IT & AI Orchestration

"Gravity" Release V13.0.0
Release notes

valid as of September 6th, 2022

Presented by

World leader in SOAP solutions
(Recognized by Gartner in SOAP 2021 Market Guide
Recognized by Gartner in Data Science Engineering platform 2022 Market Guide)

Table of Contents

1 Introduction

ActiveEon is constantly adding new capabilities, so you can leverage the latest technologies to experiment and innovate more quickly. This document provides an overview of the new ActiveEon features, expansions, and hot fixes as released in the latest release.

Browse the page below to learn about our latest innovations.

2 ProActive New Features

2.1 Proactive Workflow and Scheduling

ProActive installation within Kubernetes cluster	The ProActive instance can be easily deployed within a Kubernetes cluster environment.
	Administrators can run ProActive instances in containers within a Kubernetes cluster and apply different ProActive configuration parameters such as different databases, users, and passwords.
	This function enables the customer to take full advantage of Kubernetes capabilities in terms of high availability, disaster recovery, and full portability across platforms.
A single ProActive instance can serve multiple groups of users	Administrators can now create multi-tenants within the same instance of ProActive making the solution more convenient, future-proof, and scalable for large projects across the company.
	Each tenant can leverage the available computing resources within the organization while keeping data and workloads completely separate from other teams.
	Better usage of resources: By sharing machines among multiple users and using the same infrastructure, the organization can optimize available computing resources. **Lower costs:** Thanks to the centralized nature of the orchestrator, each tenant can benefit from centrally managed infrastructure and software, allowing the IT department to offer its services at a much lower price due to lower operational costs. **Streamlining release:** Instead of installing new versions separately on each teams' servers, the multi-tenant package only needs to be installed on a single server.
Significant increase in task processing	Our new capability enables customers to deliver automation much faster.
	The R&D team has delivered an improved version of the workload scheduler, doubling the number of tasks processed by seconds.
	By upgrading to the new release, your ProActive solution will now process the workload twice as fast as before and can process more than 2 million tasks per day.

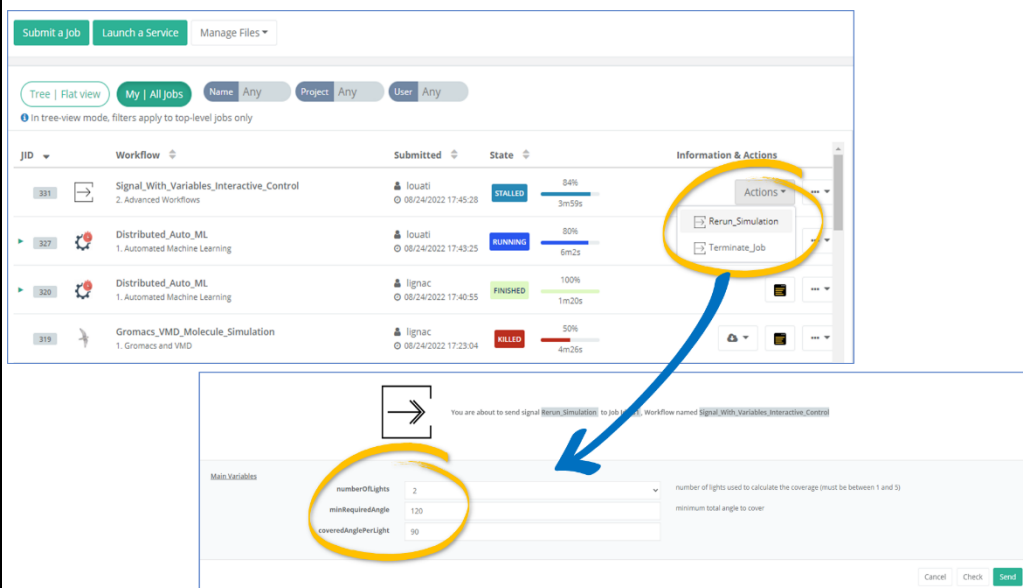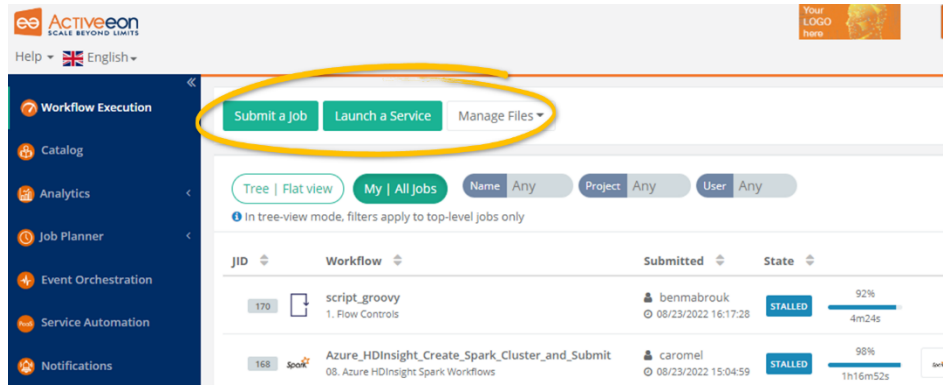| Run dynamic workflows that adjust their behavior based on parameter values	ActiveEon provides a great amount of flexibility with respect to how you want to trigger your workflow. With ActiveEon's new "Action/Signal" function, administrators can dynamically overwrite parameter values for a given run without having to redeploy their workflow.
	It's often convenient to have a workflow that is capable of handling or responding to different inputs. For example, a workflow might represent a series of steps that could be repeated for information coming from different APIs, databases, or credentials — all of which reuse the same processing logic. Alternatively, you might want to use an input parameter that we call "Action/Signal" to affect the workflow processing itself.
	The Action/Signal function is a powerful way to act upon a workflow in real-time during its process. Some examples of use cases are: ● An agent can pause the workflows and change the input parameters in real-time. (Changing the compute target, for example) ● A third-party software using REST can alter the workflow inputs during the execution of the workflow. ● The workflow can pause and request an agent to select different options: such as where to archive the data: 1- in your favorite web service 2- in the on-premises storage for confidentiality requirements Below is a screenshot where users can select to re-run the ML model by changing some parameters
	Changing input dynamically for workflow parameters enables the process to be more flexible and robust, providing a more refined user experience.

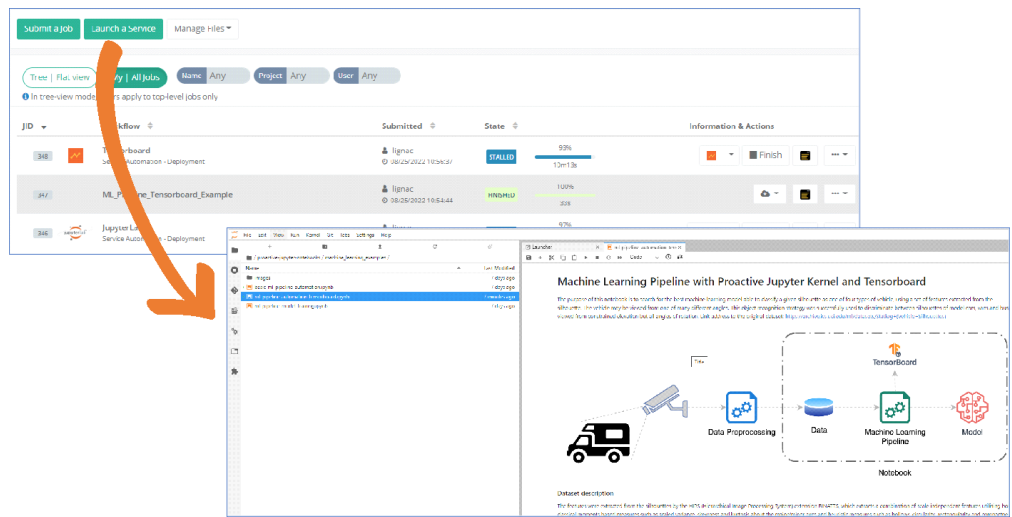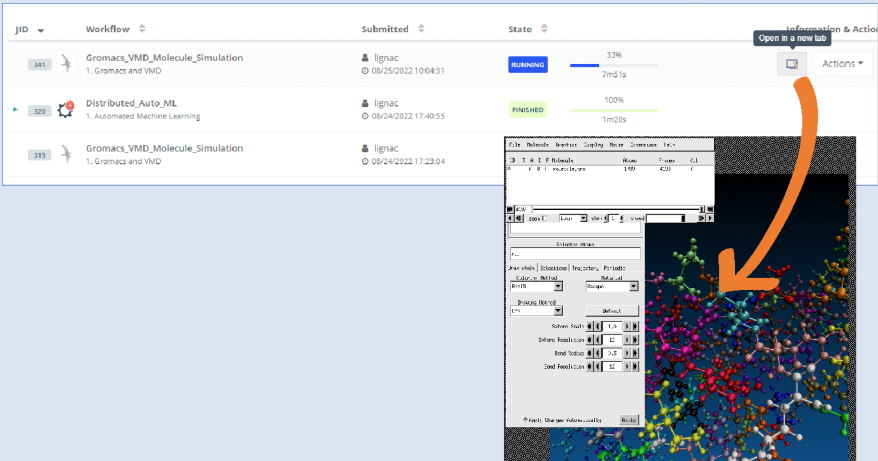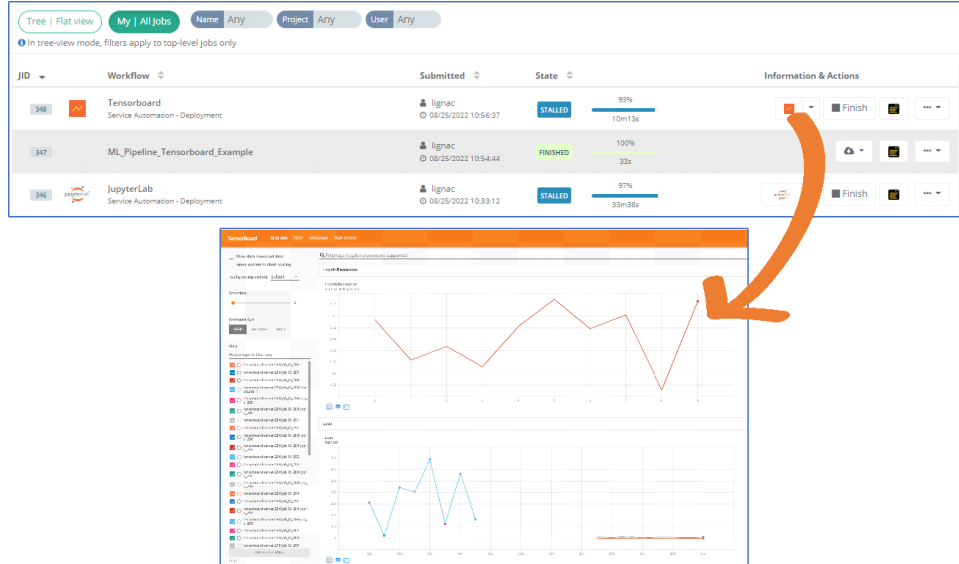| Advanced functions for monitoring workloads	Administrators will benefit from a full set of new features added to the Automation Dashboard – please see below the top five functions
Feature 1	Launching third-party services directly from the dashboard: Users can launch third-party software directly from the ActiveEon dashboard. This is particularly useful when users need information from another application and access it at a press of a button. Two customers use cases: - An Administrator/Monitoring person could dynamically open the company's centralized database describing the input parameters for each workflow template and launch a new workflow with the relevant input parameters. - Data Scientists may want to open their notebooks in Jupyter Lab and directly launch the model from Jupyter Lab into ActiveEon (using Pragma which will be described later in this document). We will discuss later on how Data scientists who have launched their experimentations can also dynamically launch Tensor Board without leaving the ActiveEon interface to visualize the progress of the model training or experiment. The service life cycle is entirely controllable by the administrator, like any other job. They can pause, run, or stop the service directly from the dashboard.

	This feature demonstrates the power of ActiveEon's open architecture, where any service can be launched from ProActive in real-time and reciprocally third-party solutions can call any Proactive service through REST API.
Feature 2	<u>Launching a third-party software during a workflow execution</u> Very similar to feature 1 but even more powerful. Administrators/Users can launch a third-party solution in real-time during the execution of a workflow. The newly launched application will be launched in parallel with the running workflow.
	Use case 1: - Engineers who launch very long processes on HPC for complex simulations could visualize in real time the output during the execution of the process. The diagram below shows a bioscience simulation using Gromacs. - Users can access Gromacs in real time to visualize the simulation as it happens

	Use case 2: - Data Scientist**s** or AI Engineers launching long ML model training or ML model optimization could visualize the early output in real-time during the execution process using for example TensorBoard. They could then decide to **terminate** the process or alternatively use the new "Action/signal" feature to change the parameters for the model input.
Feature 3	<u>Data Spaces advanced search and filtering</u> ActiveEon's Data Space is a secure and powerful way to share information between teams during the different stages of a pipeline. Because of the success of this feature, we further improved it by adding powerful search and filtering capabilities.
	This new function corresponds to the trend by our customers to use the platform to automate more sophisticated workloads, specifically for hybrid and AI pipelines, and requires many teams to collaborate.

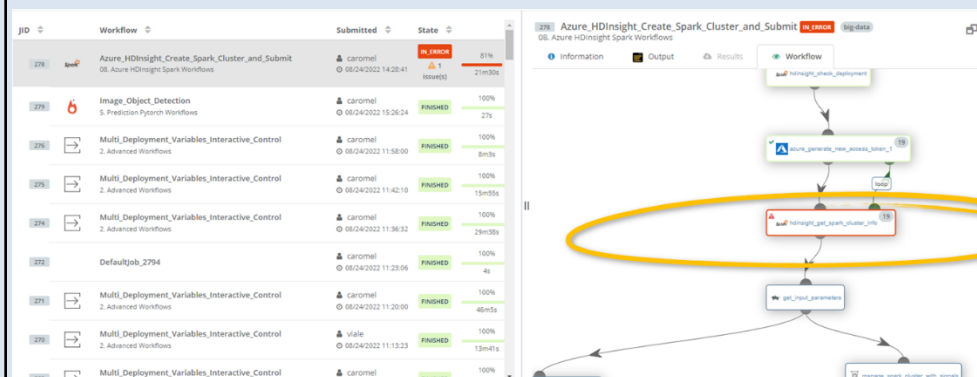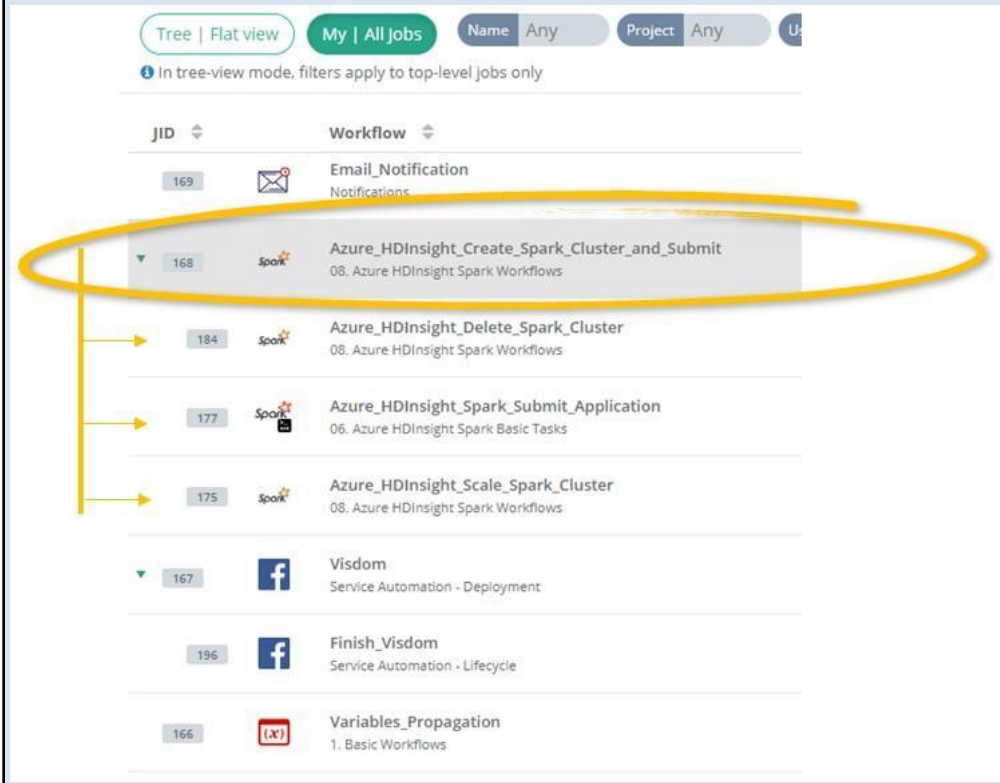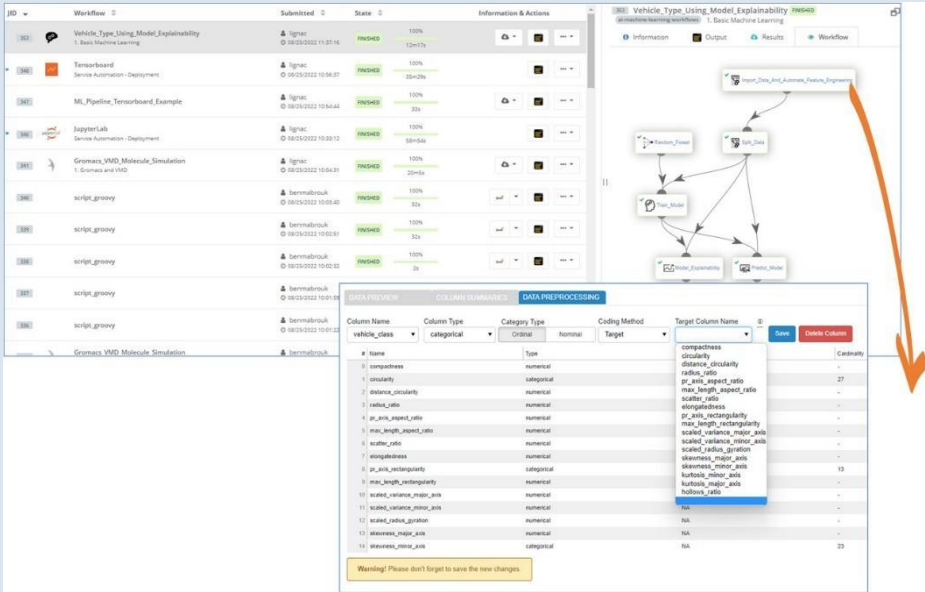| Feature 4	Advanced improved user interface for job monitoring
	Information tab - Users can see information about jobs execution status as well as the job variable (input parameters) used when the workflow has been launched.
	Example of parameters for provisioning Spark services through Azure HDInsight
	The latter becomes even more important since administrators can allow specific workflow parameters to be changed dynamically.
	As described in feature 2, users could quickly glance at the original input parameters and change them accordingly if needed.
Feature 5	Workflow tab
	Real-time visualization of workflow execution
	Workflow visualization helps improve business operations. In fact, there is no better way to troubleshoot and improve than by identifying where the process went wrong.
	As ActiveEon simplifies the process of measuring progress and success, it becomes easier for teams to spot issues and opportunities for improvement.
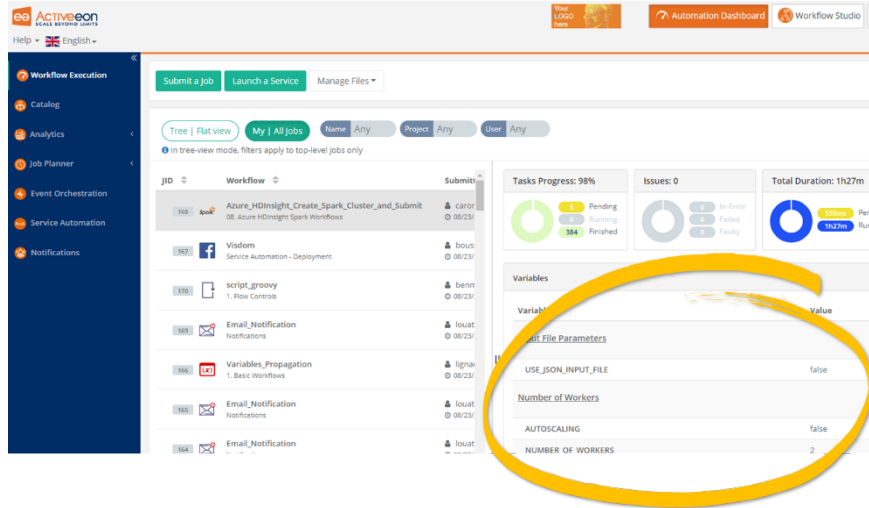
	ActiveEon displays the workflow which is being executed and graphically shows the progress – the key benefit for operations is being able to pinpoint immediately where the job has failed or paused.
Feature 6	Pipeline monitoring
	Pipelines are at the top level of the workflow hierarchy. An IT procurement pipeline might contain several workflows (provisioning a VM, allocating an IP address…) that in turn include sequential tasks (connecting to the server, creating VM…).
	ActiveEon's customers have been using pipeline management for many years, however, with the growth of advanced automation like hybrid and AI management, we decided to initiate a new advanced monitoring project and, the new "pipeline monitoring" project represents the first feature of this ambitious strategic plan.
	We are devoting serious efforts to offer our customers a growing number of features that will enable complete end-to-end visibility of pipelines – wherever the process runs and whatever applications are used. In the above screenshot the pipeline 168 – "Azur_HDInsight_Create_Spark_Cluster…" is composed of three workflows 184, 177 and 175. Administrators can monitor the overall pipeline as well as each individual workflow.
Advanced Role-Based-Access-Control (RBAC) security functions	Administrators can further secure the ProActive solution with significant new security features.
	The new advanced RBAC feature enables:

	- Administrators can limit the access of folders to specific groups or specific individuals. - Administrators can set up multiple groups of users and set access right to their group folder. - Administrators can restrict the use of a "pool" of resources to specific users or groups of users.
Additional pre-built connectors for major players	Without any coding, administrators can create sophisticated automation with the following applications: - ServiceNow (ticketing and business workflows), - Informatica, - PeopleSoft, - VMware, - Azure HDInsight, - Apache HBase, - Apache Phoenix - FTPS
	ActiveEon is continually developing more connectors. Do not hesitate to contact us for the latest list.
Improved Workflow input **parameters** (Variables) management	A workflow parameter is a value that you define before the workflow runs. It is used to set values for tasks in the workflow or to set some user-defined mapping parameters. They also reduce the overhead of creating multiple workflows when you need to change certain attributes of a workflow. This new management function will help customers who are developing sophisticated workloads that require many parameters. This is often the case for Big Data, Hybrid and AI pipelines. This new capability will help administrators to better structure the parameters for quicker and more efficient management. - Administrators can now group parameters per topics - Administrators can tag them - Administrators can also hide them if they are important for the background execution but not important for the user. This allows to provide a cleaner user interface when users launch their workflows.
	This new function corresponds to the trend by our customers to use the platform to automate more sophisticated workloads, specifically for hybrid and AI pipelines and allowing "citizen users" accessing the platform to launch their own workflows

2.2 ProActive AI Orchestration

ProActive AI new "feature engineering" Module	Data Scientists can now use the "AutoFeat" module during the Feature Engineering process. Feature Engineering is a very important step in machine learning, and it refers to the process of extracting relevant features from the data to train ML algorithms.
	One of the key features of the ProActive AutoFeat is how the algorithm can semi-automatically identify the best method for encoding each categorical column of data, with validation from a Data Scientist during the process. Alternatively, Data Scientists can select the data encoding methods they prefer: Label, OneHot, Dummy, Binary, Base N, Hash, and Target.
	AutoFeat is used right at the beginning of the model training – AutoFeat will help Data Scientists in this experimental phase by allowing them to select a relevant encoding method and by creating data transformation.
	Once the data frame is finalized, it can be sent to either a feature store solution such as: Feast, Hopsworks or used directly to train an ML model.
ProActive – "Model as a Service" can be used over multiple models	AI engineers can now deploy multiple models at the same time, as well as deploy the same model with different versions.
	This version will enable enterprises to manage registry and serving for multiple models – and or multiple versions of the same model.
ProActive Dashboard for iterative model deployment has been improved for model comparisons	Once the model is deployed – and needs to be reviewed due, for example, to drifting - an AI engineer is now able to compare the baseline data (metadata) and analyze if the new model is better or worse than the previous deployed one.

	The dashboard is able to display diagrams (plots) comparing the several versions of the same model.
More powerful use of JupyterLab	As seen above, a Data scientist can easily launch an Instance of JupyterLab directly from ActiveEon. Our latest development provides even more power to Data Scientists by using pragmas (small directives specified on the notebook cells) directly with the notebook code. In other words, Data Scientists running JupyterLab on their laptops can take advantage of the entire available enterprise hybrid compute resources by simply using pragmas and be able to train their models in very powerful machines without lifting a finger. Example of JupyterLab coding with #% pragma integrated within Jupyter Notebook:
	We have developed a list of extensive pragma types so that Data Scientists can take control of the way their models are launched, data visualized, process stopped and so on. The key element is that Data Scientists can access the entire enterprise compute resources or, at a push of a button, provision new ones. This allows to abstract the whole complexity and eases the automatic resource provisioning.
New Dashboard for model monitoring	Along with the extensive set of features, we have already developed around model deployment, model serving, and model drift, we have just launched our first ModelOps dashboard.

The dashboards enable to manage audit and traceability, Analysis of dataset, Prediction review, as well as drift analysis.

It is the first step towards answering customers who are starting to deploy 10s or 100s of models with our framework - and it will be followed very soon by some nifty ModelOps reporting tools – stay tuned.

3 List of new features, improvements, and hotfixes

3.1 ProActive Workflow And Scheduling

3.1.1 General

- *[New Feature]* Kubernetes support: ActiveEon provides different Kubernetes configurations to run ProActive (with different databases, users, passwords, etc.) in your Kubernetes cluster. Overall, offering an effortless deployment of ProActive in a fully containerized environment using Kubernetes.
- *[New Feature]* Allow defining user tenants, which represent organization units.

3.1.2 Scheduling

- *[Improvement]* Allow the usage of workflow scripts stored in ProActive Catalog's group-restricted buckets.
- *[Improvement]* Allow selective housekeeping of jobs that are terminated with errors.
- *[Improvement]* Added the possibility to submit multiple jobs from URLs
- *[New Feature]* Added the capability to have variables and values when sending a signal to a job
- *[New Feature]* Allow filtering jobs according to user tenants.
- *[Improvement]* Improved scheduler loop performance
- *[Improvement]* Update SSH libraries used by ProActive (for the communication between the server and nodes) in order to support recent SSH encryption algorithms.
- *[Improvement]* PNPS protocol now uses TLS 1.2 by default
- *[Improvement]* Updated embedded JRE version to java8u291

3.1.3 Workflow Variables

- *[New Feature]*   Workflow variables now support properties description, group, advanced and hidden.
- *[Improvement]* Added optional parameters (kind and contentType filters) for the workflow variable (PA:CATALOG_OBJECT model).
- *[Improvement]* Users can now create a pipeline by simply browsing the ProActive Catalog of workflows and select the linked workflow variables (PA:CATALOG_OBJECT) during the workflow submission.
- *[Improvement]* Workflow variables containing credentials (i.e., the type PA:CREDENTIAL) can now be more easily added or edited in the ProActive credentials management popup window.
- *[New Feature]* Added global variables which are job variables that are configured inside ProActive server and apply to all workflows or to certain categories of workflows (e.g. workflows with a given name).

3.1.4 Workflow Execution

- *[New Feature]* Support the launch of third-party software (PSA services actions) workflows as sub-jobs of the master third party software (PSA deployment) workflow.
- *[New Feature]* Launch third party software (PSA workflows) from Workflow Execution portal
- *[New Feature]* Added the capability to have variables and values when sending a signal to a job
- *[Improvement]* Added filtering capabilities to User/Global dataspace file browsers
- *[New Feature]* Display advanced, hidden, group, models, and the description of variables on workflows
- *[Improvement]* Display generic information in job details
- *[Improvement]* Added a *Launch Service* button to allow starting and monitoring a service
- *[Improvement]* Added a *Finish* button to terminate a service
- *[Improvement]* Improved RBAC enforcement (unauthorized actions are disabled)
- *[Improvement]* Improved Global/User space browser performance

3.1.5 Service Automation

- *[Improvement]* Added a *Refresh* button to retrieve the latest version of services

3.1.6 Scheduling Portal

- *[Improvement]* *RBAC:* all unauthorized actions are now disabled (menus, sub-menus, tabs)
- *[New Feature]*   Added Description, Group, Advanced, Hidden variables to Scheduling Portal ("Submit job" window and "Job variables" tab)
- *[New Feature]* Added the capability to have variables and values when sending a signal to a job

3.1.7 Resource Manager

- *[Improvement]*Administrators can now deploy resources using AWS spot instance - administrators can use the following ActiveEon system variables: provision of AWS EC2 Instances "AWSEC2Infrastructure" and provision of AWS EC2 Instances with auto-scaling "AWSAutoScalingInfrastructure".
- *[Improvement]* Disable all non-authorized items (menus, sub-menus, tabs, buttons) in the Resource Manager portal
- *[New Feature]* Node sources (pool of resources) can now be restricted by users' tenants.

- *[Improvement]* Improved the SSH infrastructure deployment (SSHInfrastructureV2) to allow various deployment scenarios (such as dynamic installation of ProActive libraries and Java Runtime Environment).
- *[Improvement]* Administrators can now allow the provisioning for more than 100 Azure instances at a time.
- *[Improvement]* Azure storage accounts automatically created by the infrastructure can now be restricted to a virtual network.

3.1.8 Automation Dashboard
- *[Improvement]* Collapse in/out the list of portals
- *[Improvement]* Fixed front-end vulnerabilities
- *[New Feature]* Added global variables within the submission windows

3.1.9 Catalog
- *[Improvement]* Display the generic info metadata of all catalog objects
- *[Improvement]* Allow changing the user group when uploading proactive examples
- *[New Feature]* Added descriptions and groups to basic-examples, controls, notifications, and data connectors workflows
- *[New Feature]* Added RBAC support for the Catalog: Fine access control of catalog objects and buckets based on username or user's group
- *[Hotfix]* Fixed the "select-all" issue when switching between buckets
- *[Hotfix]* Fixed the objects' operations button enabling/disabling when selecting objects
- *[Improvement]* When a bucket (folder) is created, it will automatically open it
- *[New Feature]* Display advanced, hidden, group, model and the description of variables on workflows
- *[Improvement]* Added the possibility to copy the workflow URL from the import windows
- *[Improvement]* Added filtering by bucket name
- *[Improvement]* Improved search & filter panel using auto-complete to easily find buckets and objects

3.1.10   Job-Planner
- *[Hotfix]* Fixed the displayed numbers of days of each month in the yearly tab CRON generator
- *[New Feature]* Added the possibility to submit a workflow from a Calendar-Workflow association
- *[Improvement]* Calendar Association search improvement
- *[Improvement]* *When* template workflows are launched, users can now append their names to workflows, and it becomes a Variable, that can be measured and monitored. This is specifically important when dealing with a large team of people launching the same workflow for their own requirements. Replace $var_name and ${var_name} by var.value when it's possible in Calendar Association and Execution Planning.
- *[Improvement]* Improved the performance and process time of backend functions by applying a cache mechanism
- *[Improvement]* Improved job-planner loop performance
- *[Improvement]* Allow skipping notifications of missed submissions for frequent calendars
- *[Hotfix]* Fixed the "select-all" issue when switching between calendars
- *[Hotfix]* Fixed the associations' operations button enabling/disabling when selecting objects

- *[Hotfix]* Fixed possible duplicate workflow submission when creating an association
- *[New Feature]* Display advanced, hidden, group, model and the description of variables on workflow associations
- Automatically add two Generic Information (GI) to the jobs submitted by the Job-Planner: **calendar.name** containing the calendar name and the **next.execution** representing the next execution date of the job.
- *[New Feature]* New panel which displays associations as a list
- *[Improvement]* Added a search box within the selected calendar
- *[Improvement]* Multi-select of associations by *Shift* key in both views (the block view and the list view)

3.1.11   Studio
- *[Hotfix]* Fixed front-end vulnerabilities
- *[Improvement]* Added filtering capabilities to User/Global dataspace file browsers
- *[Hotfix]* Fixed *Open in Studio from Catalog* causing an infinite reload of the Studio portal
- *[Improvement] In the Import/Publish Workflow window, the list of buckets/objects is now sorted. Also added filtering capabilities.*
- *[Improvement]* Keep the previously selected bucket when we publish/import a workflow to the catalog
- *[Improvement]* Added the possibility to copy the workflow URL from the import window
- *[Improvement]* Allow multi-line values in variables
- *[Improvement]* Improved Global/User space browser performance

3.1.12   ProActive Service Automation
- *[Documentation]* Added basic and advanced service creation tutorials

3.1.13   Analytics
- *[Improvement]* Have alphabetical order of Buckets

3.1.14   Proactive Workflow Templates

Self-service workflow templates for customer use
- *[New Feature]* Added 3 ServiceNow connectors to create an incident, mark an incident as resolved and send emails through ServiceNow
- *[New Feature]* Added an Informatica connector to create GET requests to an Informatica API.
- *[New Feature]* Added VMWare connectors templates to manage VMs lifecycle (Start, Stop, Restart, etc)
- *[New Feature]* Added a PeopleSoft connector to create GET requests to PeopleSoft API.
- *[New Feature]* Added a Nice DCV connector to enable an easy management of DCV sessions.
- *[New Feature]* Added workflow examples for Expect SSH, i.e., execute commands and scripts remotely using SSH while doing an interactive SSH authentication via Expect4j.
- *[New Feature]* Added an Apache Phoenix connector

3.2   ProActive AI Orchestration

3.2.1 ML Model as Service:

- *[Improvement]* Integrate multiple models' deployment and model versioning. The user can deploy multiple models and several versions of the same model type. The model version can be chosen by the user explicitly. If not, the version number will be the last model version number plus one. The models and their versions can be updated or deleted by specifying the model name and model version number.
- *[Improvement]* Update Model-As-A-Service (MaaS) ML workflows in the catalog within the MaaS bucket. An interactive self-service workflow example has also been added to run batch inferences through signals. The existing workflows have been also updated to support model versioning.
- *[New Feature]* We added the data analytics dashboard to compare baseline data coming with deployed model versions. By choosing the model's name, some charts are previewed, which provide the comparison, applied according to some chosen statistical methods, between the baseline data (data used in training each model version) of all model versions. This provides an overview of how the data is changing (drifting) from one model version to another.
- *[New Feature]* We added the option to save and visualize the performed predictions on different model versions. The user can apply predictions on different model versions of the same model type. AI engineers can choose to save and/or visualize the stored predictions.
- *[Documentation]* Add the documentation for the data analytics dashboard.
- *[Documentation]* Update the documentation for MaaS_ML.