



Operational Intelligence



Optimizing computing operations

*Federica Legger on behalf of the
operational-Intelligence@cern.ch community*



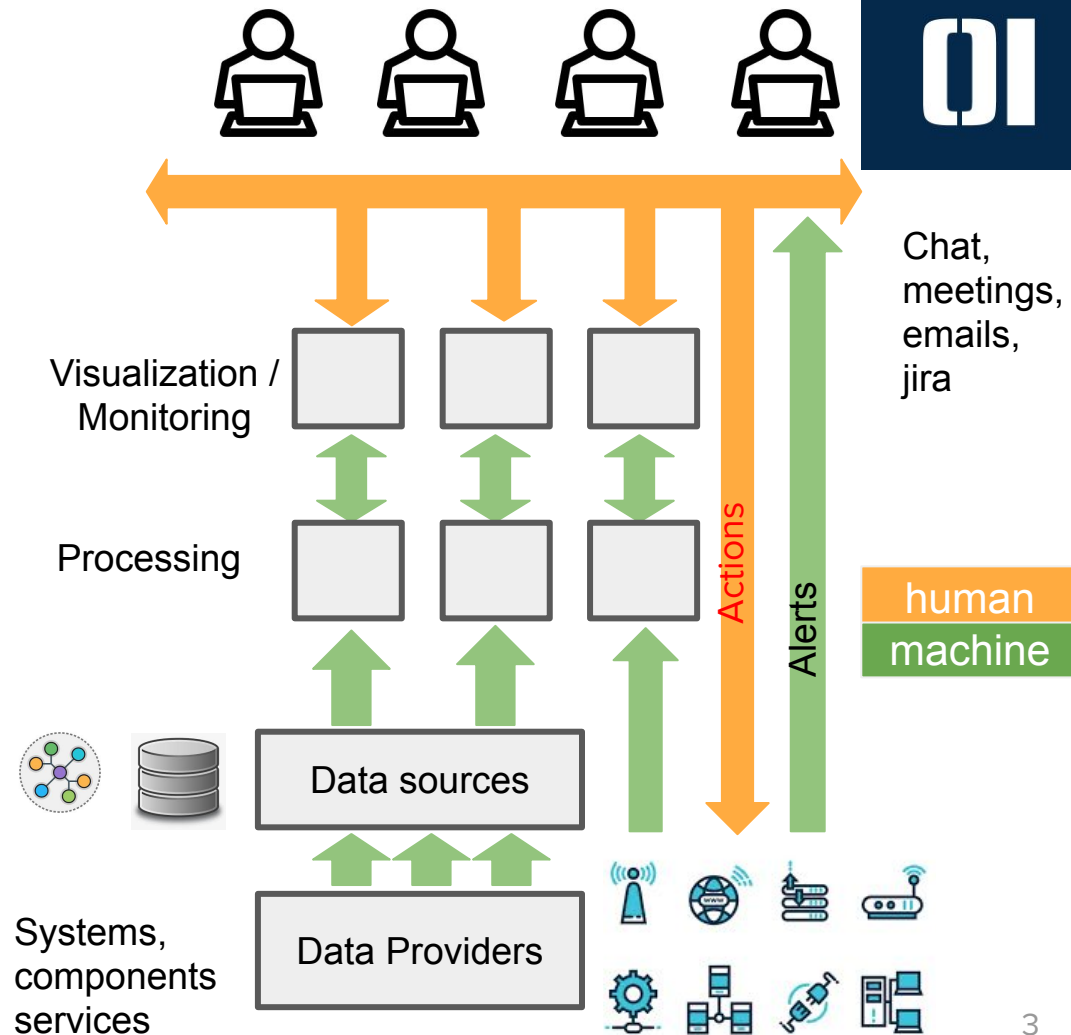
WLCG

- 200 sites
- >40 countries
- 750000 cores
- 2 million jobs/day
- 600 PB storage
- 10-100 GB links



Operations today

- Several teams responsible for own services + central team
- Multiple ways to interact (mails, meetings, chats, ...)
- Several monitoring components (many custom)
- Trivial tasks already automated (via custom scripts)
- Non-trivial tasks hard to automate since collection and correlation of information from several sources is needed
- Documentation scattered in several places+expert knowledge



The cost of operations



- As a result, operations require:
 - a lot of **human effort**
 - **learning** a lot of tools
 - Long training time for newcomers
 - Hours or days to address complex issues requiring interactions among several experts
- ATLAS/CMS report 100 people involved in computing operations (**50 FTEs/experiment**)!
 - In 1 year, > 1k GGUS tickets for ATLAS, > 2k for CMS

- Process information
- Triage
- Take decisions
- Escalate/fix
- Record actions

Can we do better?



- Standardize operation procedures:
 - Use common technology to store and access monitoring information
 - Use common visualization tools, alerts system
 - Automate common operational workflows
 - Keep flexibility to fit all/new use cases
- Add intelligence to operations:
 - LHC experiments built a successful computing ecosystem for LHC Run1-2, but at which depth do we fully “understand” it?
 - Can we perform precise modelling of the system?
 - Can we use this modelling to make predictions and/or take actions?

- A cross-experiment effort aiming to streamline computing operations:
 - Improve resource utilization by reducing the time needed to address operational issues
 - Minimize human effort for repetitive tasks by increasing the level of automation
 - Build a community of technical experts: critical mass to have impact on concrete and common issues while setting up sustainable tools.
- Our **mission**:
 - Identify common projects
 - leverage **common** tools/infrastructure
 - Collaborate, share expertise, tools & approaches
 - Across experiments
 - Across teams (operations, monitoring, developers)
 - **Bottom-up** approach

- Computing systems mature and well-understood
- Clear request from funding agencies: push on commonalities
- Many experiments use industry standard tools (ELK stack, Hadoop+Spark ecosystems)
 - ATLAS/CMS rely on common analytics infrastructure@CERN (MONIT)
 - Easier to interest students/engineers (with background different from HEP)
- More experiments starting (or considering) using LHC-developed tools
 - for example **Rucio**, and **FTS!**
- Share efforts with wider (than LHC) community

OpInt: forum where people can share and think together how to build sustainable & reusable tools, exploiting already hardly gained experiences

- Bring people onboard: successfully started Operational Intelligent activities
 - Kick-off meeting at HOW19
 - Regular bi-weekly meetings
 - Experiment-agnostic: the community is the driving force
 - Data-formats, schema, tools/layers
 - ATLAS, CMS, LHCb, HammerCloud, CERN Monit, FNAL and more...
- Work in parallel on different areas
 - Study the literature, in WLCG but not only
 - **data standardization**: achieve global schema among data (difficult)
 - (More realistic) use data as is and apply **analytics/ML** methods
 - Close the loop and provide feedback to upstream tools (for example, to improve error reporting)
 - Do not rush on implementations, but architect system(s) which can cover multiple experiments needs

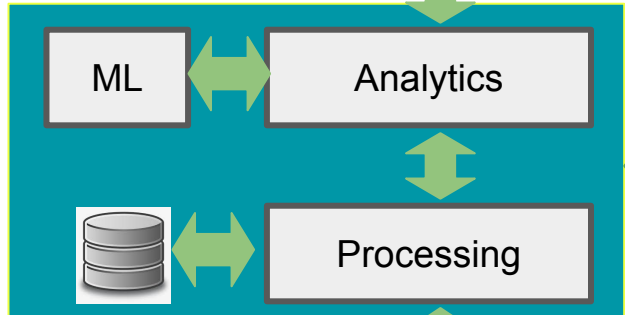
Tomorrow



Frontend: aggregated views, suggestions, collects feedback

Visualization / Monitoring

Backend: Fetches, stores, filters, and analyses information about alerts, issues and solutions

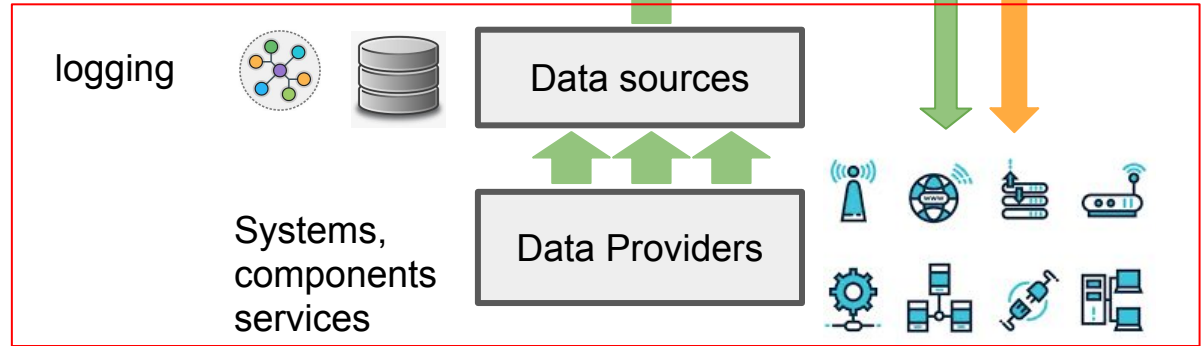


Actions/alerts

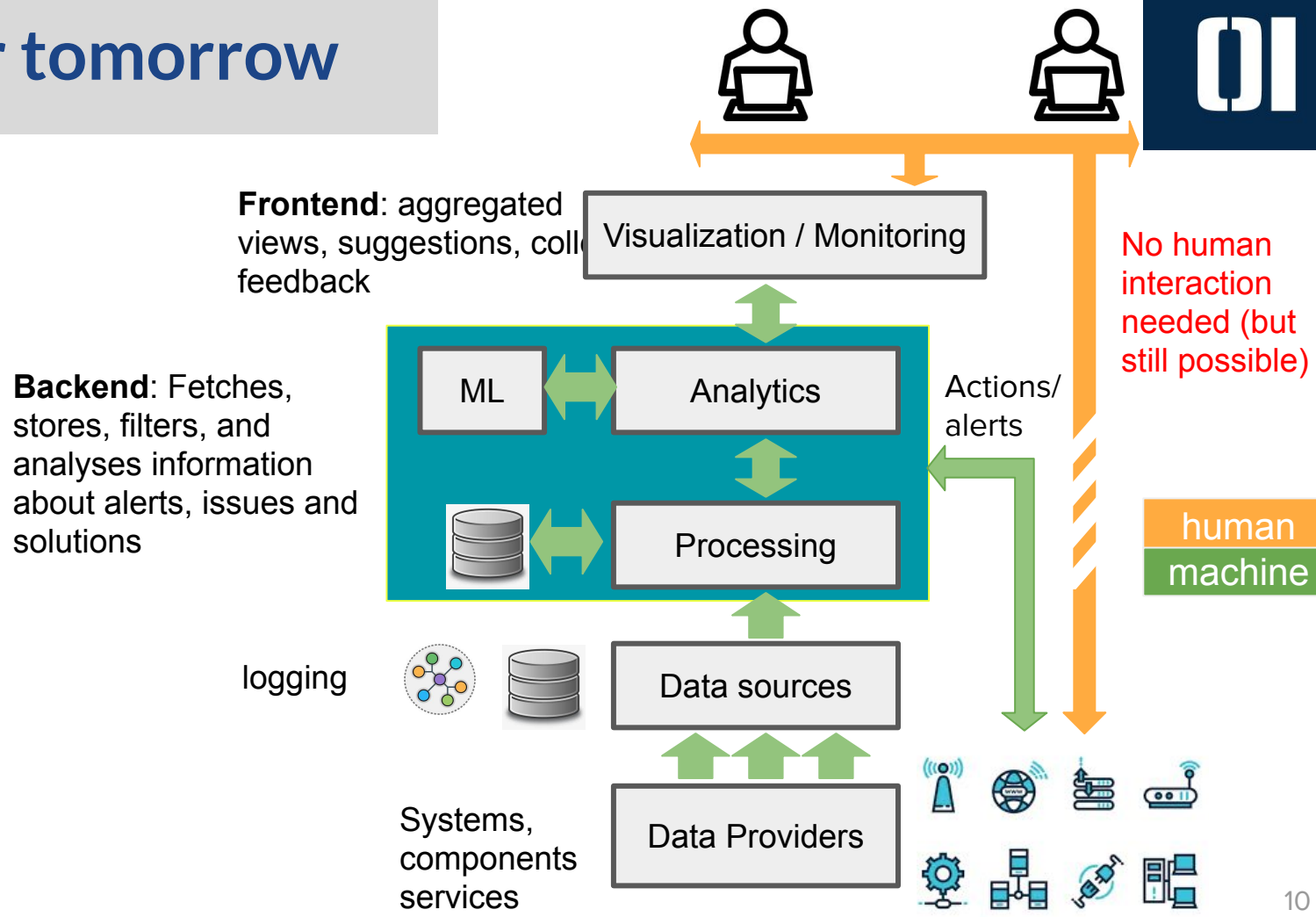
Actions

human
machine

We already have this →



Day after tomorrow



- **Anomaly detection** in time series (Data quality, Network issues, Site performance)
 - Despite importance, not many off-the-shelf tools
- **Classification and clusterization** (Data popularity, error classification, FTS, Rucio, Frontier)
 - Logs - *almost free style text* - NLP?
- Lack of
 - well annotated data
 - common data formats/flow
- Deploy in **production!**
- Foster the Op Int **community**
 - For now based on **best effort**, need people with dedicated time



- Infrastructure/tools:
 - CERN, FNAL, INFN-CNAF et al. Analytics Infrastructure, ATLAS ML Chicago platform
 - MLaaS for HEP
 - Visual Analytics
 - Log analysis
 - **Oplnt Framework**
- Projects:
 - Jobs: CMS Operator Console and Alert Triage, **Jobs Buster** in Atlas
 - Data management:
 - ATLAS, CMS, and LHCb Data Popularity
 - **Rucio data transfers**
- Sites:
 - FNAL black node detection and **Google AutoML Tables**
 - INFN Predictive Site Maintenance and Site operations anomaly detection

Quick recap

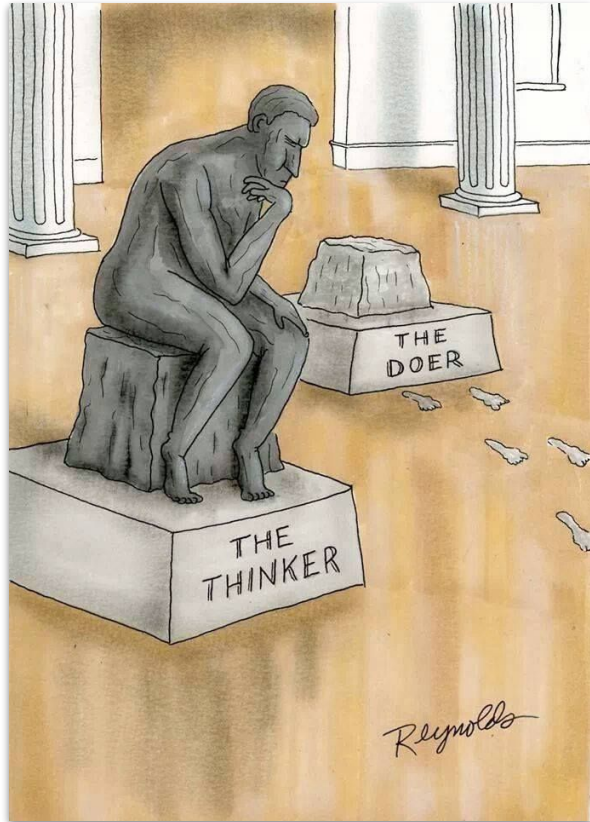


- Oplnt [website](#)
- Github repository: <https://github.com/operationalintelligence>
- Dockerhub: <https://hub.docker.com/orgs/operationalintelligence/repositories>
- E-group for communication:
 - operational-intelligence@cern.ch
 - operational-intelligence-framework@cern.ch
 - operational-intelligence-nlp@cern.ch
- Regular meetings (2x/month): **3-4 pm on Mondays**
- Framework meetings: **Fridays at 2pm**
 - Indico category: <https://indico.cern.ch/category/11205/>



Today

Tomorrow



Thanks for Listening

