# Industry Needs for ML in the Real World

Lessons from deploying ML at PSI.

Edouard Servan-Schreiber, PhD
TRANSMUTEX

# Transmutex, in brief.

**2019**

FOUNDED IN GENEVA, SWITZERLAND

**40+ (28 PhDs)**

NUCLEAR PHYSICS, PARTICLE ACCELERATION, AI MODELING, INDUSTRIAL ENGINEERING

**$61M**

RAISED FROM TOP-TIER VCS & GOVERNMENTS

MEDIUM-TERM REVENUE

## HI-BEAM — Nuclear Medicine

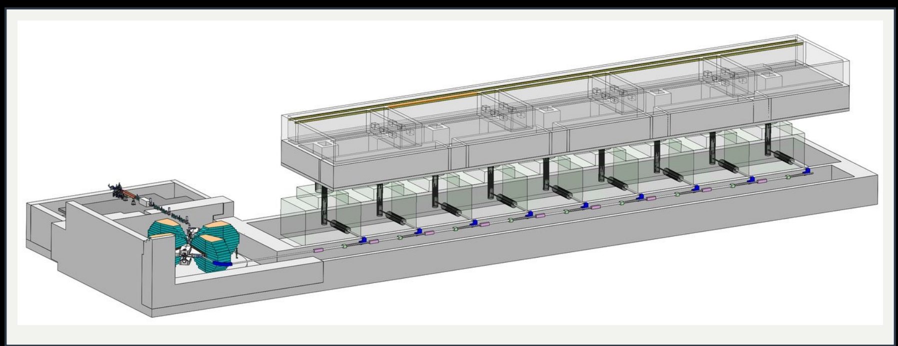
One of the world's most powerful cyclotrons (1.2 MW) producing Actinium-225 — a scarce alpha-therapy cancer isotope — at industrial scale. Up to 1.5M doses/year (≈100× today's global supply) at 5–10× lower cost.

LONG-TERM SOLUTION

## START — Energy & Waste

An accelerator-driven, subcritical lead-cooled fast reactor delivering clean baseload power while destroying long-lived nuclear waste. ≈8× less storage volume and 1,000× shorter radioactivity per unit.

Validated by the U.S. DOE (NEWTON program) and Germany's SPRIN-D · Built on proven technology with CERN, PSI, TRIUMF & CHUV.



# We build the machine, then we build the model.

Transmutex's mission is to produce isotopes for nuclear medicine — starting with Ac-225, the fuel for targeted alpha therapy: arguably the most promising cancer treatment of the past 30 years. We make it by transmuting thorium in a single-stage cyclotron.

| 200 MeV | 5 mA | Ac-225 |
|---|---|---|
| BEAM ENERGY | BEAM CURRENT | + MEDICAL ISOTOPES |

*Running it economically — getting the best performance out of every irradiation cycle — will require ML.*

PHASE 1 – NOW

**Get the hardware right**

Commissioning, I&C calibration — operational data starts accumulating.

→

PHASE 2 – ~2030

**ML earns its economic edge**

Best-performance operation during irradiation, once the data and architecture are ready.

The question this talk answers: **what architecture decisions must we make now**, so ML control becomes possible — and economically attractive — once we've accumulated operational data?

**ML-compatible I&C**

Data logging, timestamping, operator records, full instrumentation, data quality.

**The ML approach, identified**

Safety protocols tested and refined. Overall architecture proven.

**High-fidelity digital twin [KEY]**

Design support, commissioning, ML algorithm verification, suggestion testing.

# What everyone wants — and why the incentives align.

| Party | Objective | Incentive |
| --- | --- | --- |
| **Machines** | Stay intact, stay on. | Every moment off costs money. |
| **Management** | Reduce labor costs, increase accountability, reduce dependency on heroes. | Good footing for the future. |
| **Operators** | Control and reliability. | Backup to deliver reliably, and facilitate training. |
| **Business** | Uptime and throughput, now and tomorrow. | Deliver quickly and reliably, and get better with time. |

# What makes this hard — and rare.

| | |
|---|---|
| Data access | ACCESS |
| Confidentiality obligations | ACCESS |
| Sim-to-real gap | TECHNICAL |
| Integrating with — not replacing — the safety layer | TECHNICAL |
| Facility clearance | ACCESS |
| National security constraints | ACCESS |
| Building the simulation environment itself | TECHNICAL |
| Damage to the machine or the business is often irreversible | STAKES |

IMPORTANT: Accountability is <u>non-negotiable</u>, and AI/ML can’t claim it.

# “Safety” is doing four jobs at once.

### Machine safety

The accelerator, its magnets, its targets — physically undamaged.

### Operator safety

The people in the room — no exposure, no injury, no forced trust in a system they can't override.

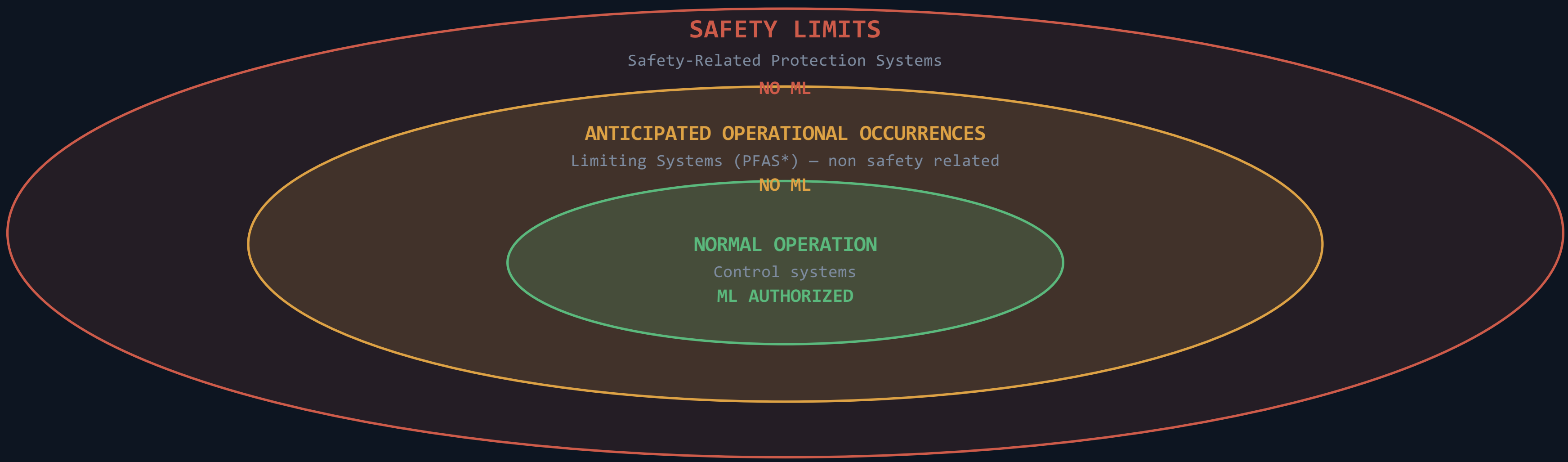
### Result safety

The isotope output meets spec — a “safe” run that produces garbage isn't safe.

### Business safety

The facility, the contract, the license to operate — none of it put at risk by an automation shortcut.

# ML is authorized in exactly one layer.

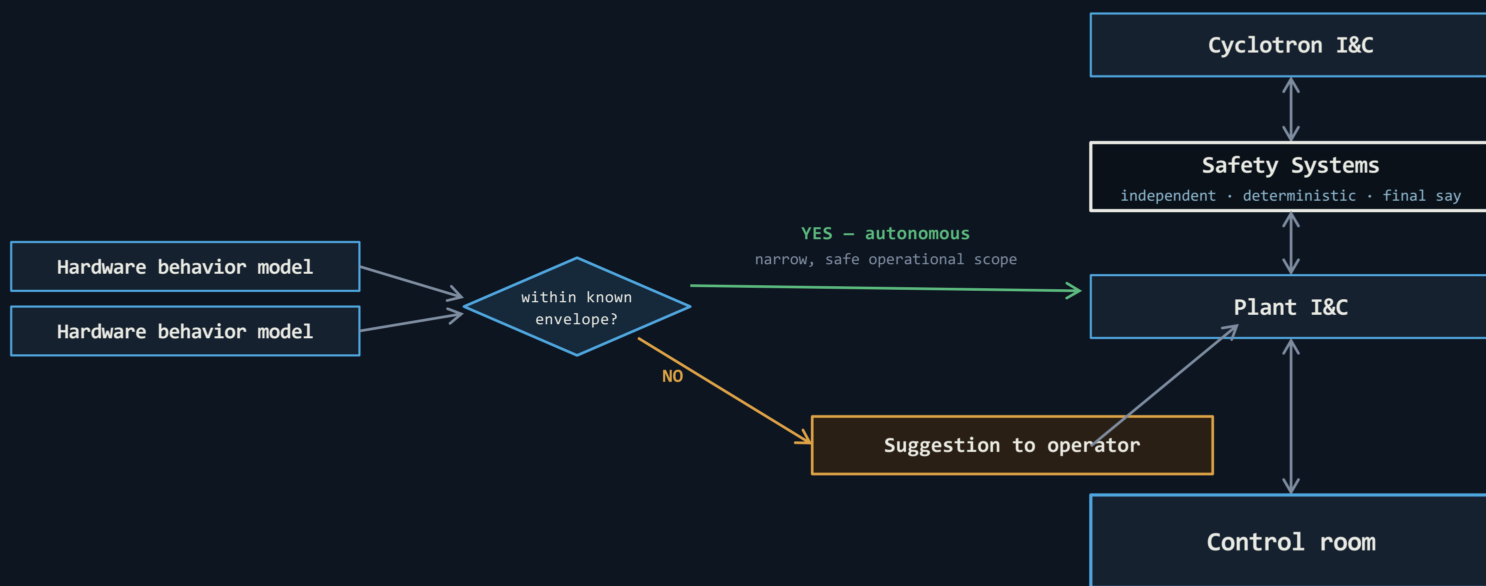



* Protection Feature Actuation Systems. Additional defense-in-depth layers not shown: severe accident and emergency planning.

07 ML WITHIN THE CONTROL SYSTEM

# ML sits inside the loop — not above it.

MPS ~ms → RL Agent ~s → Supervisory ~min → Operator (HITL) min-hr



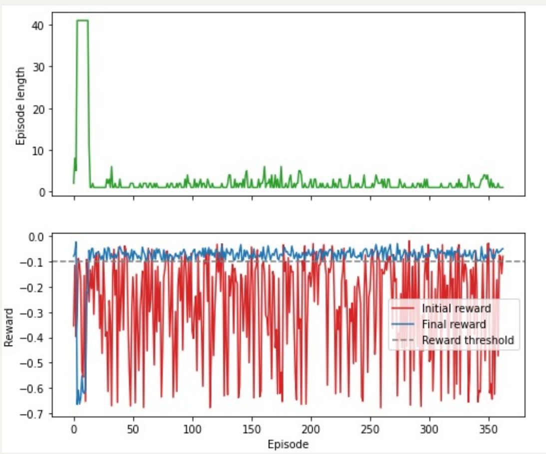
# Five ingredients

| Ingredient | What it is | The catch |
|---|---|---|
| **Historical data** | Years of logged operational behavior — the baseline the model learns from. | Often confidential; access is a negotiation, often messy and requiring lots of wrangling. |
| **Simulation environment** | A high-fidelity twin, where mistakes cost nothing. | Usually has to be built — rarely exists off the shelf. |
| **Safety enforcer** | The Machine Protection System (MPS) — untouched and authoritative. | Lives in the customer's existing software layer. You integrate with it; you don't own it. |
| **Operator** | The human whose expertise is the ground truth we're building toward. | A sensitive human — trust, blame, and job security are all live wires. Any problem is his/her problem. It’s personal. |
| **Task-complexity crescendo** | Never the hardest task first. | Bespoke per facility — not particularly repeatable. |

# Secure gradual adoption and operator confidence.

NOT ACCURACY. NOT AUTONOMY. EARNED TRUST.

# Three steps, each one de-risking the next.




RL agent trained from scratch on beam-tracking simulation — convergence achieved within 1000 iterations.

**1. Mockup functions — fast behavior testing.**

Simple analytic functions emulate system response to check stability; identify tuning, normalization, and convergence issues early.

**2. Surrogate models — data-driven realism.**

Built from historical machine logs (phases, trims, losses); realistic testing under controlled virtual beam conditions; bridges abstract simulation and hardware constraints.

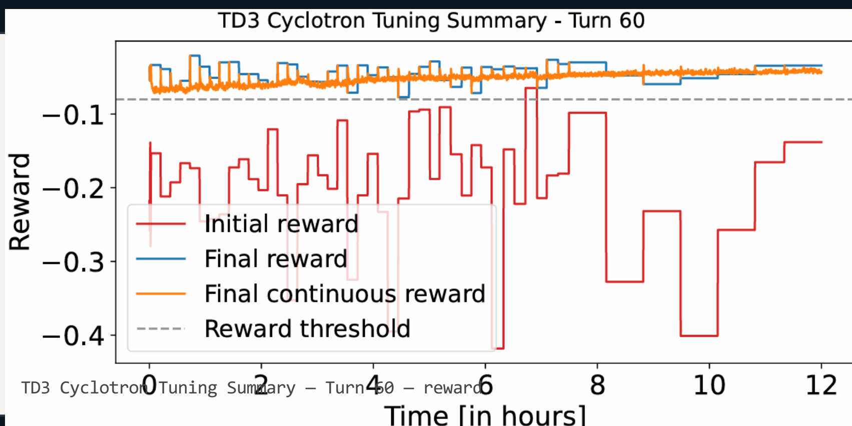
**3. High-fidelity simulations — physics validation.**

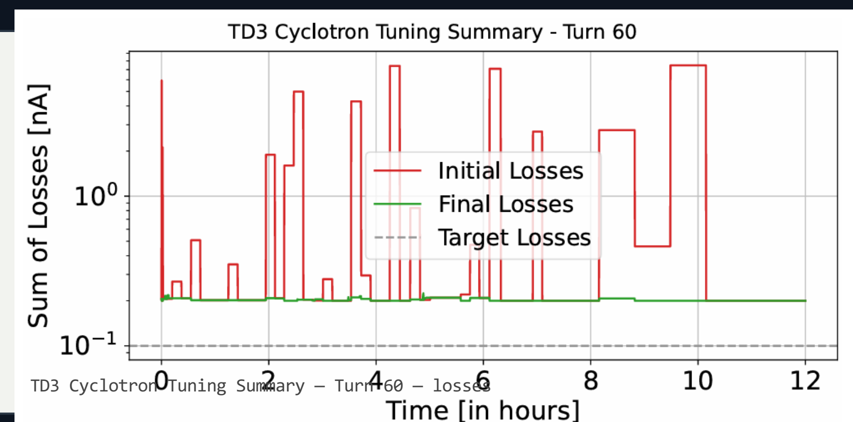
Full beam dynamics model; validates reward structure, response time, and safety limits under real physical dynamics.

This approach maximized learning efficiency and minimized risk throughout the two-week campaign.

# 12 Hours Autonomous. Zero Interlocks. Zero Interventions.

Trained during the day, launched in autonomous evaluation mode at 9pm. No operator until 9am. The agent compensated for thermal drifts, maintained beam quality, and kept losses at the same level all night.




This is not a simulation. This is the real machine, at the real facility, under real safety constraints.

# From 20 µA to 800 µA.

| Beam current | Initial losses | Final losses | Nb of steps | Nb of interlocks |
|---|---|---|---|---|
| 50 µA | 0.2 nA | 0.2 nA | 1 | 0 |
| 100 µA | 2.0 nA | 0.2 nA | 3 | 0 |
| 200 µA | 17.1 nA | 0.2 nA | 2 | 0 |
| 300 µA | 95.0 nA | 0.3 nA | 3 | 0 |
| 500 µA | 7.0 nA | 0.2 nA | 3 | 0 |
| 800 µA | 23.3 nA | 0.3 nA | 5 | 0 |

**1.5 hour** effective beam time in total — across the entire 50–800 µA sweep, zero interlocks.

- Starting from model pre-trained the day before at 20 µA
- RL consistently converged in ≤5 steps.
- Final losses always dropped to <0.3 nA (0.2 nA corresponds to absolute minimum)
- **No interlocks triggered - even at 800 µA ≡ 58 kW average beam power**

# Measured impact on Injector 2

*Results from 20+ full day deployment on PSI HIPA, ML in control*

| | Before ≡ manual operation | | After ≡ ML-assisted |
|---|---|---|---|
| **Configuration change** | **10 – 47 min**<br>Manual tuning per change | → | **< 2 min**<br>Once agent is pretrained (< 10 steps) |
| **Interlocks during setup** | **3 – 7 per change**<br>Beam loss during manual exploration | → | **0 – 3 total**<br>Turn 89 (degraded mode): zero |
| **Overnight operation & drift** | **Reactive**<br>Losses accumulate until interlock fires | → | **12h stable, proactive**<br>Inspector detects drift, prompts re-centring |
| **Operational data quality** | **Unstructured**<br>Not time-aligned, not documented | → | **Structured by design**<br>Every action + observable timestamped; reveals hidden correlations |
| **New working point** | **Full re-optimisation**<br>Manual re-tuning from scratch | → | **Immediate transfer**<br>Sim-pretrained policy, loaded and worked |
| **Knowledge capture** | **In operators' heads**<br>Retires when they do | → | **In the reward function**<br>Transferable, auditable, reusable as training sim |

Operator expertise encoded formally in the reward function → transferable, auditable, and reusable as a risk-free training simulator for the next generation

Full results phase I published, M. Haj Tahar et al, arxiv 2512.03829 (2025)

# Take Aways

| Topic | Why | What's hard |
|---|---|---|
| **Need a grounding outside of ML** | Need simulation, virtual env, digital twin, to accelerate learning without hardware, like CAD for architecture | That's where all the domain challenges show up |
| **Earn trust gradually** | This is an arranged marriage, be considerate | Define milestones of progress WITH operators |
| **Work with accountability** | No machine will ever be accountable | Help to maintain control and facilitate accountability |
| **Help to protect data** | You can obfuscate data (like the Netflix challenge) and still do ML | What is dangerous is often highly confidential and requires data protection. |
| **Align on "success"** | Figure out what REALLY matters | Not just WHAT, but WHO |

# Thank you.

TRANSMUTEX · QUESTIONS WELCOME

Backup slides available on request: MPS integration detail, simulation environment architecture.