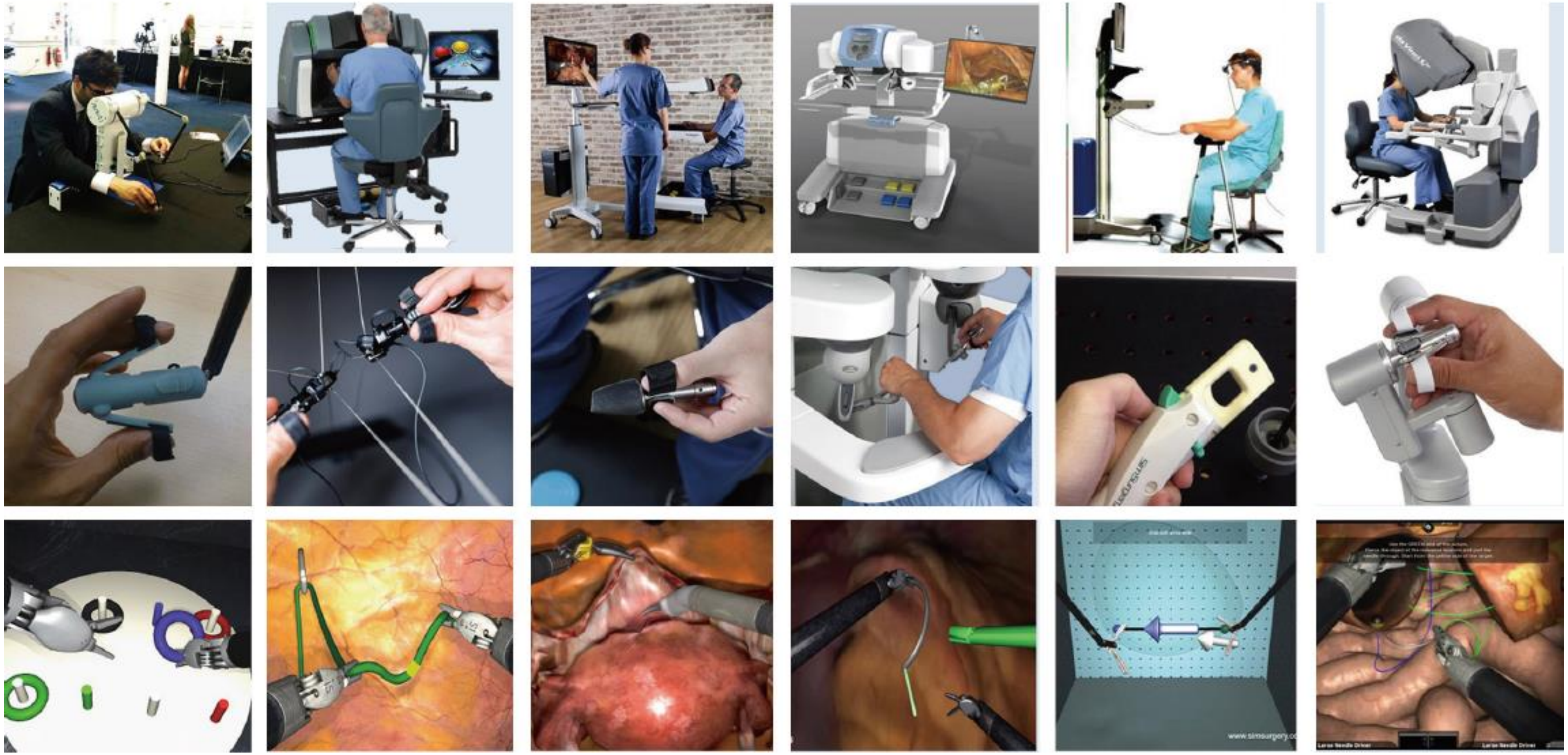
- Now possible thanks to Big Data, Machine Learning and Deep Learning
- Truly objective measures (no human judgment)
- Automatically recorded (requires no time from evaluators)

Disadvantages

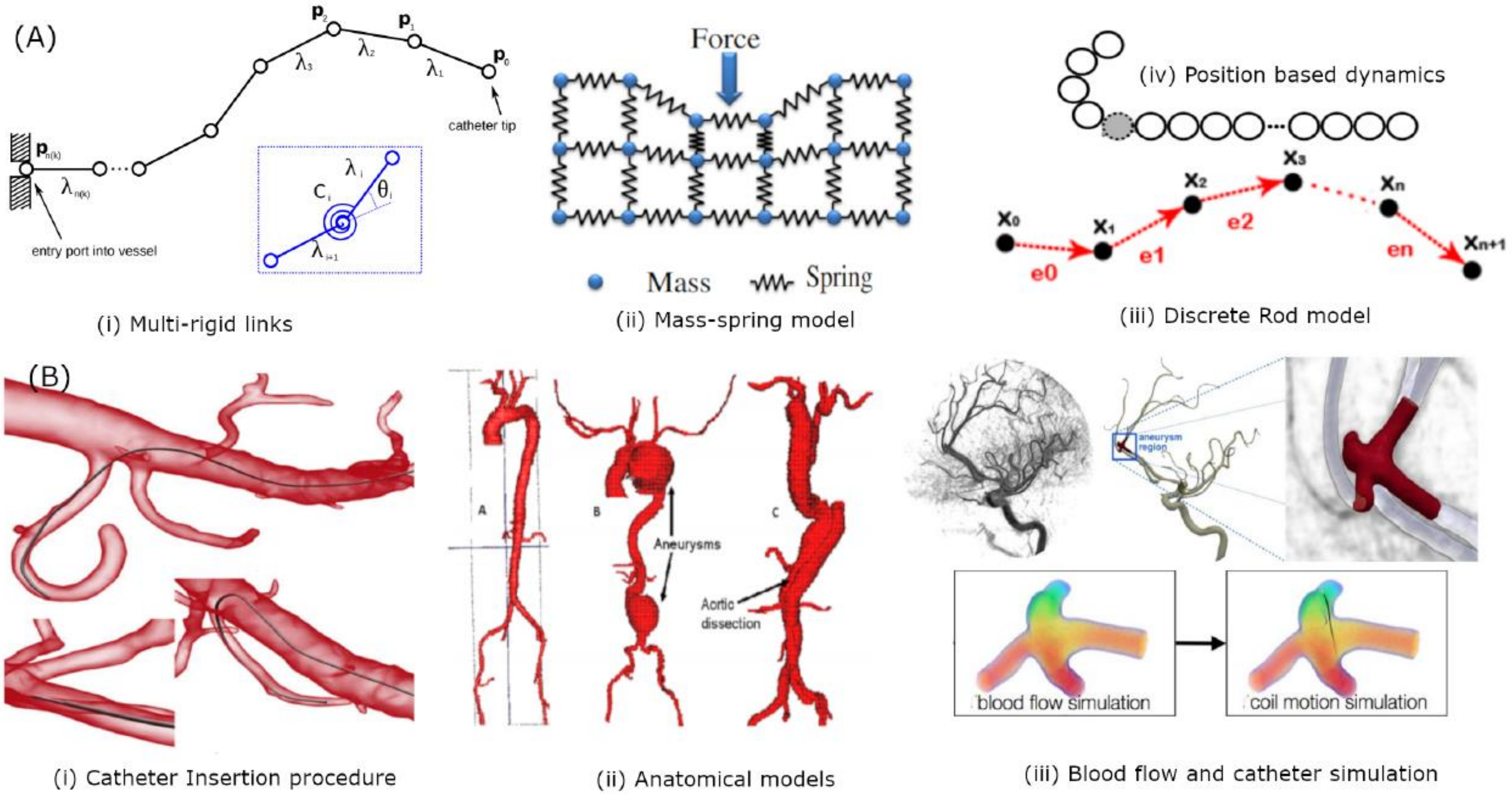
- Still in development phase
- Additional devices required to record APMs
- Large volumes of data need post processing and segmentation by algorithms
- Surgical performance measurements without surgical context will not result in meaningful surgical feedback.

Simulation based evaluation: virtual environments

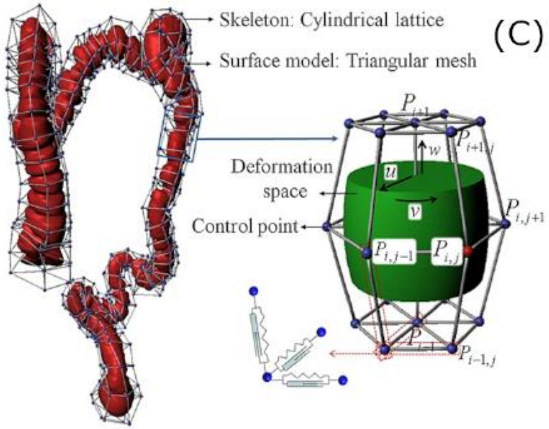
Actaeon	dV Trainer	RobotiX Mentor	ROSS	SEP robot	da Vinci Skills Simulator
BBZ	Mimic technologies	3D systems	Simulated surgical systems	SimSurgery	Intuitive surgical



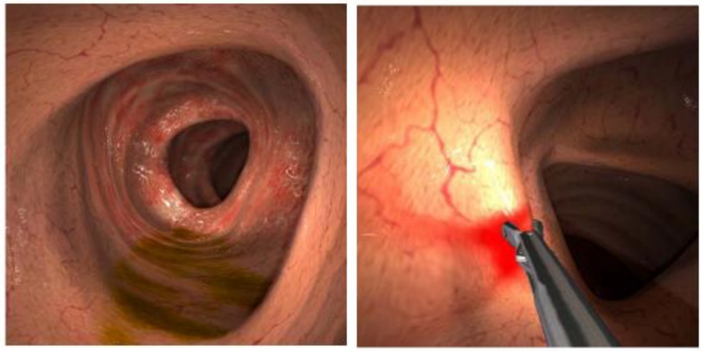
Simulation based evaluation: virtual environments



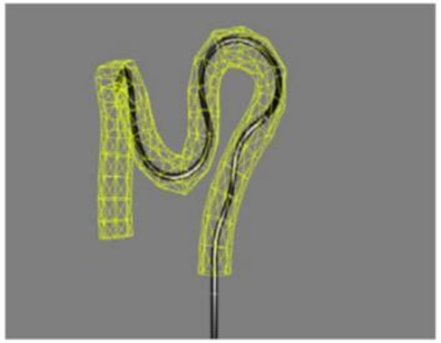
Simulation based evaluation: virtual environments



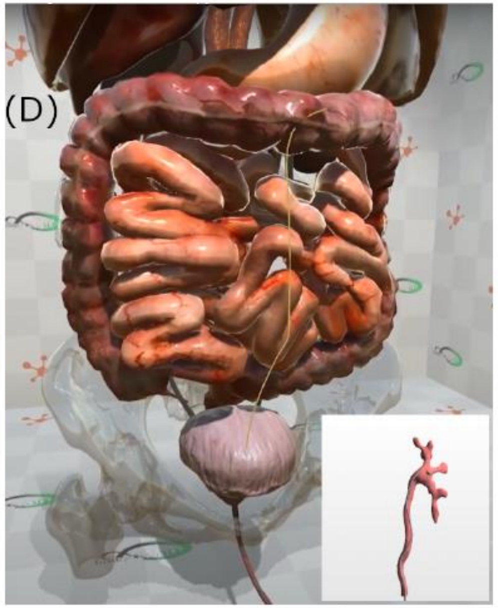
(i) Deformable colon model



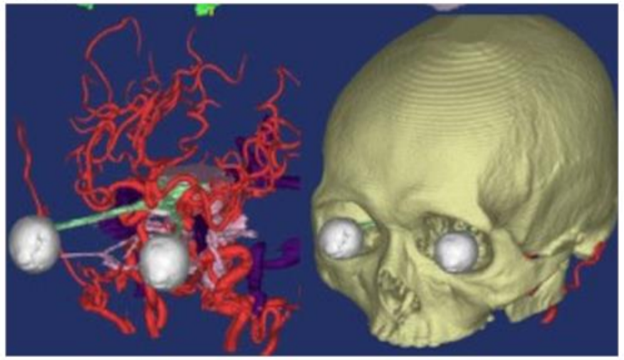
(ii) Simulated colonoscopy



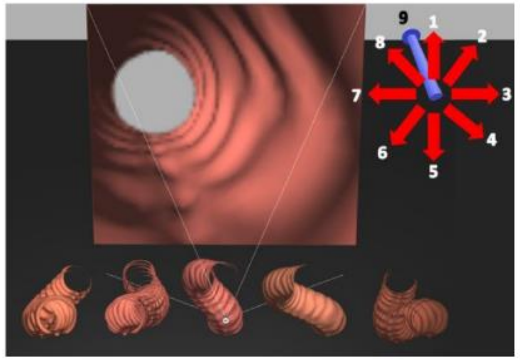
(iii) Colonoscope deformation



(E)

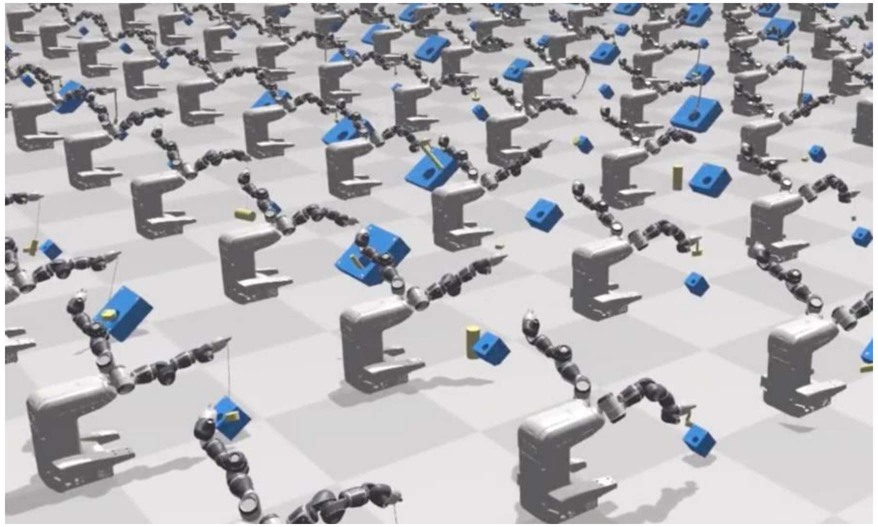


(F)



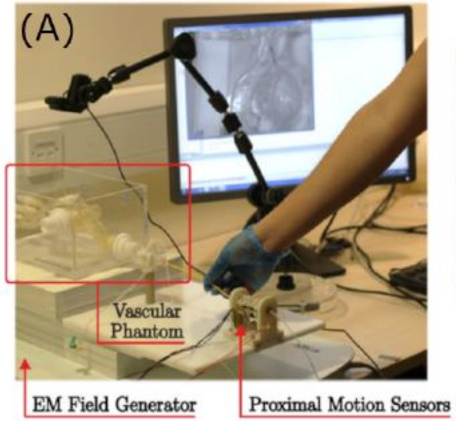
Virtual simulation main limitation: reality gap

- The problem of the limited realism of virtual simulations is known in literature with the term “reality-gap”.
- You could improve the characteristics of the simulation environment, but it is impossible to completely close this gap.
- In general, numerous methodologies have been proposed to reduce the gap, for instance:
    - Calibration procedures
    - Domain randomization
    - Deep learning approach

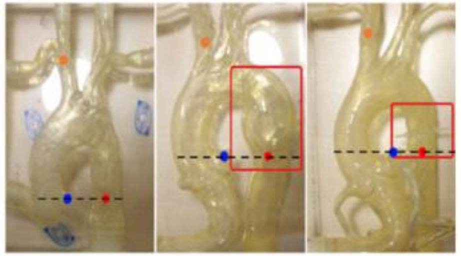


Validated Simulator are needed!

Evaluation in synthetic environments (phantom)



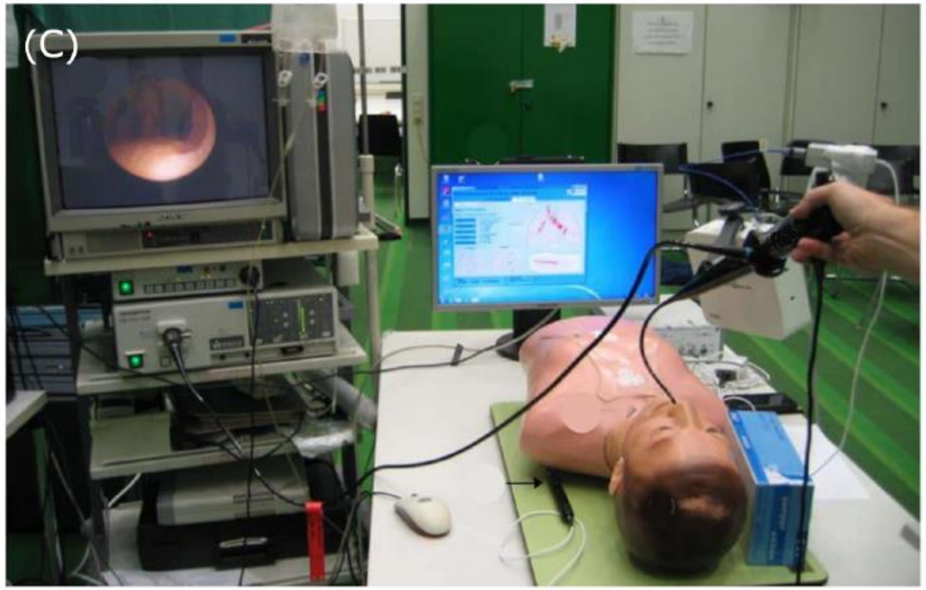
(i) Experimental setup

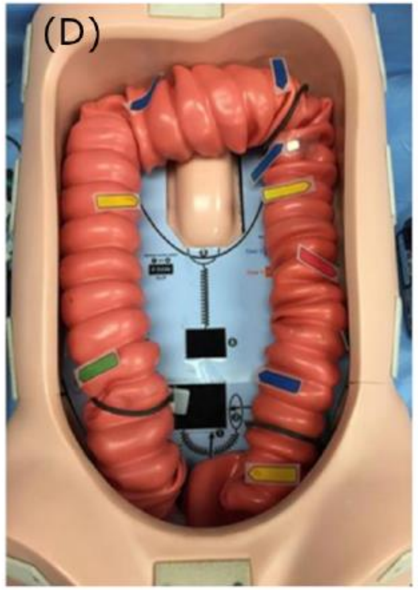


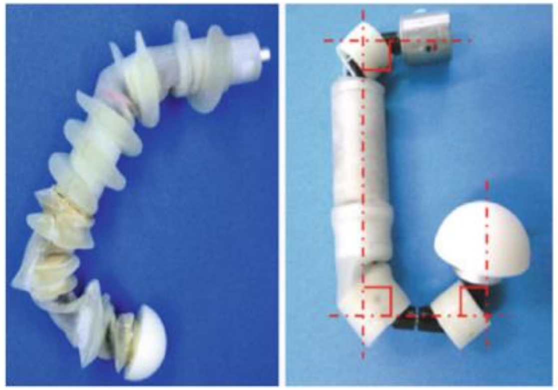
(ii) Different silicon phantom

(i) Pumper simulating blood flow

(ii) Deformation of a phantom using string







Simulation based evaluation: final considerations

- It is an essential steps before moving any autonomous system to any more complex scenario
- All simulation environments have advantages and disadvantages, it is always important to choose the best approach, <u>avoid overfitting to a specific simulation env</u>
- Learn from your errors, failed experiments are essential part of your development process.
- Verification of autonomous system is still in early stages, there is space to contribute in many aspect:
  - Validated simulation environment
  - Automatic surgical performance evaluation
  - Legal and regulatory aspects

Questions?

For more information please write to: diego.dallalba@univr.it