About Autonomous Access

ADD-ON

Advanced Identity Cloud add-on capability

Contact your Ping Identity representative if you want to add PingOne® Autonomous Access to your Advanced Identity Cloud subscription. Refer to Add-on capabilities.

Autonomous Access leverages artificial intelligence (AI) and machine learning (ML) techniques to analyze incoming authentication events and identify abnormal online behavioral patterns. It provides an advanced threat detection solution powered by AI, protecting against account takeover and fraudulent activities at the identity perimeter.

By accelerating and simplifying access decisions, Autonomous Access empowers your organization to efficiently thwart threats while delivering personalized user experiences that enhance online interactions for legitimate users.

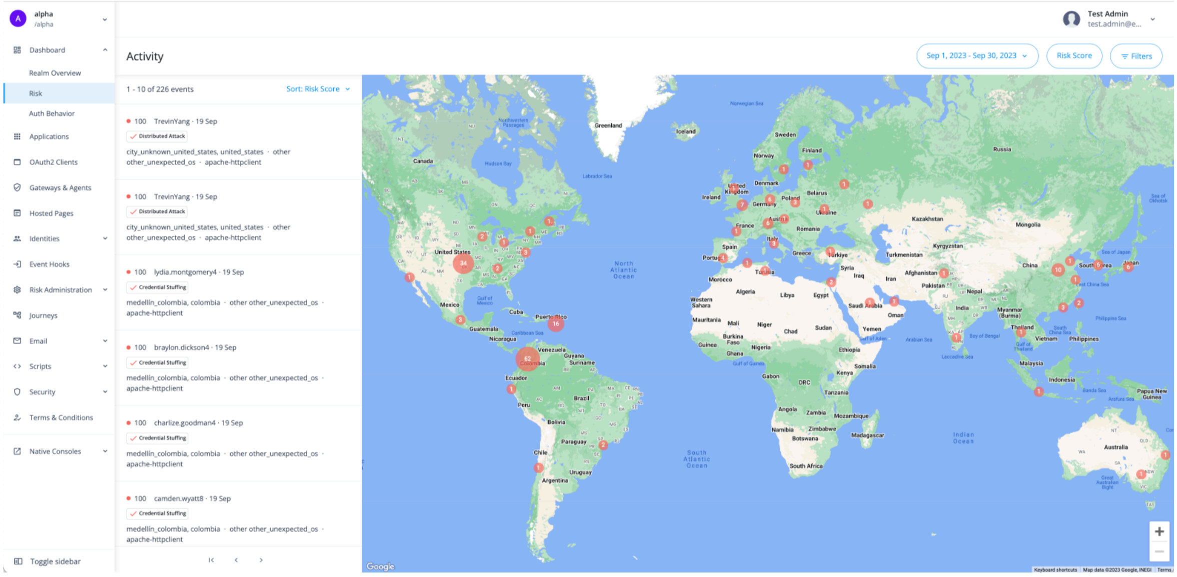
Ultimately, Autonomous Access seeks to provide insights based on the following questions:

- Is the user’s online behavior unusual compared to their typical behavior?
- If the user typically exhibits similar behavior to a specific group (for example, a department), is their current behavior deviating from the norm in this context?
- Does the user’s behavior differ from any other patterns observed on the platform?

Ping Identity implements Autonomous Access within your new or existing tenants (development, staging, and production), ensuring that your users' data and personally identifiable information (PII) remain exclusively within the tenant’s boundaries.

Activity dashboard

The Activity dashboard shows risky access activity and lets users drill in and investigate across time, risk reason, and risk score in the realm that you are currently in (for example, Alpha or Bravo).



1 - 10 of 226 events

100 TrevinYang · 19 Sep
Distributed Attack
city_unknown_united_states, united_states · other other_unexpected_os · apache-httpclient

100 TrevinYang · 19 Sep
Distributed Attack
city_unknown_united_states, united_states · other other_unexpected_os · apache-httpclient

100 lydia.montgomery4 · 19 Sep
Credential Stuffing
medellin_colombia, colombia · other other_unexpected_os · apache-httpclient

100 braylon.dickson4 · 19 Sep
Credential Stuffing
medellin_colombia, colombia · other other_unexpected_os · apache-httpclient

100 charlize.goodman4 · 19 Sep
Credential Stuffing
medellin_colombia, colombia · other other_unexpected_os · apache-httpclient

100 camden.wyatt8 · 19 Sep
Credential Stuffing
medellin_colombia, colombia · other other_unexpected_os · apache-httpclient

Figure 1. Autonomous Access activity dashboard

Autonomous Access nodes

Autonomous Access provides three prebuilt nodes and a journey template, eliminating the need for custom coding or connectors to implement these journeys. By adding these three nodes, you can fully leverage over 100 nodes to trigger actionable outcomes when high risk scores are discovered.

Risk scores reflect a combination of anomalous behavior and identified threats. You have the flexibility to determine the appropriate actions for different ranges of risk scores. For example, when Autonomous Access returns lower risk scores (e.g., numbers between 0 and 30), you can configure Autonomous Access to allow users to proceed without additional authentication steps. Conversely, for high-risk scores (e.g., numbers between 71 and 100), you can configure Autonomous Access to flag these events for escalation, such as requiring multifactor or step-up authentication, blocking, reviewing, informing the user, or other actions.

The following nodes are available:

- Autonomous Access signal node: Assesses risk based on anomalous user behavior, credential stuffing, suspicious IPs, automated user agents (bots), impossible travelers, and brute force attacks using AI/ML analytics. The result is a risk score from 0 (no risk) to 100 (high risk).
- Autonomous Access decision node: Maps the risk score to a high, medium, low, or unknown branch of a journey to direct the user experience.
- Autonomous Access results node: Sends data back to Autonomous Access for the dashboards and model learning.

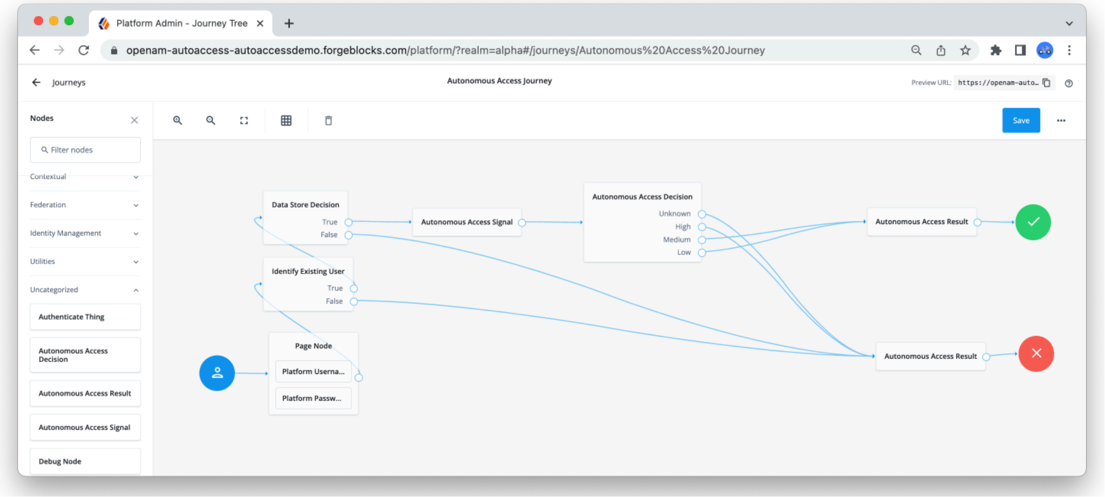


Figure 2. Autonomous Access nodes

Key features

ADD-ON

Advanced Identity Cloud add-on capability

Contact your Ping Identity representative if you want to add Autonomous Access to your Advanced Identity Cloud subscription. Refer to Add-on capabilities.

Autonomous Access features

Feature	Description
Fully-native Advanced Identity Cloud deployment	Ping Identity’s Autonomous Access and its components are fully cloud-native, deployed into your new or existing development, staging, and production tenants. The data collected by Autonomous Access is stored for three months in the Risk dashboard and six months in the cloud to ensure optimal artificial intelligence/machine learning (AI/ML) analytics.
Machine learning-based anomaly detection	Autonomous Access uses AI/ML-based detection analytics centered around user behavior and geospatial contextual information at authentication. Anomaly detection includes location, time of day, operating system version, device model and type, browser type and version, and other data. Autonomous Access’s AI/ML decisions are designed to be explainable, providing clear reasoning for its scoring instead of generating black box results with limited transparency.

Feature	Description
Fully GDPR compliant	The General Data Protection Regulation (GDPR) is a collection of European Union (EU) regulations designed to protect the privacy and personal data of users. GDPR grants an organization’s users greater control over their personal information and requires organizations to obtain explicit consent to access and remove their personal data. It also requires organizations to provide clear information about data processing and security measures to safeguard user data. Autonomous Access stores user data in the cloud for six months. Users can request to access or remove their personal data processed through Autonomous Access. Refer to Handling GDPR requests.
Real time threat detection	Autonomous Access AI/ML analytics engine discovers the risk threats described in Real time threat detection.
Autonomous Access dashboards	Autonomous Access presents multiple UI dashboards providing insights in the online behavior for tenant and individual users. - **Risk dashboard**. Displays an intuitive risk activity page showing all suspected access threats occurring in the past three months across a world-wide company. Authorized users can click on an event to drill down to examine the details. The dashboards display the risky events specific to the realm that you are in. - **Activity detail**. Displays a summary page for a selected risk event on the Risk dashboard. - **User access behavior**. Displays a graphical summary page of the typical access behavior for a selected user. You can access this page from the Activity detail page. - **Tenant access behavior**. Displays a graphical summary page of the typical access behavior for all users in the tenant. You can access this page on the Advanced Identity Cloud admin UI.

Feature	Description
Autonomous Access nodes	Three Autonomous Access nodes integrate within your journeys. No custom coding and connectors are required for these nodes. The following Autonomous Access nodes are available: • **Signal node**: The signal node determines the heuristics and anomaly detection to include in the risk score generation. The node begins making API calls to the Autonomous Access AI server to collect and extract data from a pre-defined data source. After you run an AI/ML training workflow to generate the risk scores and models, the Autonomous Access AI server returns the risk score and accompanying information for each event to the decision node. • **Decision node**: The decision node determines the actionable journey paths based on where a risk score lies within a predetermined range of scores. • **Result node**: The result node collects the risk predictions and results for successful and failed outcomes and writes them to the Autonomous Access AI server. The nodes are all specific to the realm that you are in. For further customizations, you can leverage the more than 100 Ping Identity nodes within your journeys to implement in your use cases. For more information, refer to Learn about the Autonomous Access nodes.
Out-of-the-Box journey	Advanced Identity Cloud provides a preconfigured Autonomous Access journey with nodes. You can use this journey as a starting template for your specific use cases and requirements. Advanced Identity Cloud Analytics dashboard also reports successful or failed Autonomous Access journeys. For more information, refer to Create journeys.

Feature	Description
Custom features	Autonomous Access lets you add custom features using YAML-based risk configuration and scripted nodes. For example, you can configure Autonomous Access with the following custom features: • **Multiple policies**: Companies typically require multiple risk policies for its various use cases. Autonomous Access provides a single risk policy out-of-the-box, but you can configure multiple policies. • **Custom logic**: Autonomous Access uses the highest risk score of all triggered signals by default. For example, if you have a UEBA signal with a score of 30 and an impossible traveler score of 60; the resulting score of these events is 60. However, you can also change the logic to use the sum of all triggered signals for your applications. For example, a UEBA score of 30 and an impossible traveler score of 60 results in a sum score of 90, which triggers a high risk.

Real time threat detection

Autonomous Access AI/ML analytics engine discovers the following risk threats:

- **Anomaly detection**. Autonomous Access's User and entity behavior analytics (UEBA) signal effectively identifies online anomalies in a user's behavioral profile. UEBA is a powerful security tool that utilizes machine learning to analyze network activity, detecting any deviations from a user's typical online behavior. This complementary tool can be seamlessly integrated with other threat signals for enhanced security measures.
- **Prevent double jeopardy**. Avoids flagging a user for the same reason or risk score if they already passed multifactor authentication. For example, if a user in France visits Singapore and gets flagged for an unusual location but successfully completes multifactor authentication, Autonomous Access will *not* flag the user again during their next login within a default time window (60 minutes) from the same city (Singapore).
- **Credential stuffing**: Identifies instances where a single IP address attempts to access multiple user accounts over a period of time by counting the total number of users accessed by that IP.
- **Suspicious IP**: Tracks the overall count of authentication attempts made by a single IP address across all users. An IP is flagged as *suspicious* if it exceeds a certain threshold of authentication attempts within a specified timeframe.
- **Automated user agent filter**: Detects if automated bots exist in the user-agent string. An automated bot is a program that operates independently, performing

tasks automatically without the need for human interaction. Hackers utilize automated bots to launch large-scale attacks, such as distributed denial-of-service (DDoS) attacks or credential stuffing, by leveraging the bots' ability to carry out malicious activities rapidly and at scale. can detect such malicious activity using its automated user agent filter heuristic.

- **Impossible Travel**: Detects if users are authenticated from two locations too far apart for a person to travel between these points at an impossible speed.
- **Brute force**: Detects the frequency of authentication attempts for a user over a period of time. If the frequency is high, then Autonomous Access flags the event as a possible brute force attack.
- **Distributed attack**: Detects whether the number of authentication attempts by a single user exceeds a predefined threshold of unique IP addresses within a specified time period. For example, if the threshold is set to 7 and the window is set to 10 minutes, Autonomous Access raises a *distributed attack* flag if the same user makes authentication attempts from 8 or more distinct IP address within a span of 10 minutes. The only action is to display the risk score on the Risk dashboard, so that the administrator can adjust the login journey to block or challenge this activity.
- **Allow/block IP addresses**: Autonomous Access provides two important features to mitigate against cases where known IPs can be triggered as false positives and known malicious IP addresses that are associated with harmful activities on the Internet: allow IP lists and block IP lists.
    - **Allow IP Addresses**. This feature allows you to override a risk score when dealing with specific IP addresses triggering high-risk scores. Instead of assigning a high-risk score, it sets the risk score to 0. For example, many users and organizations use VPNs to access online services. However, VPN usage can often trigger a false positive related to credential stuffing because multiple users are coming from the same IP address. To address this, you can add the VPN’s IP address to an *allow list*. When an IP address is on this list, Autonomous Access assigns it a risk score of 0, bypassing heuristic and machine learning processes.
    - **Block IP Addresses**. This feature allows you to override any calculated risk score and set it to 100 for IP addresses known to be malicious. For example, if you want to block access from known malicious IP addresses completely, you can add them to a *block list*. When an IP is on this list, Autonomous Access subjects it to all configured heuristics and machine learning processes, calculates a risk score, and then overrides the calculated risk score by assigning a score of 100, indicating a high-risk state.

> IMPORTANT
>
> Autonomous Access is not a firewall. You must consume the output risk score in a succeeding node in the journey for actionable outcomes. Autonomous Access cannot allow or block any IP address by itself.

Use dashboards to detect risks

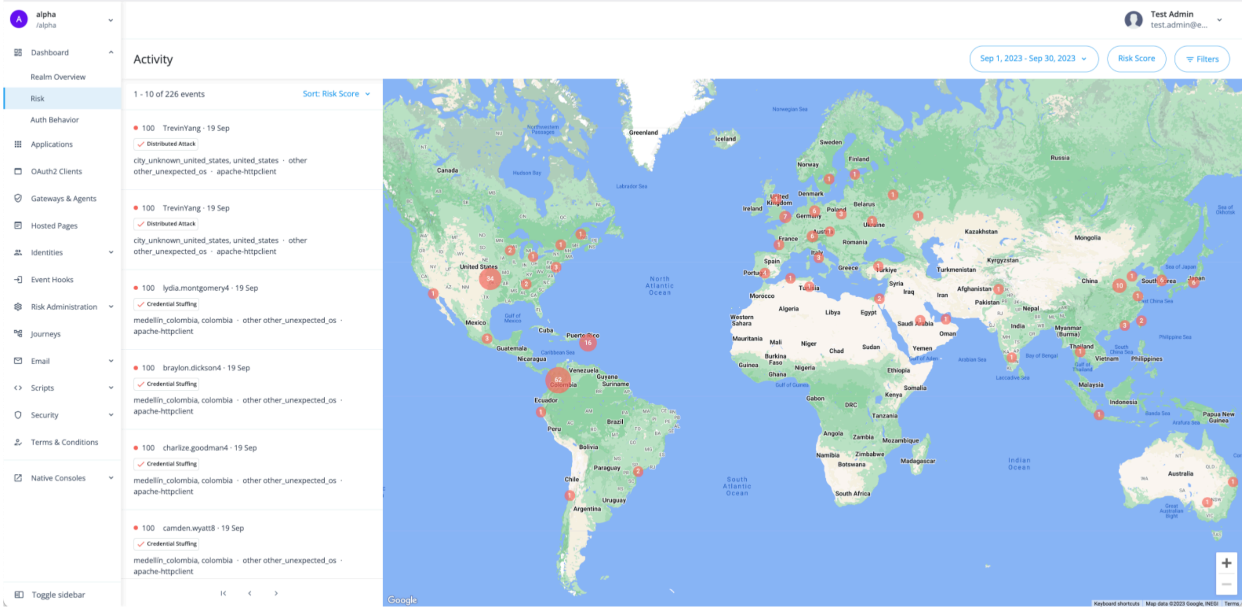
Autonomous Access provides a simple and intuitive graphical UI displaying the risks detected across your organization. The dashboard displays the risky events specific to the realm that you are in.

This section provides an overview of the available Autonomous Access UIs with your deployment.

Access the Risk dashboard

On the **Risk** dashboard, Autonomous Access displays a global map featuring sortable list of "anomalous" or risky events discovered during the AI/ML analytics run.

To access the **Risk** dashboard, from the left navigation pane of the Advanced Identity Cloud admin UI, click **Dashboard > Risk**.



1 - 10 of 226 events

100 TrevinYang · 19 Sep
✓ Distributed Attack
city_unknown_united_states, united_states · other other_unexpected_os · apache-httpclient

100 TrevinYang · 19 Sep
✓ Distributed Attack
city_unknown_united_states, united_states · other other_unexpected_os · apache-httpclient

100 lydia.montgomery4 · 19 Sep
✓ Credential Stuffing
medellín_colombia, colombia · other other_unexpected_os · apache-httpclient

100 braylon.dickson4 · 19 Sep
✓ Credential Stuffing
medellín_colombia, colombia · other other_unexpected_os · apache-httpclient

100 charlize.goodman4 · 19 Sep
✓ Credential Stuffing
medellín_colombia, colombia · other other_unexpected_os · apache-httpclient

100 camden.wyatt8 · 19 Sep
✓ Credential Stuffing
medellín_colombia, colombia · other other_unexpected_os · apache-httpclient

Each event displays a summary of the event with the following information:

Data Element	Description
Risk Score	Risk score associated with the event. A risk score is a combination of the likelihood an anomalous behavior event and/or a known threat is detected. Therefore, a risk score of 100 indicates the highest likelihood an access attempt was an anomaly and/or a known threat, such as suspicious IP, credential stuffing, or others.
User Identifier	Username on the account.
Date of Occurrence	Date the anomalous event occurred.

Data Element	Description
Heuristics	Type of heuristic used in the AI/ML analytics. The heuristic indicates the type of risk threat. These include: • `Automated User Agent` • `Brute Force` • `Credential Stuffing` • `Distributed Attack` • `Impossible Travel` • `IP Blocklist` • `Suspicious IP`
City, Country	Geolocation information for city and country.
Browser Type	Browser type can include: • operating system (for example, 'Linux', 'Android', 'Windows', 'iOS') • operating system version (for example, '10', 11', 'other_unexpected_os') • browser (for example, 'Firefox', 'Chrome', and others) • user agent type (for example, user agent request header, such as 'apache-httpclient')

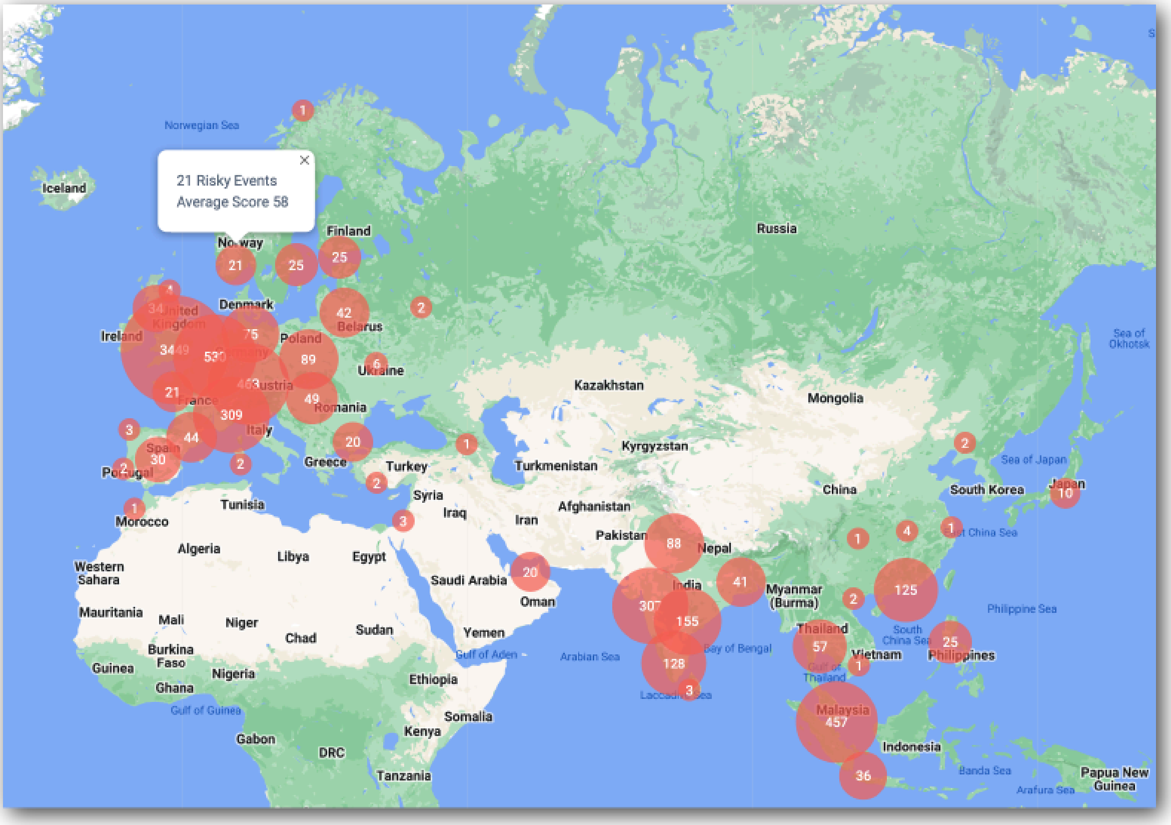
TIP

You can sort the activity events list by clicking **Sort: Risk Score** and selecting one of the following:

- Descending risk score (default sorting)
- Event time
- Username

In the right column, the Risk dashboard displays a world map with the number of risky events in each circle. By default, the dashboard displays all risk scores of 50 and above. You can change this setting by change the **risk_score_threshold** value in the Risk Configuration. Refer to Configure the risk settings.

You can click and drag directly on the map to access events in other countries. If you click one of the red numbers, the dashboard displays a summary of risky events and the average risk score associated with those events.

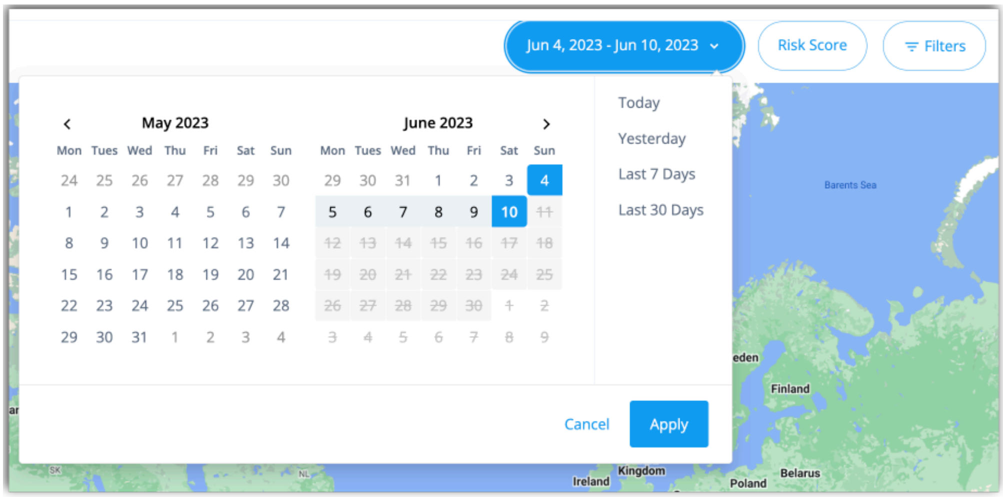


21 Risky Events
Average Score 58

Filtering by date range

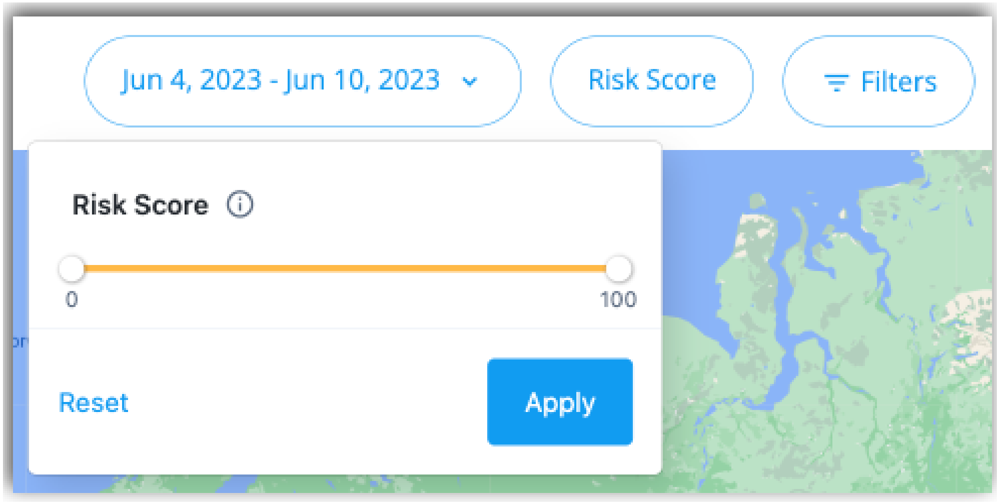
You can filter the display and risky events in a number of useful ways. If you select the **Date** filter, you can filter the dashboard to display only certain dates and ranges. The options are:

- **Today**: Display only events occurring today.
- **This Week**: Display only the events occurring this week.
- **This Month**: Display only the events occurring this month.
- **Custom**: Display the events based on your selected calendar date range.



Filter by risk score

The Risk dashboard supports filtering by risk score range.



Use advanced filtering

The Risk dashboard supports advanced filtering based on attribute type and value as well as heuristic filter. The attributes are derived from the information in the user agent string. The AI/ML jobs use the heuristics occurring in the selected timeframe.

1. On the **Risk** dashboard, click **Filters**. A filters popup window appears.
2. Under **Attribute Filters**, click the **Feature** drop-down list, and select one of the following attributes:

Attribute	Description
City	City where the risky event occurred, for example `Bristol or Toronto.`

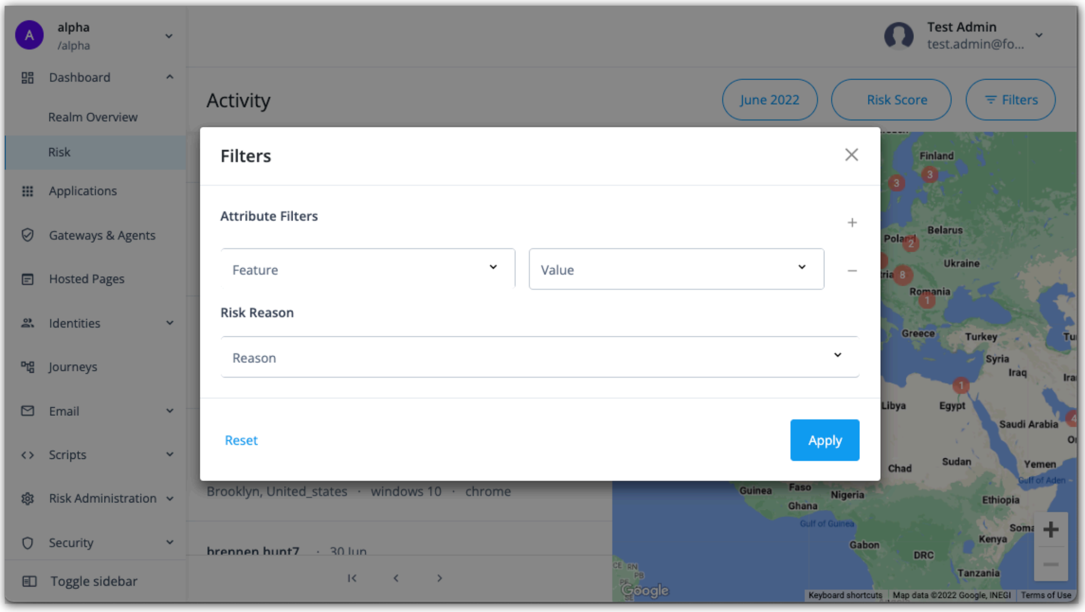
Attribute	Description
Country	Country where the risky event occurred, for example `United States`, `Singapore` , or `Great Britain`.
Device	Device that may have made the risky action, for example `iPad`, `Mac`, or `Other`.
Device Type	Device type of the system that made the risky action, for example `Apple, or Samsung`.
OS	Operating system of the computer, for example `iOS`, `Linux`, or `Android`.
OS Version	Operating system version, for example `10`, `11`, or `14`.
Time of Day	Timestamp of the risk event.
User Agent	User-Agent request header lets servers identify the browser and operating system to the web server.
User ID	User ID of the account that may have been compromised.

3. Enter a value for the attribute filter you created in the previous step.
4. Click the **Risk Reason** menu to select one of the following heuristics:

Heuristic	Description
Automated User Agent	A detected threat where an automated bot is in the user-agent string.
Brute Force	A detected threat where direct users are failing multiple authentication attempts.
Credential Stuffing	A detected threat where an IP address is attempting to access a number of different users in a period of time.
Distributed Attack	A detected threat where the total number of unique IP addresses making authentication attempts exceeds a predetermined threshold within a specified time frame.
Impossible Travel	A detected threat where an attacker runs multiple authentication attempts from various locations in a short time span, making such travel impossible for a single person.

Heuristic	Description
IP Blocklist	A detected threat where the user’s IP address is on the Block IP Address list.
Suspicious IP	A detected threat where a user at an IP address is making many authentication attempts over a period of time.
Unusual Browser	A detected threat where a user at an IP address is making authentication attempts on a different browser.
Unusual City	A detected threat where a user at an IP address is making authentication attempts at a different city.
Unusual Country	A detected threat where a user at an IP address is making authentication attempts in a different country.

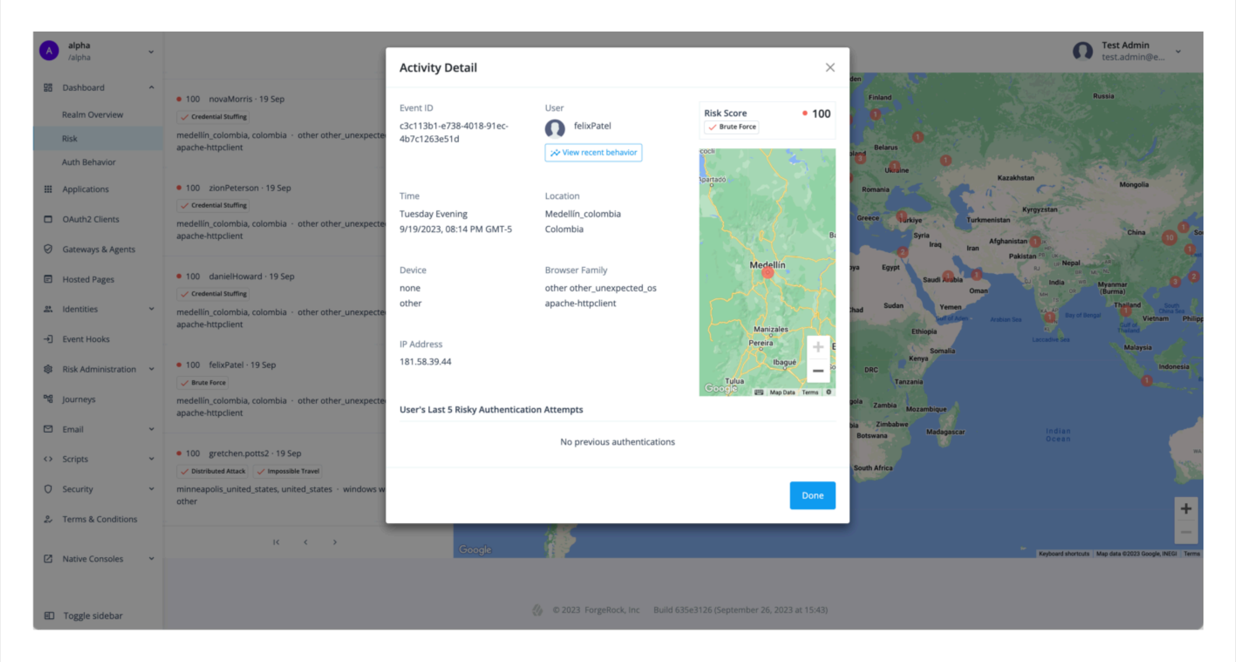
5. When done, click **Apply**. The Risk dashboard displays only those criteria that matches your filter.



Access risk activity detail

To investigate a particular risk event, click the activity event in the left-hand column. The **Activity Detail** popup window displays the specific details of the event including transaction ID, user, risk score, time, location, device, IP, and user agent plus a map of where the activity took place.

The popup window also lists the user’s last five risky authentication attempts and the possible type of heuristic discovered (for example, Automated User Agent). Any category that differs from normal behavior is marked with "Unusual <category>" (for example, "Unusual City").



Activity Detail

Event ID
c3c113b1-e738-4018-91ec-4b7c1263e51d

User
felixPatel
View recent behavior

Risk Score 100
Brute Force

Time
Tuesday Evening
9/19/2023, 08:14 PM GMT-5

Location
Medellín_colombia
Colombia

Device
none
other

Browser Family
other other_unexpected_os
apache-httpclient

IP Address
181.58.39.44

User's Last 5 Risky Authentication Attempts

No previous authentications

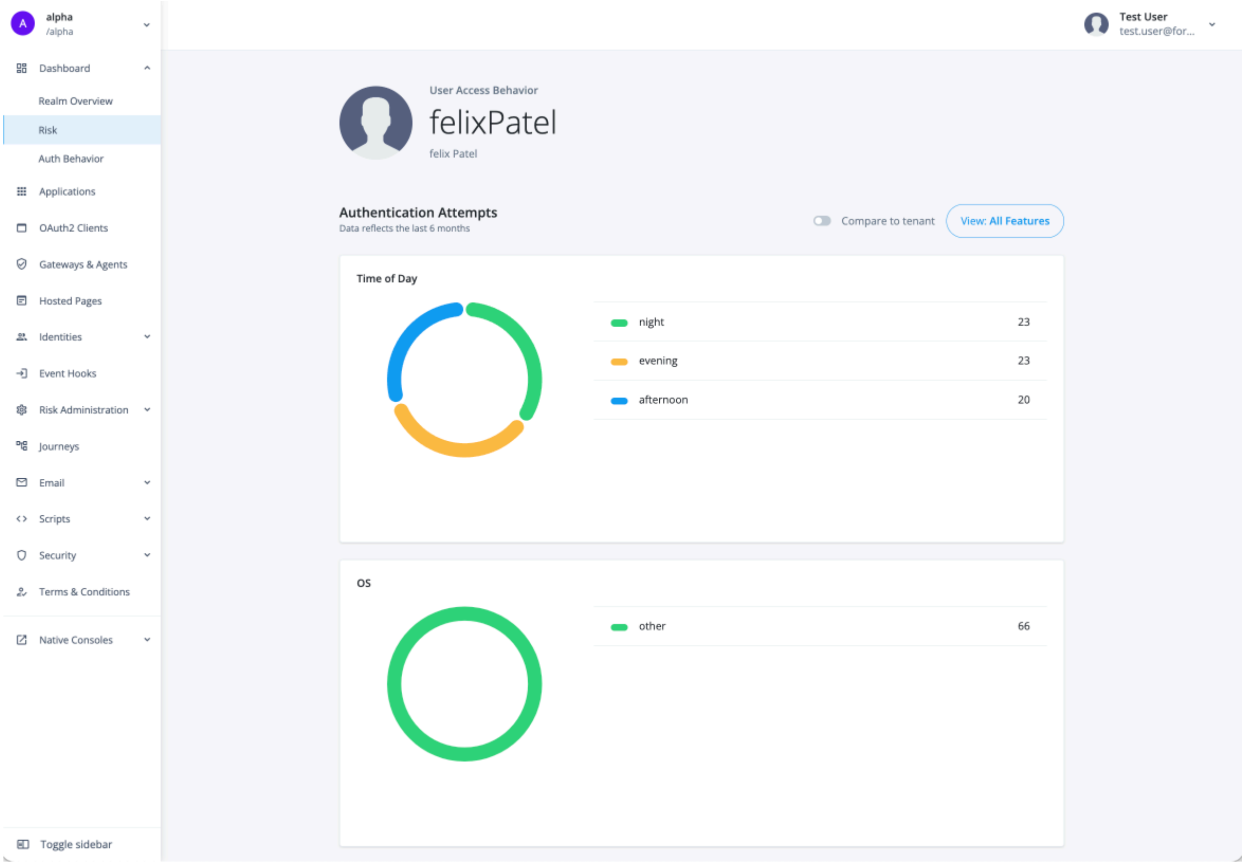
Done

© 2023 ForgeRock, Inc   Build 635e3126 (September 26, 2023 at 15:43)

Get the user access behavior

Autonomous Access provides two innovative UI features: User access behavior and Tenant access behavior pages. These UIs pages provide insights into the *typical* authentication behavior for a specific user and for all users in the tenant.

User access behavior provides administrators with a comprehensive graphical overview of a user’s typical authentication attempts over the past six months. The User access behavior page lets administrators gain insights into their users' typical login attempt patterns, making it easier to spot and respond to any irregularities swiftly. An additional feature lets administrators compare a user’s authentication behavior to that of all users in the tenant.



User Access Behavior

felixPatel

felix Patel

Authentication Attempts

Data reflects the last 6 months

Time of Day

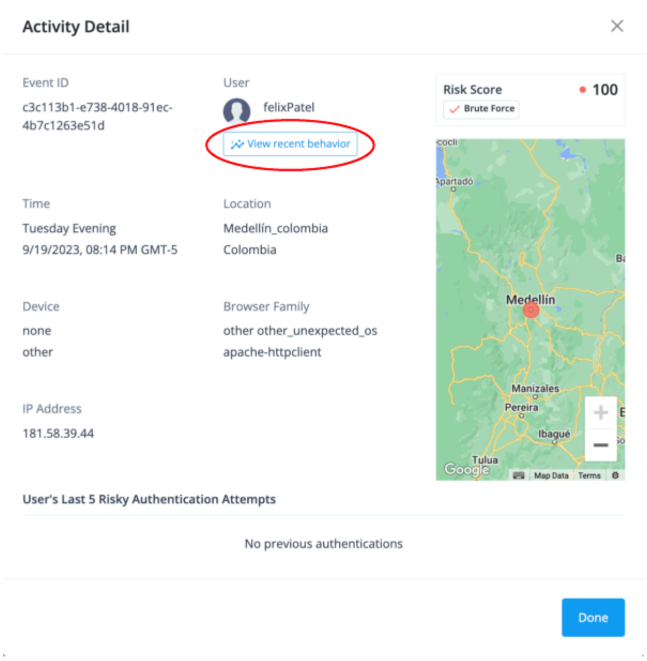
night	23
evening	23
afternoon	20

OS

other	66

The **User access behavior** page displays the *typical* behavior for a specific user in a tenant for the past six months.

1. Log in to your tenant.
2. On the Advanced Identity Cloud admin UI, click **Risk** to access the Risk dashboard.
3. On the Risk dashboard, under **Activity**, select a specific user to display the user’s Activity Detail page.
4. On the Activity Detail page, click **View recent behavior** under the user’s username.



Activity Detail

Event ID
c3c113b1-e738-4018-91ec-4b7c1263e51d

User
felixPatel

Risk Score • 100
Brute Force

Time
Tuesday Evening
9/19/2023, 08:14 PM GMT-5

Location
Medellín_colombia
Colombia

Device
none
other

Browser Family
other other_unexpected_os
apache-httpclient

IP Address
181.58.39.44

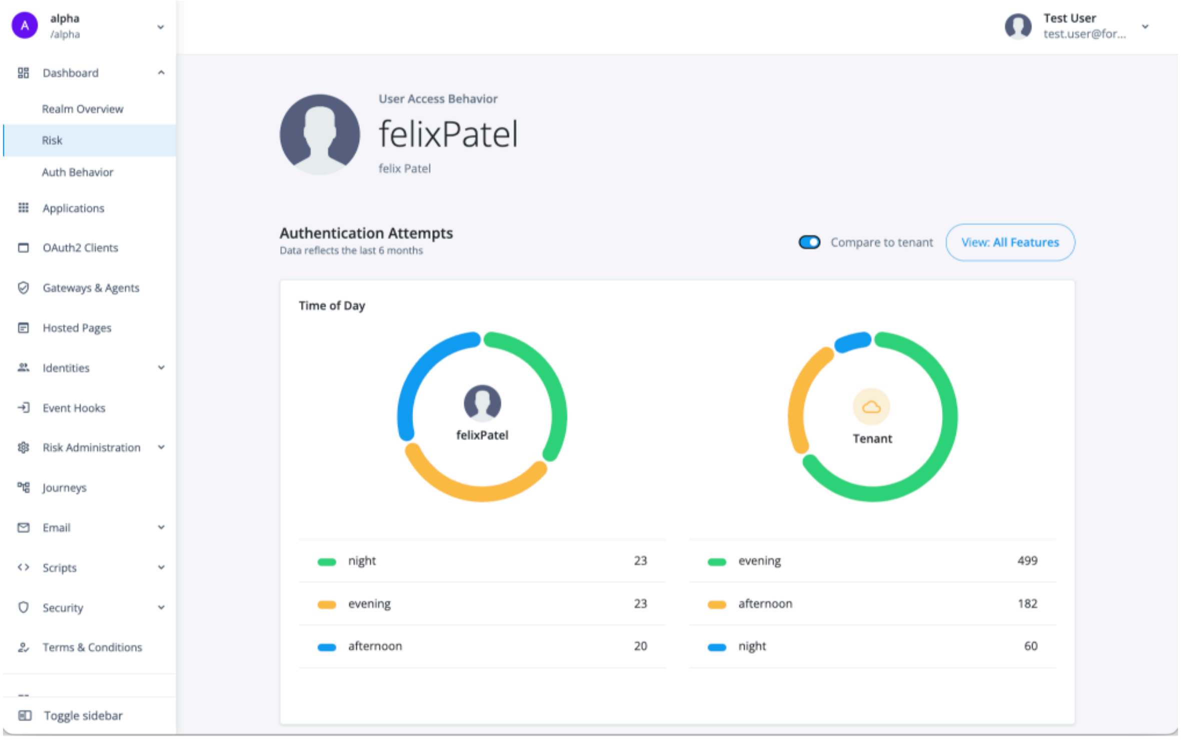
User's Last 5 Risky Authentication Attempts

No previous authentications

The User Access Behavior page appears.

5. You can compare the user’s typical authentication attempts statistics to that of all users in the tenant. Click **Compare to tenant**.

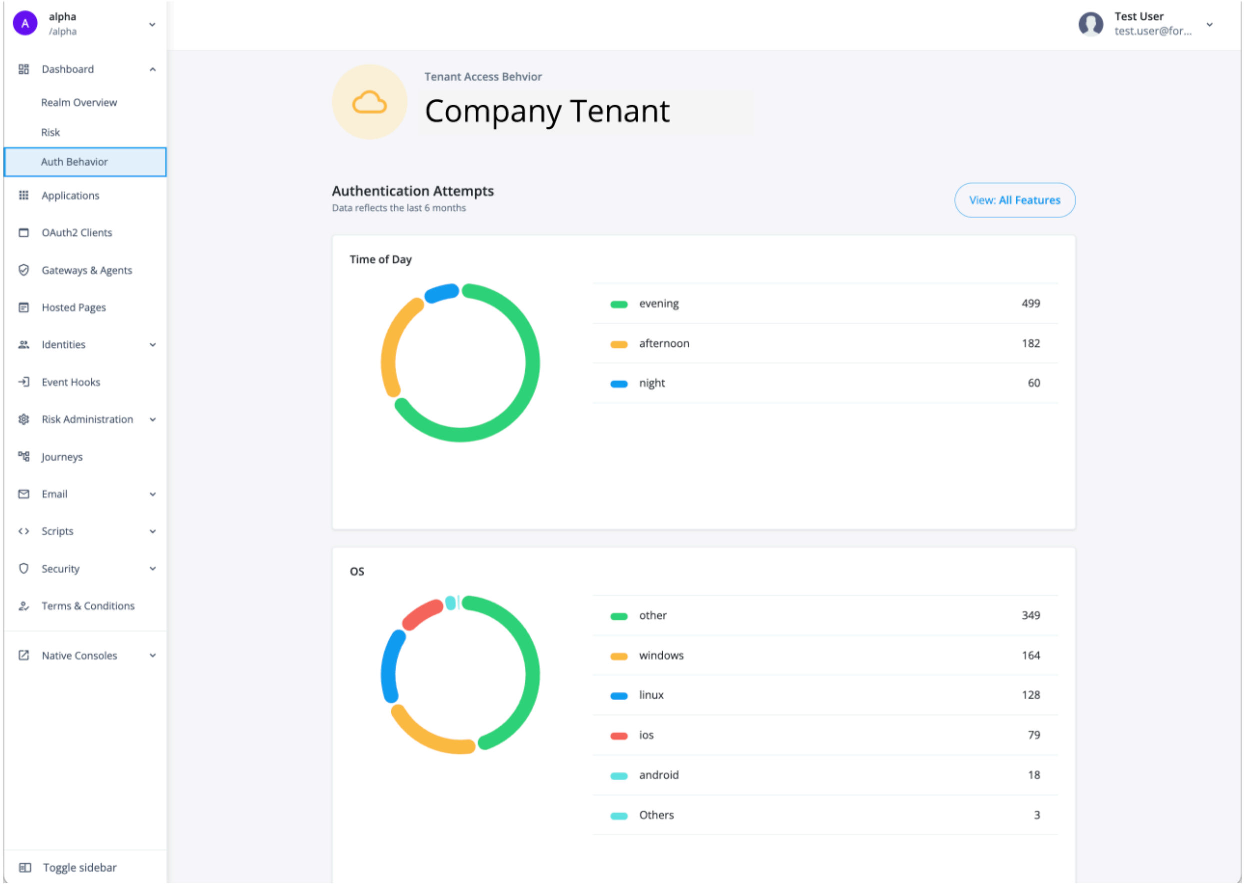
   The User Access Behavior page displays the user’s typical authentication attempts behavior on the left-hand side and all users' access attempts behavior on the right-hand side.



🟩 night	23	🟩 evening	499
🟧 evening	23	🟧 afternoon	182
🟦 afternoon	20	🟦 night	60

Get the tenant access behavior

Tenant access behavior provides administrators a comprehensive graphical overview of authentication attempts considered typical for all users in the tenant over the past six months. This page proves valuable when pinpointing policy anomalies by contrasting them with the usual authentication behaviors of your tenant’s users. For example, administrators can use this page when multiple individuals are violating an IT policy that restricts the use of personal computing devices.



Tenant Access Behvior

Company Tenant

Authentication Attempts
Data reflects the last 6 months

Time of Day

evening	499
afternoon	182
night	60

OS

other	349
windows	164
linux	128
ios	79
android	18
Others	3

The **Tenant access behavior** page displays the typical behavior of all users in a tenant for the past six months.

- On the Advanced Identity Cloud admin UI, click **Dashboard** > **Auth behavior**.

Learn about the available categories

Both the User access behavior and Tenant access behavior pages display the "typical" access information for the following categories:

Stat	Description
Time of Day	Time indicator based on the hour and specific to each region as to when the user(s) attempted a login: • **morning**: If the hour is greater than or equal to `5` and the hour is less than or equal to `11` , then set the time of day to "morning". • **afternoon**: If the hour is greater than or equal to `12` and the hour is less than or equal to `17` , then set the time of day to "afternoon". • **evening**: If the hour is greater than or equal to `18` and the hour is less than or equal to `22` , then set the time of day to "evening". • **night**: Any other time of day other than `morning` , `afternoon or evening` is set to `night` .
OS	Indicator for the typical operating system when the user(s) attempted a login. • windows • linux • ios • android • Others: Indicates that is the sum of all other features that are not classified in the top 5 or top 10 categories in the UI. • other: Indicates that the ML was not able to ascertain an OS value from the user-agent string, or an unknown empty value was encountered.
Device	Indicator for what device was used at login. • mac • iphone • generic_smartphone • other: Indicates that the ML was not able to ascertain the device type from the user-agent string, or an unknown empty value was encountered.

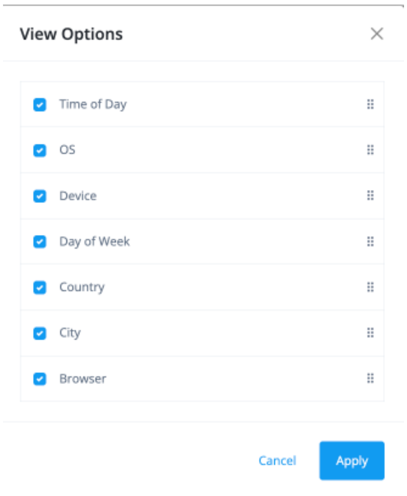
Stat	Description
Day of week	Indicator for the typical day users log in: • sunday • monday • tuesday • wednesday • thursday • friday • saturday
Country	Indicator for the country where the login typically occurs. Examples are: • united_states • colombia • germany • Others: Indicates that the ML was not able to ascertain a country value from the user-agent string, or an unknown empty value was encountered.
City	Indicator for the city and country that the login typically occurs. Examples are: • beijing_china • tokyo_japan • paris_france • new_york_united_states • city_unknown_united_states • Others:Indicates that the ML was not able to ascertain a city value from the user-agent string, or an unknown empty value was encountered.

Stat	Description
Browser	Indicator for the typical browser type used at login. • chrome • firefox • apache-httpclient • mobile_safari_ui_wkwebview • other: Indicates that the ML was not able to ascertain the browser type from the user-agent string, or an unknown empty value was encountered.

Filter viewable categories

You can filter the viewable categories on the User Access Behavior and Tenant Access Behavior pages as needed.

1. On the user's **User Access Behavior** (or **Tenant Access Behavior**) page, click **View All Features**.



View Options

☑ Time of Day

☑ OS

☑ Device

☑ Day of Week

☑ Country

☑ City

☑ Browser

Cancel Apply

2. On the **View Options** modal, click or clear a specific category, and then click **Apply**.

   The User Access Behavior or Tenant Access Behavior page displays the categories you enabled accordingly.

Deploying Autonomous Access

Autonomous Access is an add-on capability available to new and existing PingOne Advanced Identity Cloud customers who sign up for this feature.

Learn about the deployment steps

Once a customer has signed up for Autonomous Access, ForgeRock staff deploys the service in the following manner:

1. **Integration**. ForgeRock adds Autonomous Access to the private tenants of new or existing PingOne Advanced Identity Cloud customers. For more information, refer to Learn about integration.
2. **Set up journeys**. After ForgeRock integrates the Autonomous Access service into your tenants, the ForgeRock Professional Services (FPS) team configures an Autonomous Access journey for initial data collection. For more information, refer to Create journeys.
3. **Set or confirm data sources**. The next step is to define a training dataset, or data source. For Advanced Identity Cloud, Autonomous Access uses a default data source, `autoaccess-ds` that pulls in data from the data lake. You only need to verify that the `autoaccess-ds` data source is present and active.

   Typically, you also need to define a mapping to match customer data attributes to the Autonomous Access schema for training. However, you can skip this step due to the `autoaccess-ds` data source, which already maps its attributes. For more information, refer to Set the data sources.
4. **Set the risk configuration**. Set the AI/ML configuration and threshold properties required by Autonomous Access on the Risk Config page. For more information, refer to Configure the risk settings.
5. **Run training**. The training pipeline job is initiated to train the AI/ML models. When the Autonomous Access nodes are configured in an authentication journey, heuristics are enabled. ForgeRock encourages our customers to create and run the training pipelines and evaluate the models for accuracy. Administrators can run training on a periodic basis (e.g., bi-weekly or bi-monthly) and as soon as the Autonomous Access journeys begin to collect data. For more information, refer to Run training.
6. **Tune the models**. Once you get an initial risk model, you must tune it for improved training performance. After you tune the models, you must rerun the training job to update your models. For more information, refer to Tune the model.

1. **Publish the model**. After the predictions job completes, you will be asked to publish the model for later use.
2. **Grant roles**. Grant roles to users to access either Autonomous Access dashboards, configuration, or both. For more information, refer to Grant Autonomous Access roles.

Learn about integration

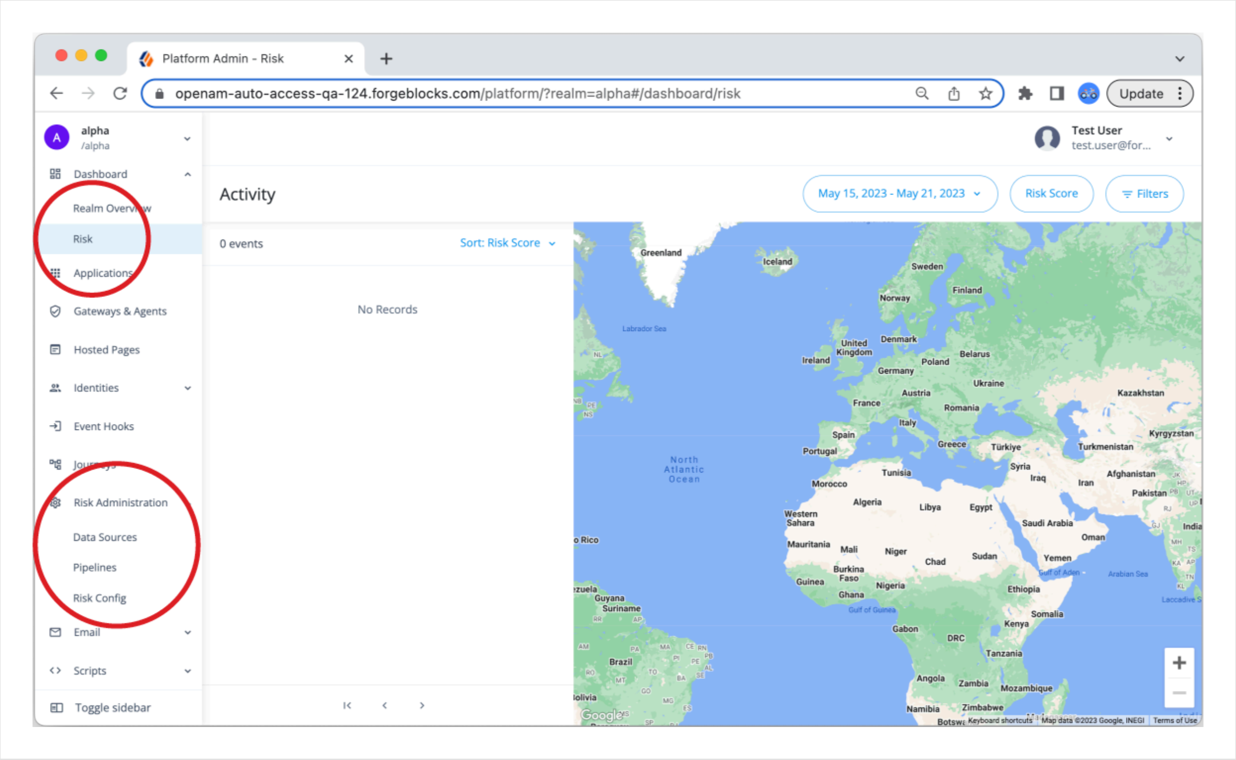
Ping Identity has automation in place to quickly add the Autonomous Access capability to the private tenants (dev, staging, and production) of new or existing PingOne Advanced Identity Cloud customers.

Your SLA determines the turnaround time for adding the service.

Once ForgeRock integrates the service into your tenants, administrators are automatically granted two Autonomous Access roles needed for operation:

- **Fraud analyst**. The fraud analyst role lets users access the Risk dashboard, but they cannot run any training or configuration.
- **Data analyst**. The data analyst role lets users access and run the risk administration menu options on Advanced Identity Cloud, but they cannot access the Risk dashboard.

Verify that the Autonomous Access links exist on the Advanced Identity Cloud UI:



NOTE

If you do not have the Autonomous Access roles, contact Ping Identity or refer to Autonomous Access roles to add them.

Create journeys

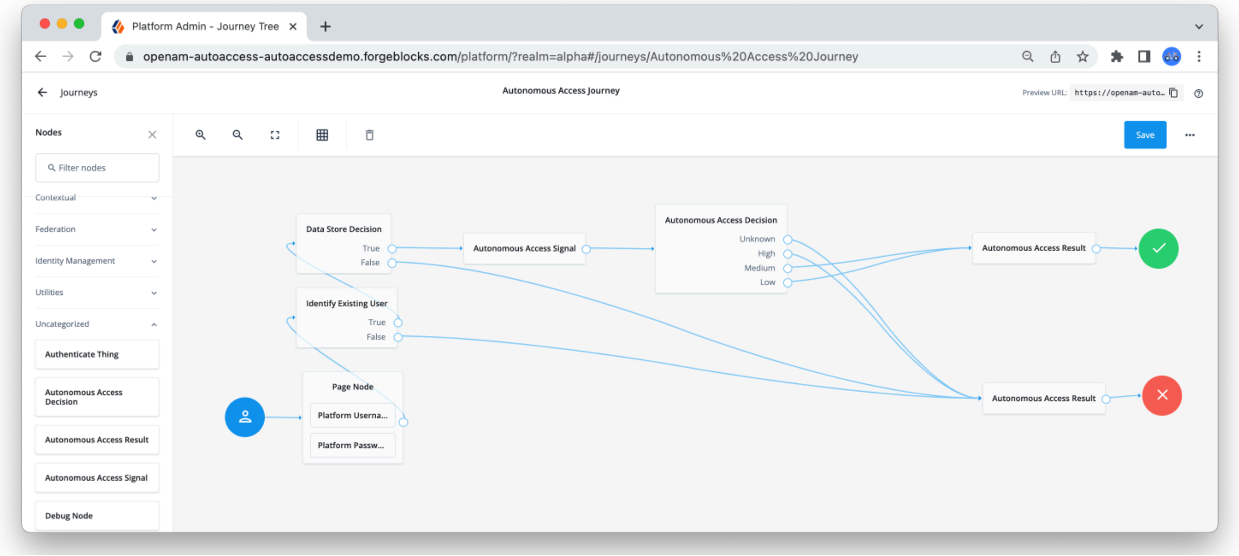
When your company purchases the Autonomous Access add-on capability, Ping Identity Professional Services configure an example journey in your tenants using three Autonomous Access nodes.

This section outlines several example journeys to illustrate how can configure journeys for your specific use cases. Consult with your ForgeRock representative to formulate your specific journeys.

Example journey

ForgeRock uses the following example journey for data collection and API calls to the Autonomous Access AI server to get the risk scores. This journey is also a good starter template to create more advanced journeys for production purposes.

Example Autonomous Access journey



1. In the Advanced Identity Cloud admin UI, go to **Journeys**, and edit the Autonomous Access template.
2. Provide details for these nodes in the journey:
   - Page node
   - Platform Username node
   - Platform Password node
   - Identify Existing User node
   - Data Store Decision node
   - Autonomous Access Signal
   - Autonomous Access Decision
   - Autonomous Access Result
   - For information about all available nodes, refer to Authentication nodes reference. For Autonomous Access nodes, refer to Learn about the Autonomous Access nodes.
3. To test the journey, copy the Preview URL, and paste the URL into a browser using Incognito or Browsing mode.
4. When you’re satisfied with your journey, click **Save**.

Example journey with step-up multifactor authentication

The following example journey illustrates a step-up multifactor authentication (MFA) with Autonomous Access nodes.

> NOTE
>
> There are various ways to implement MFA using journeys. Consult with your Ping Identity representative to discuss your particular application.

This example MFA setup uses the ForgeRock Authenticator (OATH) module, which supports HMAC one-time password (HOTP) and a time-based one-time password (TOTP) authentication method. It is assumed that the user has an OATH-compliant device that can provide a password.

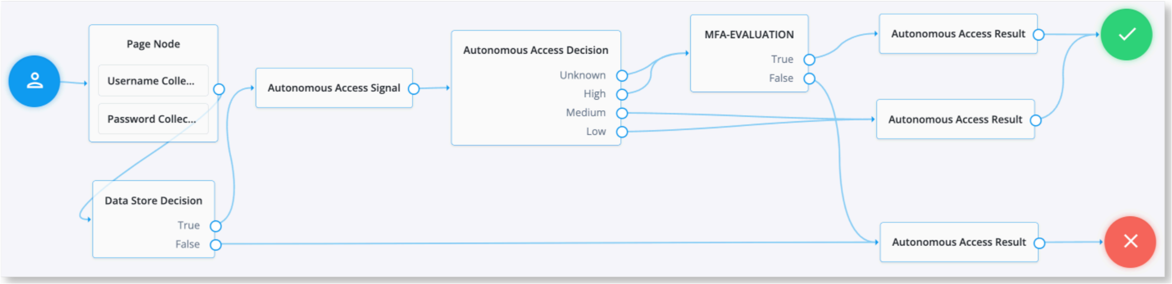
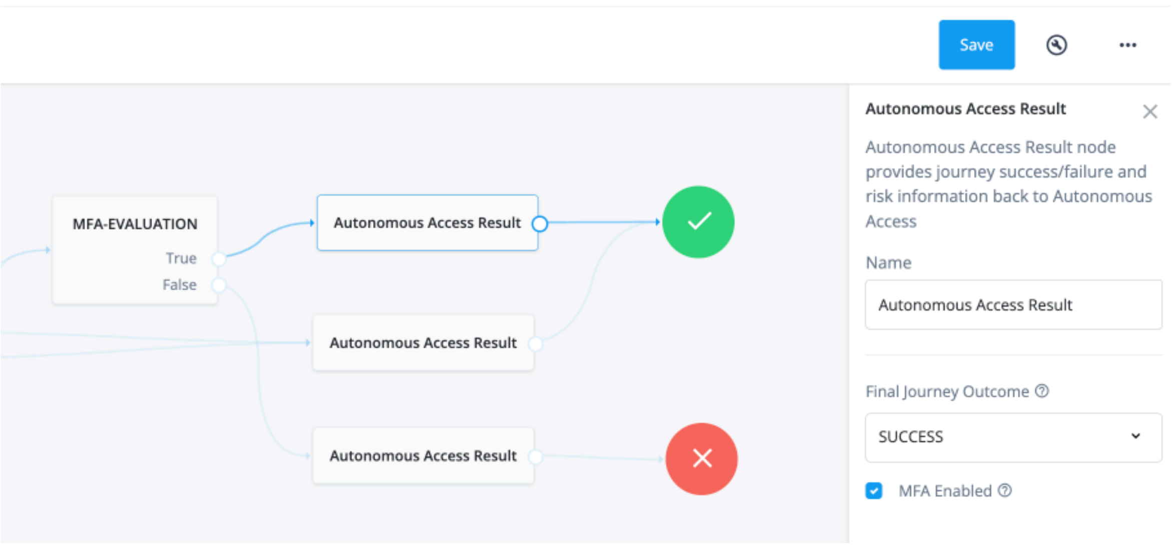


Figure 3. Autonomous Access journey with step-up authentication

The journey starts off in the same manner as the Example journey in the previous section, except that the journey uses an Inner Tree Evaluator node, relabelled as "MFA-EVALUATION" that calls a subtree called "PushStepUp."

The journey also implements three Autonomous Access result nodes: two "SUCCESS" nodes and one "FAILURE" node. The first result node takes the destination of a true outcome in the "MFA-EVALUATION" node. The node has a "SUCCESS" outcome and **MFA Enabled** checkbox enabled, so that the AAI server can keep track of users and knows that MFA has completed successfully.

The second result node has a "SUCCESS" outcome for low decision scores. This node does not have `MFA Enabled` as low and medium risk scores are passed through successfully. The third result node is for "FAILURE" states and has `MFA Enabled` checked, so that Autonomous Access knows that the MFA journey has failed during the its process.



Autonomous Access Result

Autonomous Access Result node provides journey success/failure and risk information back to Autonomous Access

Name

Autonomous Access Result

Final Journey Outcome

SUCCESS

☑ MFA Enabled

Figure 4. Autonomous Access result node

The MFA EVALUATION node calls a subtree, PushStepUp, as follows:

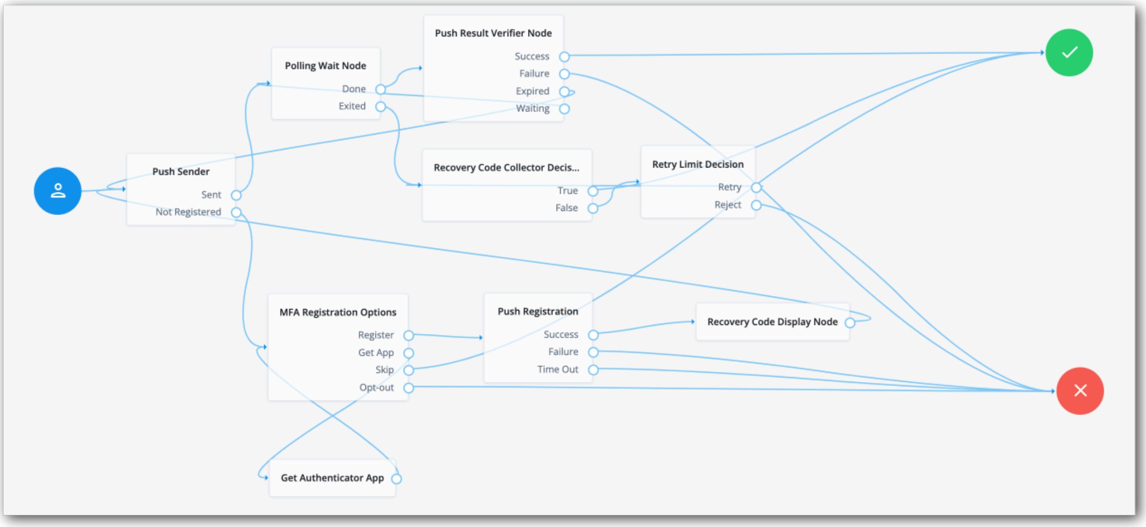


Figure 5. Push step up subtree journey

1. In the Advanced Identity Cloud admin UI, go to **Journeys**, and edit the Autonomous Access template.
2. Provide details for these nodes in the MFA journey:
   - Page node
   - Platform Username node
   - Platform Password node
   - Autonomous Access Signal
   - Autonomous Access Decision
   - Autonomous Access Result

    - The MFA EVALUATION node is an Inner Tree Evaluator node that calls the PushStepUp journey.
    - The Validate Credentials node is a renamed Data Store Decision node.
    - Push Sender node
    - Polling Wait node
    - MFA Registration Options node
    - Get Authenticator App node
    - Push Result Verifier node
    - Recovery Code Collector Decision node
    - Push Registration node
    - Retry Limit Decision node
    - Recovery Code Display node
    - For information about all available nodes, refer to Authentication nodes reference. For Autonomous Access nodes, refer to Learn about the Autonomous Access nodes.

3. To test the journey, copy the Preview URL, and paste the URL into a browser using Incognito or Browsing mode. A login screen prompts you to enter your user ID and password.
4. Verify that you can use the ForgeRock Authenticator application to perform MFA. For more details, refer to Authentication nodes and journeys.
5. When you're satisfied with your journey, click **Save**.

Set the data sources

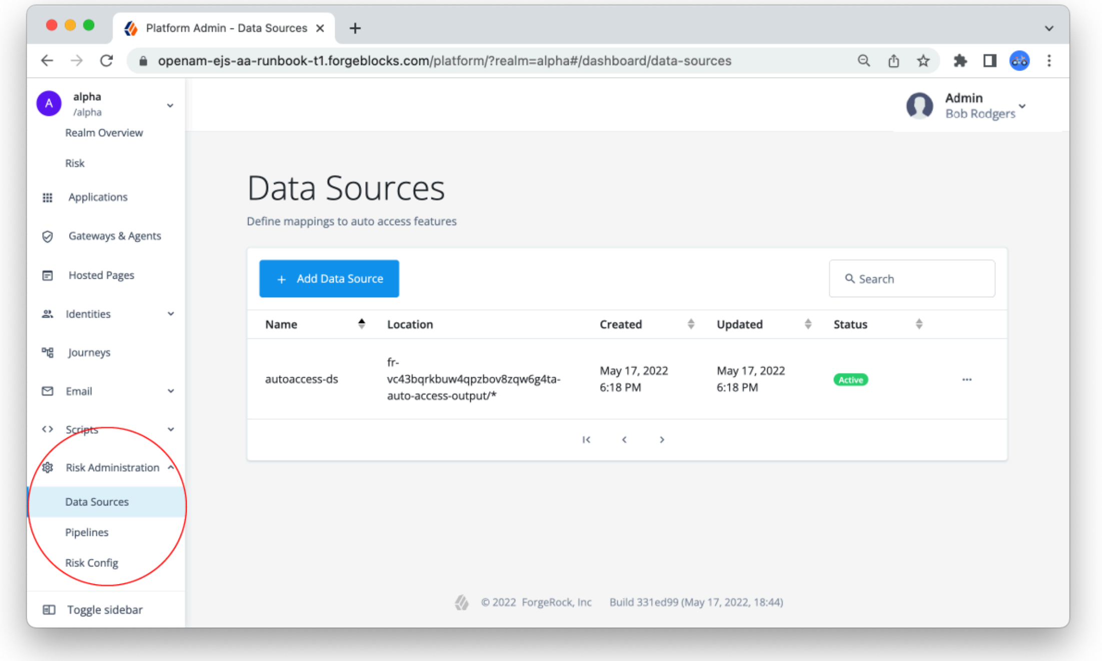
IMPORTANT

IMPORTANT

You can skip this section for Advanced Identity Cloud tenants. This section is presented for information only.

Autonomous Access automatically uses an out-of-the-box data source, `autoaccess-ds`, that accesses the customer's data lake within the Advanced Identity Cloud tenant's cloud storage data for ML training. You do not need to define any data sources in this case.

The out-of-the-box data source also does not require attribute mapping. You simply define the source on the Autonomous Access Advanced Identity Cloud UI (refer to Run training) when setting up your training run.



Data Sources

Define mappings to auto access features

Name	Location	Created	Updated	Status
autoaccess-ds	fr-vc43bqrkbuw4qpzbov8zqw6g4ta-auto-access-output/*	May 17, 2022 6:18 PM	May 17, 2022 6:18 PM	Active

© 2022 ForgeRock, Inc Build 331ed99 (May 17, 2022, 18:44)

If you do not have the `autoaccess-ds` data source, contact ForgeRock for assistance.

NOTE

NOTE

Autonomous Access uses cloud storage data for its ML training runs and Elasticsearch data for its heuristic predictions. The Autonomous Access result journey node collects data and stores it in cloud storage and stores risky events in Elasticsearch.

The general guidelines for customer data storage is as follows:

- **Three months of access logs**. The Autonomous Access activity dashboard displays the anomalous accesses that occurred over the past three months. As a result, Autonomous Access requires *three* months or more of customer data in Elasticsearch for optimal results.
- **Google Cloud storage**. Autonomous Access requires *six* months of customer data for optimal AI/ML analytics results.
- **Secure data**. All customer data resides within each customer's private tenant and cannot be accessed externally to the tenant.

In general, you may have to define data sources for the following cases:

- **Demos**. Customers who want to do a POC with their own data but are not current Advanced Identity Cloud customers may need to set up and define their own data sources. Typically, we load *three* months of access logs into a GCP data bucket and point to that location within the data source configuration presented below.

Verify the default data source

1. On the Autonomous Access UI, click **Risk Administration** > **Data Sources**.
2. Verify that the `autoaccess-ds` is present and activated.

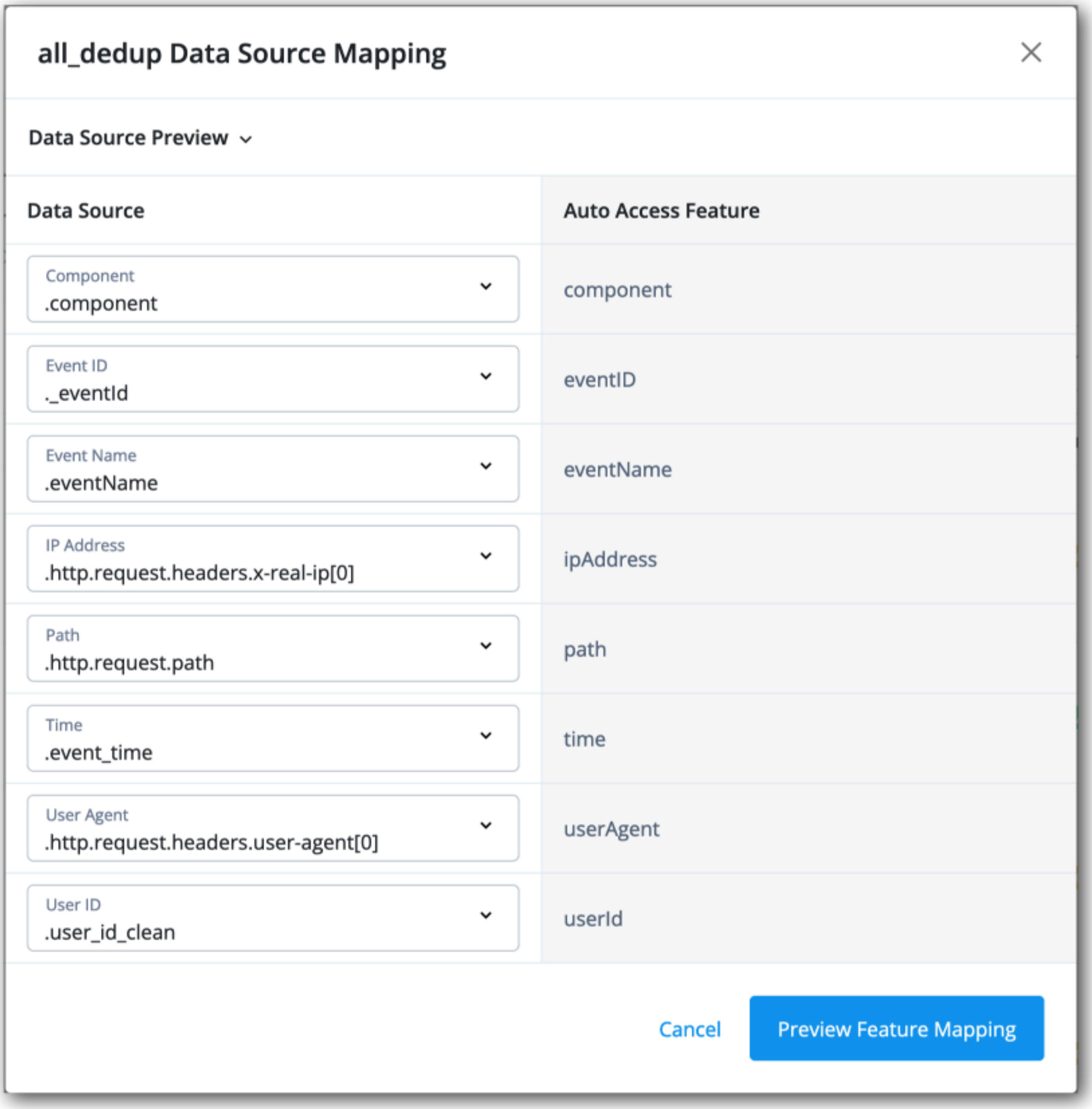
You do not have to set the mapping as it is configured already. Next, set the Risk Configuration.

Set the data source (if not using the default data source)

1. On the Autonomous Access UI, click **Risk Administration** > **Data Sources**.
2. On the Data Sources page, click **Add Data Source**.
3. On the Add Data Source dialog, select the data bucket in the **Bucket Search** field.
4. For Object Prefix, click **Define from Prefix**, and enter the following:
    a. **Name (of the Data Source)**. Add a descriptive name for the data source.
    b. **Bucket Name**. Add the data bucket for the data source.
    c. **Prefix**. Add a prefix.

5. Click **Save**. The new data source is displayed on the page. The Status column displays the current state of the data source.
6. At this stage, you need to set attribute mapping between your data source and the schema. Click the trailing dots, and select **Create or Edit Mapping**.
7. Under Data Source, select the attribute to map to the Auto Access feature. Repeat for as many attributes as you can. Note that you cannot add attributes to the list of attributes.

▼ *Display an example of a data source mapping.*



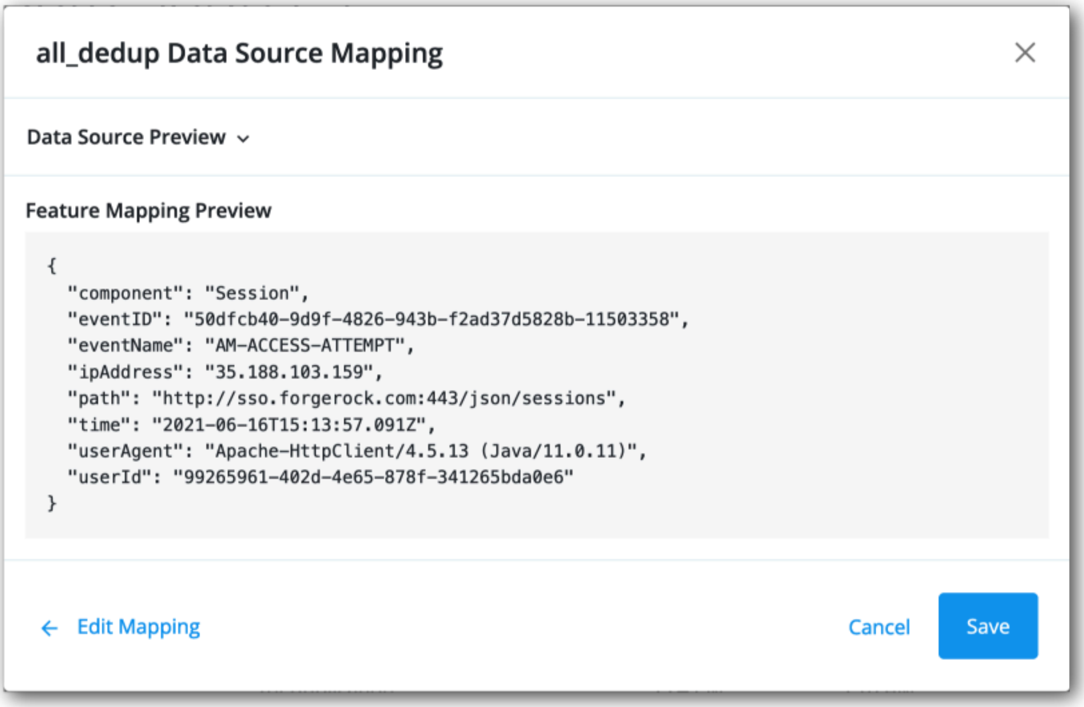
all_dedup Data Source Mapping

Data Source Preview

Data Source	Auto Access Feature
Component .component	component
Event ID ._eventId	eventID
Event Name .eventName	eventName
IP Address .http.request.headers.x-real-ip[0]	ipAddress
Path .http.request.path	path
Time .event_time	time
User Agent .http.request.headers.user-agent[0]	userAgent
User ID .user_id_clean	userId

8. Click **Preview Feature Mapping** to review the mapping, and then click **Save** if satisfied.

▼ *Display an example of a data source mapping preview.*



```
{
  "component": "Session",
  "eventID": "50dfcb40-9d9f-4826-943b-f2ad37d5828b-11503358",
  "eventName": "AM-ACCESS-ATTEMPT",
  "ipAddress": "35.188.103.159",
  "path": "http://sso.forgerock.com:443/json/sessions",
  "time": "2021-06-16T15:13:57.091Z",
  "userAgent": "Apache-HttpClient/4.5.13 (Java/11.0.11)",
  "userId": "99265961-402d-4e65-878f-341265bda0e6"
}
```

9. The data source will be in an `Inactive` state, you must now activate the data source to use it in the training run. Click the trailing dots, and select `Activate`. The data source is now in an active state.

   You have successfully set up or confirmed your data source(s) and mappings. Next, you can set the Risk Configuration.

Configure the risk settings

Customers in different industry verticals require varying risk policies and heuristics in their applications. Ping Identity designed Autonomous Access's risk configuration with this in mind.

To enable easy configuration of its risk parameters, Autonomous Access stores its risk configuration settings in a YAML-based file. Users who have the `Data_Analyst` role can modify the parameters to set how risk is evaluated and how the response Autonomous Access sends data back to its node. The Autonomous Access server polls the configuration file every ten minutes (default) for changes to the file.

IMPORTANT

Misconfiguration of the file can result in an inoperable service. ForgeRock Professional Services group can assist in the risk configuration process.

Grant Autonomous Access roles

Ping Identity Platform provides two roles that let users use different aspects of Autonomous Access:

- **Fraud analyst.** The fraud analyst role lets users access the Risk dashboard, but they cannot run any training or configuration.
- **Data analyst.** The data analyst role lets users access the risk administration menu options on Advanced Identity Cloud admin UI, but they cannot access the Risk dashboard.

> NOTE
>
> Administrators are automatically granted both roles and can assign these roles to other users.

Setting Autonomous Access roles

1. In the Advanced Identity Cloud admin UI, go to **Identities** > **Manage**. Search for the user(s) to whom you want to assign the roles.
2. In the Advanced Identity Cloud admin UI, go to **Native Consoles** > **Access Management**.
3. Click **Identities**, and then click the **Groups** tab.
4. If the `Data_Analyst` and `Fraud_Analyst` groups are not present, you will need to add them as follows:
    a. Click **Add Group**.
    b. In the **Group ID** field, enter `Data_Analyst` , and then click **Create**.
    c. Repeat the previous step and create a `Fraud_Analyst` group. Both groups appear in the `Groups` tab.
5. Go to **Identities** > **Manage**. Select a user to whom you want to add the role.
6. Click the **Groups** tab, and in the **Name** field, select `Data_Analyst` or `Fraud Analyst` .
7. Repeat again to add the roles to other users.

Risk configuration

The risk configuration page provides an extremely extensible and performant server configuration giving the end user full control of their Autonomous Access system.

Edit the risk configuration

1. On the Autonomous Access UI, go to **Risk Administration** > **Risk Config**.
2. Edit any of the threshold parameters for each heuristic if necessary.

```
version: "1.1"
bruteForce:
  BRUTE_FORCE_WINDOW_MS: 300000
  BRUTE_FORCE_COUNT_THRESHOLD: 20
  BRUTE_FORCE_RISK_SCORE: 100
credentialStuffing:
  CREDENTIAL_STUFFING_WINDOW_MS: 600000
  CREDENTIAL_STUFFING_COUNT_THRESHOLD: 5
  CREDENTIAL_STUFFING_RISK_SCORE: 100
impossibleTravel:
  IMPOSSIBLE_TRAVEL_SPEED_CUTOFF_MPH: 700
  IMPOSSIBLE_TRAVEL_RISK_SCORE: 100
suspiciousIp:
  SUSPICIOUS_IP_WINDOW_MS: 300000
  SUSPICIOUS_IP_COUNT_THRESHOLD: 10
  SUSPICIOUS_IP_RISK_SCORE: 100
uebaConfig:
  RISK_SCORE_RATIO: 0.25
  RISK_SCORE_CENTER_SIGMA: 50
  RISK_SCORE_BASELINE_THRESSHOLD_SIGMA: 6
  USER_COUNT_CUTOFF_FOR_SCORE: 20
userAgentRule:
  USER_AGENT_RULE_RISK_SCORE: 100
doubleJeoPardy:
  MFA_TIMEOUT: 60
heuristicsConfig:
  HEURISTIC_RISK_SCORE_COMPUTE_STRATEGY: max
processConfig:
  risk_score_threshold: 50
  UEBA_AGGREGATION_STRATEGY: max
  HEURISTIC_AGGREGATION_STRATEGY: max
  RISK_PROCESS_TIMEOUT: 950
distributed_attack_heuristic:
  DISTRIBUTED_ATTACK_WINDOW_MS: 600000
  DISTRIBUTED_ATTACK_COUNT_THRESHHOLD: 7
  DISTRIBUTED_ATTACKC_RISK_SCORE: 100
block_and_allow_list:
  BLOCK_LIST: []
  ALLOW_LIST: []
```

The properties are defined as follows:

▼ *Risk configuration properties*

Brute force	BRUTE_FORCE_WINDOW_MS	Number of milliseconds back in time to look for brute force events. Default: 300000. Minimum: 0.
	BRUTE_FORCE_COUNT_THRESHOLD	Number of events required to trigger the rule. Default: 20. Minimum: 1.
	BRUTE_FORCE_RISK_SCORE	Brute force risk score returned if the rule is tripped. Default: 100. Minimum: 1.
Credential stuffing	CREDENTIAL_STUFFING_WINDOW_MS	Number of milliseconds back in time to look for credential stuffing events. Default: 600000. Minimum: 0.
	CREDENTIAL_STUFFING_COUNT_THRESHOLD	Number of events required to trigger the rule. Default: 5. Minimum: 1.
	CREDENTIAL_STUFFING_RISK_SCORE	Credential stuffing risk score returned if the rule is tripped. Default: 100. Minimum: 1; maximum: 100.
Impossible travel	IMPOSSIBLE_TRAVEL_SPEED_CUTOFF_MPH	Speed in miles per hour above which the events will be flagged. Default: 700 mph. Minimum: 1; maximum: 5000.
	IMPOSSIBLE_TRAVEL_RISK_SCORE	Impossible travel risk score returned if the rule is tripped. Default: 100. Minimum: 1; maximum: 100.
Suspicious IP	SUSPICIOUS_IP_WINDOW_MS	Number of milliseconds back in time to look for suspicious IP events. Default: 300000. Minimum: 60 maximum: 480.
	SUSPICIOUS_IP_COUNT_THRESHOLD	Number of events required to trigger the rule. Default: 10. Minimum: 1.

	SUSPICIOUS_IP_RISK_SCORE	Suscipious risk score returned if the rule is tripped. Default: 100.
UEBA config	RISK_SCORE_RATIO	Used for the calculation of the risk score. Default: 0.25. Minimum: 0; maximum: 1.
	RISK_SCORE_CENTER_SIGMA	Used for the calculation of the risk score. Default: 50. Minimum: 1.
	RISK_SCORE_BASELINE_THRESHOLD_SIGMA	Used for the calculation of the risk score. Default: 6. Minimum: 1
User agent rule	USER_COUNT_CUTOFF_FOR_SCORE	Used for the calculation of the risk score. Default: 20. Minimum: 10
	USER_AGENT_RULE_RISK_SCORE	User agent rule risk score when an incoming user agent string is matched against a known pattern of botnet user agents. Default: 100. Minimum: 1; maximum: 100.
Double jeopardy	MFA_TIMEOUT	Multifactor authentication timeout in minutes if double jeopardy is enabled. Default: 60.

Heuristics config	HEURISTIC_RISK_SCORE_COMPUTE_STRATEGY	Type of strategy to determine the heuristic risk score. Options are: - **max**: Final risk score is the maximum heuristic risk score assigned to the user. Default. - **avg**: Final risk score is the average of all heuristic risk scores assigned to a user. - **softmax**: Final risk score is measured as a percentage and obtained from the softmax function. - **sum_floor_to_hundred**: Final risk score is capped at 100.
Process config	RISK_SCORE_THRESHOLD	Risk score threshold for UEBA and heuristics, displayed on the risk dashboard. Default: 50. Minimum: 1; maximum: 100.

	UEBA_AGGREGATION_STRATEGY	Type of aggregation strategy for UEBA signals: ◦ **max**: Final risk score is the maximum UEBA risk score assigned to the user. Default. ◦ **avg**: Final risk score is the average of all UEBA risk scores assigned to a user. ◦ **softmax**: Final risk score is measured as a percentage and obtained from the softmax function. ◦ **sum_floor_to_hundred**: Final risk score is capped at 100.
	HEURISTIC_AGGREGATION_STRATEGY	Not needed.
	RISK_PROCESS_TIMEOUT	Max timeout in milliseconds of the risk processing time. Once the timeout occurs, risk scores are no longer processed. Default: 950ms. Min: 1ms; Max: 1000 ms.

Block and Allow list	BLOCK_LIST: [ ]	Overrides the risk score of IP addresses in the *block list* with a value of 100. You must also enable the `Allow/Block List` heuristic on your journey's Signal node. ◦ Support IPv4 and IPv6 formats. ◦ Specify single IP addresses, like 10.0.48.0, or IP subnets, like 10.0.48.0/24. ◦ Cannot use regular expressions or wildcards with the IP addresses. ◦ Subjects IPs on the *block list* to heuristics and machine learning and overrides their computed risk score with a score of 100.

	ALLOW_LIST: [ ]	Overrides the risk score of IP addresses in the *allow list* with a value of 0. Allow list IP addresses bypass heuristics and machine learning analytics. You must also enable the `Allow/Block List` heuristic on your journey's Signal node. ◦ Support for IPv4 and IPv6 formats. ◦ Specify single IP addresses, like 10.0.48.0, or IP subnets, like 10.0.48.0/24. ◦ Cannot use regular expressions or wildcards with the IP addresses. ◦ Assigns a risk score of 0 to any IPs on the *allow list* and excludes them from heuristics and machine learning.
distributed_attack_heuristic:	DISTRIBUTED_ATTACK_WINDOW_MS	Number of milliseconds back in time to look for distributed attack events. Default: 60000. Minimum: 0.
	DISTRIBUTED_ATTACK_COUNT_THRESHOLD	Number of events required to trigger the rule. Default: 7.
	DISTRIBUTED_ATTACK_RISK_SCORE	Risk score returned if the rule is tripped. Default: 100. Minimum: 1.

3. After you have made your changes to the file, click **Save**. The Preview Risk Evaluation popup window appears.
4. On the Preview Risk Evaluation popup window, do the following:
   a. Click **Bucket Search** to select your data source location or type the name of the data source location.
   b. Optional. Enter an object prefix to filter your search results.

c. Next to your desired object, click the trailing dots, and then click **Preview Object** to display your data source change(s).

d. Click **Preview Risk Evaluation** to review a simulated risk evaluation for the first event.

e. If you are satisfied with your change(s), click **Save Config**.

Run training

The Training job is the first part of a multi-step process that automates machine learning workflows to generate the machine learning (ML) models. An ML model is an algorithm that learns patterns and relationships from data, enabling it to make predictions or perform tasks without explicit programming instructions.

Initially, Autonomous Access's training job operates without a model, focusing on processing jobs to gather and correlate unstructured and unlabelled data into structured input data. Subsequently, the training workflow utilizes this input data and heuristics to generate machine learning models, iteratively refining its processing to enhance model accuracy.

IMPORTANT

Whenever you add new users after a training run, you need to create a new training job and run training again to update the model. You will need atleast twenty events for a new user to successfully be part of the UEBA and heuristics learning process.

Once the training job has completed, you must tune the models for greater accuracy and performance. Once the models are tuned, you must rerun the training job, and then publish the model, saving it for later use in production.

Finally, with the new training model and rules, run a predictions job on the historical data.

NOTE

ForgeRock encourages customers to create and run the training pipelines and evaluate the models for accuracy. Administrators can run training on a periodic basis (e.g., bi-weekly or bi-monthly) and as soon as the Autonomous Access journeys begin to collect data.

Run training

Using the default data source, `autoaccess-ds`, run the training job on the UI.

The general guidance as to *when* you can run your first training model is as follows:

- **Six months of customer data**. For optimal results, Autonomous Access requires six months of customer data, or data with 1000 or more events. Data collections with less than 1000 events will not yield good ML results.

  > NOTE
  >
  > The training job takes time to process as it iteratively runs the machine learning workflows multiple times.

Run the Training job

1. On the Autonomous Access UI, go to **Risk Administration > Pipelines**.
2. Click **Add Pipeline**.
3. On the Add Pipeline dialog box, enter the following information:
   a. **Name**. Enter a descriptive name for the training job.

   b. **Data Source**. Select the data source to use for the job.

   c. **Select Type**. Select **Training**. The dialog opens with model settings that you can change if you understand machine learning.

      - **Model A**. Model A is a neural-network module that is used to learn the optimal coding of unlabeled data. You can configure the following:
        - **Batch size**. The batch size of a dataset in MB.
        - **Epochs**. An epoch is the number of iterations the ML algorithm has completed during its training on the entire dataset.
        - **Learning rate**. The learning rate is a parameter that determines how much to tune the model as model weights are adjusted. Thus, the learning rate adjusts the step size of each iterative training pass.
      - **Model B**. Model B is a neural-network module that returns good data points that helps with predictive training. You can configure the following:
        - **Batch size**. The batch size of a dataset in MB. The smaller the batch size, the longer the training will run, but the results may be better.
        - **Epochs**. An epoch is the number of iterations the ML algorithm has completed during its training on the entire dataset.
        - **Learning rate**. The learning rate is a parameter that determines how much to tune the model as model weights are adjusted. Thus, the learning rate adjusts the step size of each iterative training pass.
      - **Model C**. Model C is a module that aggregates and groups data points.
        - **Max number_of_clusters**. Maximum number of clusters for the dataset.
        - **Min number_of_clusters**. Minimum number of clusters for the dataset.

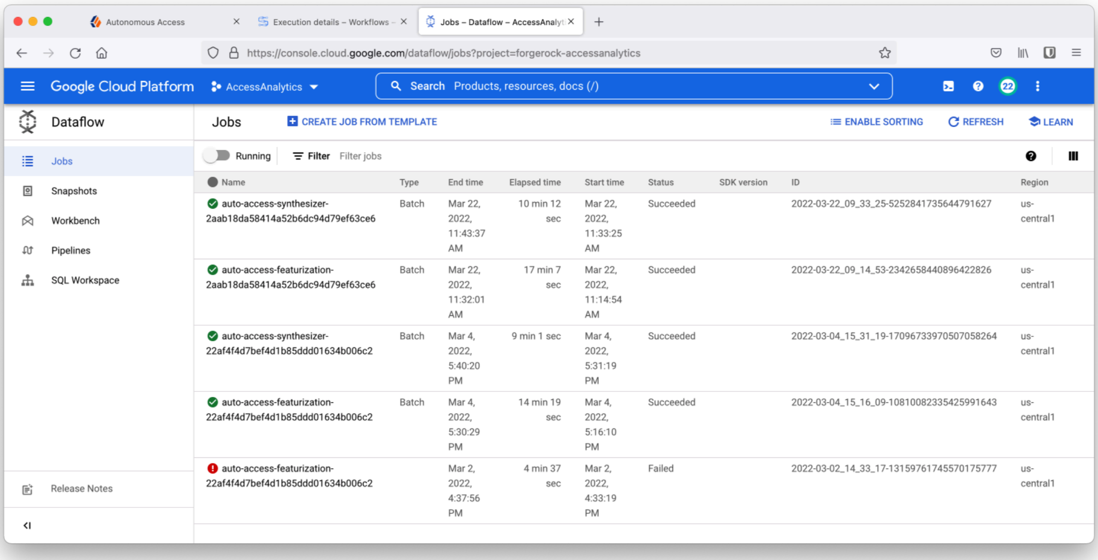
    - **Embeddings**. The Embeddings module is a meta-model and is used to convert text fields into numbers that the ML models can understand. Autonomous Access trains the embedding model, and then trains the other models on top of the Embeddings model.
        - **Embedding dimension**. The embeddings module takes actual raw text data and encodes them into tokens or numbers. The embedding dimension determines how many numbers/tokens to use in the encoding. For the purposes of Autonomous Access, 20 is the default number and is sufficient for most Autonomous Access applications.
        - **Learning rate**. The learning rate is a parameter that determines how much to tune the model as model weights are adjusted. Thus, the learning rate adjusts the step size of each iterative training pass.
        - **Window**. Indicates how far ahead and behind to look for the data. The minimum number is 1. The default number is 5.

4. Click **Save**.

5. Click the trailing dots, click **Run Pipeline**, and then click **Run**. Depending on the size of your data source and how you configured your job settings, the training run will take time to process.

6. Click **View on GCP** to display the detailed processing of the pipeline. Take note of the workflow name, you can use it to monitor the logs during the training run.

> NOTE
>
> Another useful GCP page is to view the individual training sub-jobs on the Dataflow page.



Name	Type	End time	Elapsed time	Start time	Status	SDK version	ID	Region
✓ auto-access-synthesizer-2aab18da58414a52b6dc94d79ef63ce6	Batch	Mar 22, 2022, 11:43:37 AM	10 min 12 sec	Mar 22, 2022, 11:33:25 AM	Succeeded		2022-03-22_09_33_25-5252841735644791627	us-central1
✓ auto-access-featurization-2aab18da58414a52b6dc94d79ef63ce6	Batch	Mar 22, 2022, 11:32:01 AM	17 min 7 sec	Mar 22, 2022, 11:14:54 AM	Succeeded		2022-03-22_09_14_53-2342658440896422826	us-central1
✓ auto-access-synthesizer-22af4f4d7bef4d1b85ddd01634b006c2	Batch	Mar 4, 2022, 5:40:20 PM	9 min 1 sec	Mar 4, 2022, 5:31:19 PM	Succeeded		2022-03-04_15_31_19-17096733970507058264	us-central1
✓ auto-access-featurization-22af4f4d7bef4d1b85ddd01634b006c2	Batch	Mar 4, 2022, 5:30:29 PM	14 min 19 sec	Mar 4, 2022, 5:16:10 PM	Succeeded		2022-03-04_15_16_09-10810082335425991643	us-central1
! auto-access-featurization-22af4f4d7bef4d1b85ddd01634b006c2		Mar 2, 2022, 4:37:56 PM	4 min 37 sec	Mar 2, 2022, 4:33:19 PM	Failed		2022-03-02_14_33_17-13159761745570175777	us-central1

7. Click **OK** to close the dialog box.

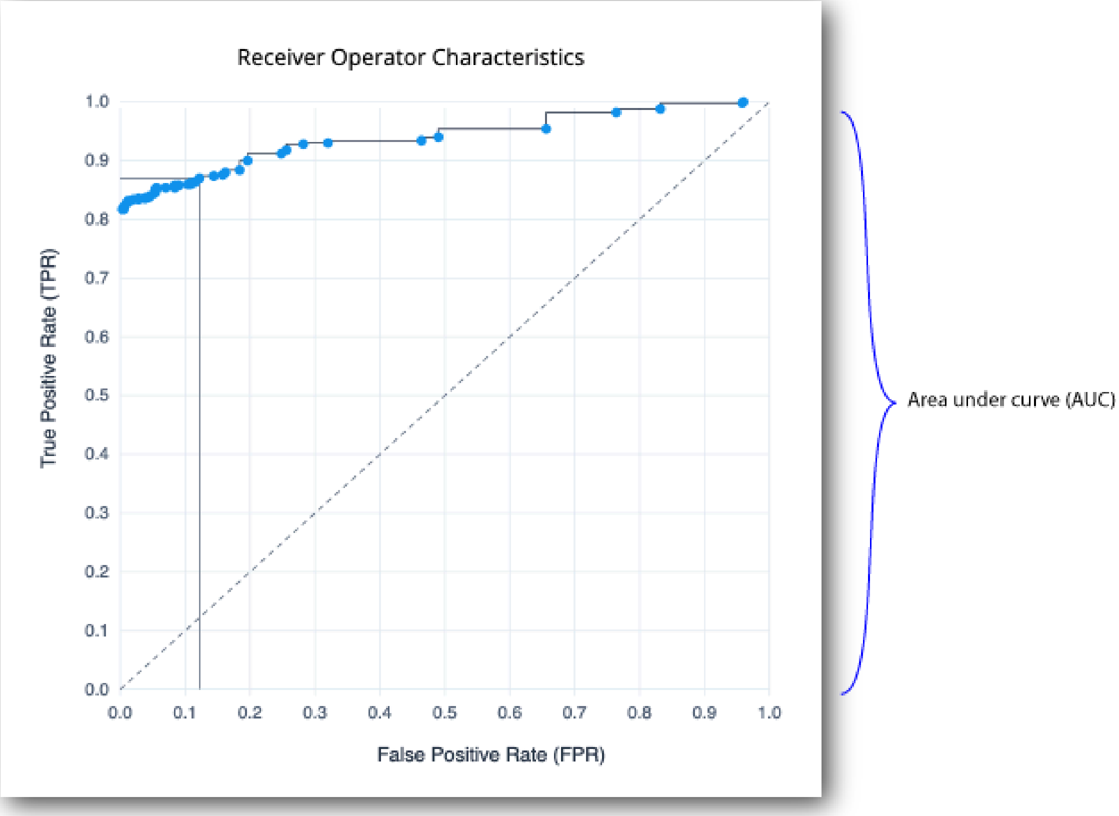
8. Upon a successful run, a `Succeeded` status message appears.

Tune the model

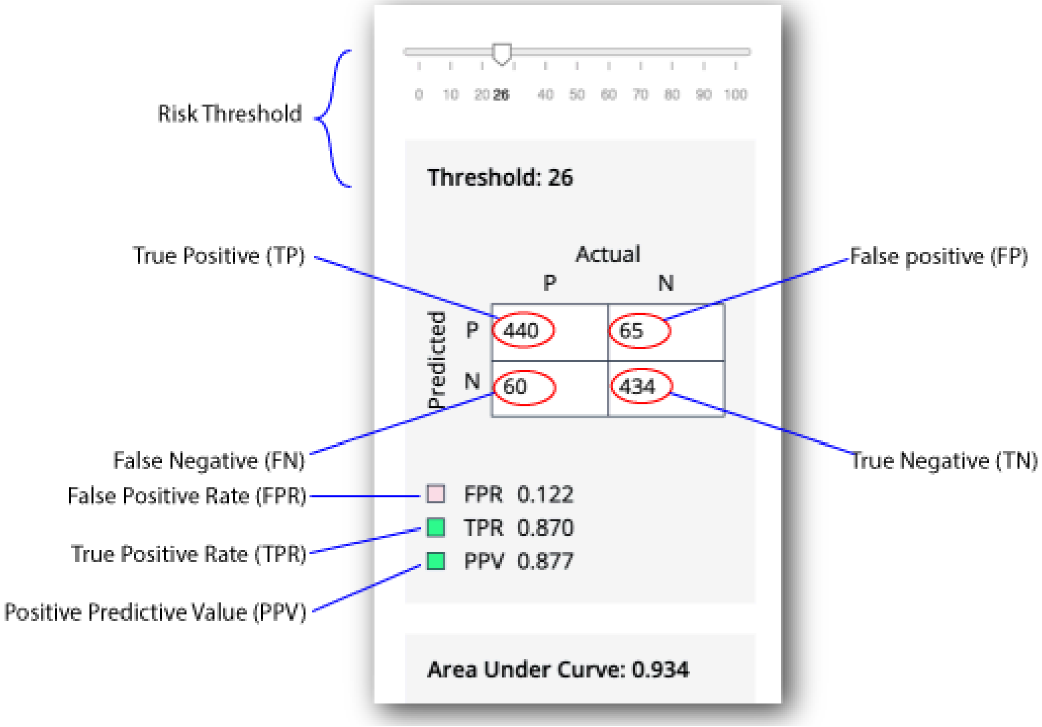
Training model terms

The following training model terms are presented to better understand the tuning models:

- **Receiver Operator Characteristics**. A *receiver operator characteristic (ROC)* curve is graphical plot that shows the tradeoffs between true positives and false positives of a model as the risk threshold changes. The x-axis shows the false positive rate (FPR), the probability of a false alarm; the y-axis shows the true positive rate (TPR), probability of correct detection. The diagonal line is a random classifier, which serves as a baseline to evaluate the performance of a classifier. The points above the diagonal represent good classification results; points below the diagonal represent bad results. The ROC curve with blue points is a probability curve, which represents the varying trade-offs between the true positive rate and false positive rate at different probability thresholds. The ideal representation is called a *perfect classification*, where (0,1) indicate no false negatives and 100% true positives, and where the graph looks like an upside-down "L". The Area under the curve (AUC) represents how well the model can distinguish a true positive and a false negative. The higher the AUC value (closer to 1), the better the model is at distinguishing between a risky threat and no threat.



- **Confusion Matrix**. A *confusion matrix* is an 2x2 (also called a binary classification) table that aggregates the ML model's correct and incorrect predictions. The horizontal axis is the actual results; the vertical axis is the predicted results. Note that each prediction in a confusion matrix represents one point on the ROC curve.



Threshold: 26

		Actual	
		P	N
Predicted	P	440	65
	N	60	434

□ FPR 0.122
🟩 TPR 0.870
🟩 PPV 0.877

Area Under Curve: 0.934

- **True Positive**. A *true positive* is an outcome where the model correctly identifies an actual risky threat as a risky threat.
- **True Negative**. A *true negative* is an outcome where the model correctly identifies a non-risky threat as a non-risky threat.
- **False positive**. A *false positive* is an outcome where the model incorrectly identifies a non-risky threat as a risky threat.

> IMPORTANT
>
> Customers care a lot about false positives, because it affects the user’s runtime experience. Also, note that false positives is the rate that we are correctly flagging *anomalies*, not real fraud.

- **False negative**. A *false negative* is an outcome where the model incorrectly identifies a risky threat as a non-risky threat.
- **True Positive Rate (TPR)**. The rate of probability of detection, where TP/(TP+FN), where TP is the number of true positives and FN is the number of false negatives. The rate is the probability that a positive threat is predicted when the actual result is positive. The TPR is also known as *recall*.
- **False Positive Rate (FPR)**. The rate of the probability of a false alarm, where FP/(FP+TN), where FP is the number of false positives and TN is the number of true negatives. "FP+TN" is the total number of negatives. The rate is the probability that a false alarm will be raised, where a positive threat is predicted when the actual result is negative.

- **Positive Predictive Value (PPV)**. The rate where TP/(TP+FP), where TP is the number of true positives and FP is the number of false positives. "TP+FP" is the total number of positives. The ideal value of the PPV with a perfect prediction is 1 (100%), and the worst is zero. The rate is the probability that a predicted positive is a true positive. The PPV is also known as *precision*.
- **Area under the curve**. On a scale from 0 to 1, the area under the curve (AUC) shows how well the model distinguishes between a true positive and false negative. If the AUC is closer to 1, the better the model.
- **Centroids**. On model C, the centroids represents a typical user and describes the profiles of known users.
- **Precision Recall**. The *precision recall* curve is a graphical plot that shows the tradeoff between precision and recall for different threshold values. The Y-axis plots the positive predictive value (PPV), or precision; the x-axis plots the true positive rate (TPR), or recall. A model with ideal results are at the point (1,1). The precision recall is a different way to view the model’s performance. Most users can use the ROC and confusion matrix.

NOTE

The Ensemble model is the best average of the three models and should show better results than each individual model. The model C chart is a choppier step graph.

Tuning Training

Autonomous Access supports the ability to tune the AI/ML training models for greater accuracy. There are three things that you can check to look at ML model performance:

- **Training logs**. Each model generates a metadata file that show the `train_losses` and `val_losses` (validation losses). `Train loss` indicates how well the model fits the training data. `Validation loss` indicates how well the model fits new training data. The losses tend to move down from a high value (near 1) to a low value (toward 0).

  Another thing to look at is how good are the train losses to the validation losses. The bigger the gap between the two loss numbers is an indicator that the model memorizes the training set but does not generalize very well. One thing you can do to improve this value is increase the number of epochs and make the learning rate smaller, which adjusts the model closer to data. For the embedding model, you can increase the window size and decrease the learning rate to improve results.

NOTE

NOTE

> If you are seeing a model with a big gap between the training loss and validation loss, it could mean you have too many parameters, meaning that an overfitting of data is taking place. One possible solution is to reduce the embedding dimension. Again, ask the ForgeRock for assistance.

- **ROC curve and confusion matrix**. You can view the receiver operator characteristics (ROC) curve and confusion matrix on the UI.
- **Risk Configuration**. The Autonomous Access lets you fine-tune the backend AAI server. For more details, refer to Configure the risk settings.

Tune the training models

1. On the Pipelines page, click the dots next to a training run, and then click **View Logs**.
2. Click the dots next to a training run, and then click **View Run Details**.
3. On the Training Execution Details, click the dots, and then click **Results**. The training results are displayed.
4. Tune each model by adjusting the threshold. You can select the model on the drop-down list, and adjust the threshold to view the optimal balance of parameters in the confusion matrix on the Decision node. You can view the graphs of the following models:
   - **Ensemble**. Displays the best average of all charts in one view.
   - **Model A**. Displays the Model A charts.
   - **Model B**. Displays the Model B charts.
   - **Model C** Displays the Model C charts.

   NOTE

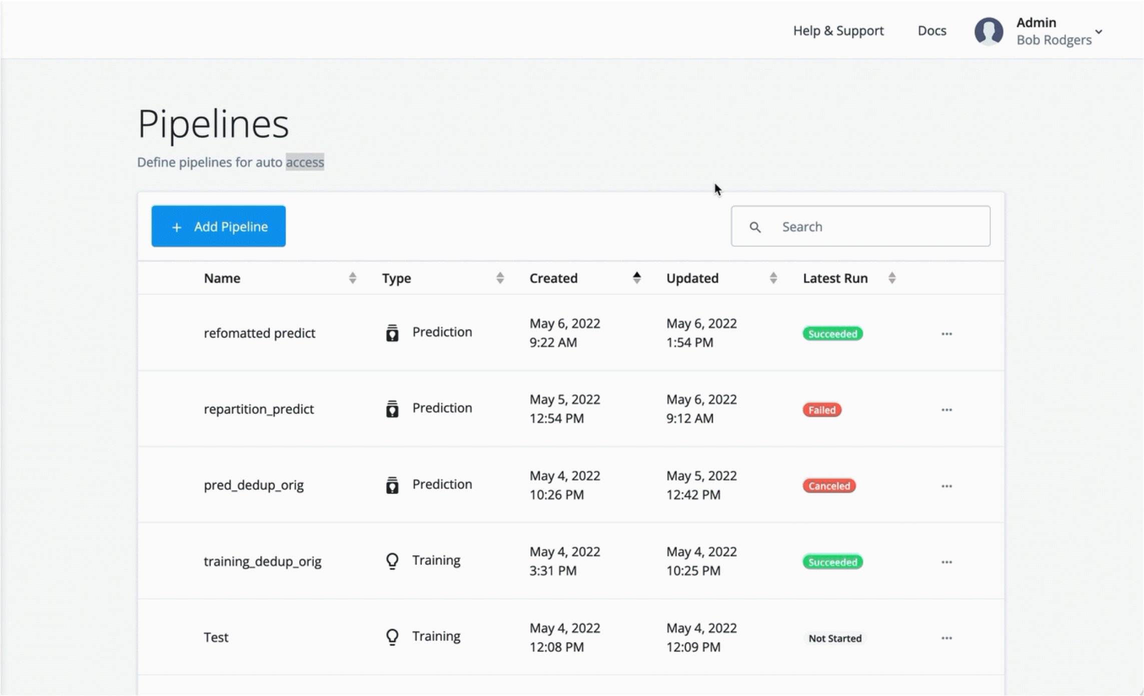
   > ForgeRock's data science experts can assist you in this process.

5. To close the dialog, click **OK**.
6. Finally, if you are satisfied with the models' performance, click **Publish** to save the training model. Once published, you can only overwrite it with another training run.

   NOTE

   > As a general rule of thumb, ForgeRock recommends to publish the models with an AOC > 0.9. For more information, contact your ForgeRock representative.

   ▼ *Display an example*



Pipelines

Define pipelines for auto access

Name	Type	Created	Updated	Latest Run
refomatted predict	Prediction	May 6, 2022 9:22 AM	May 6, 2022 1:54 PM	Succeeded
repartition_predict	Prediction	May 5, 2022 12:54 PM	May 6, 2022 9:12 AM	Failed
pred_dedup_orig	Prediction	May 4, 2022 10:26 PM	May 5, 2022 12:42 PM	Canceled
training_dedup_orig	Training	May 4, 2022 3:31 PM	May 4, 2022 10:25 PM	Succeeded
Test	Training	May 4, 2022 12:08 PM	May 4, 2022 12:09 PM	Not Started

Learn about the Autonomous Access nodes

This section provides a description of each Autonomous Access node you can use within your journeys.

The Autonomous Access nodes do not require any custom coding or connectors to implement within a journey.

> NOTE
>
> When you purchase the Autonomous Access add-on capability, the Autonomous Access nodes appear automatically in the Advanced Identity Cloud **Journeys** section under **Autonomous Access**.

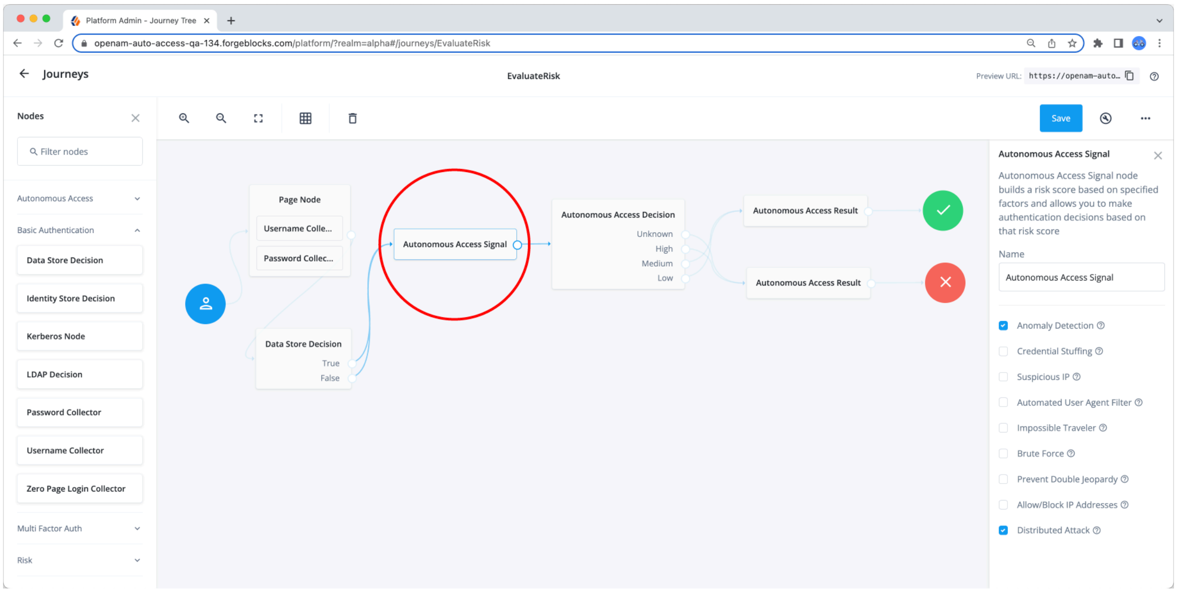
Signal node

The Autonomous Access Signal node is a construction node where you specify the heuristics and/or anomaly detection to be included in risk score generation during the AI/ML pipelines. By default, all heuristics and anomaly detection are enabled, but you can specify multiple combinations depending on the use case.

The input typically comes from the Data Store Decision node, but can come from other similar nodes where Autonomous Access can obtain a user ID (for example, Platform Username node).

> NOTE
>
> The output must connect to the Autonomous Access Decision node for actionable paths.

The signal node creates a transaction ID and sends an API call to the Autonomous Access server for information. In response, the Autonomous Access returns the risk scores and additional information associated with the transaction ID. The output connects to the Autonomous Access Decision node or a Scripted Decision node for some actionable paths.



Autonomous Access Signal

Autonomous Access Signal node builds a risk score based on specified factors and allows you to make authentication decisions based on that risk score

Name: Autonomous Access Signal

☑ Anomaly Detection
□ Credential Stuffing
□ Suspicious IP
□ Automated User Agent Filter
□ Impossible Traveler
□ Brute Force
□ Prevent Double Jeopardy
□ Allow/Block IP Addresses
☑ Distributed Attack

Table 1. Signal node configuration

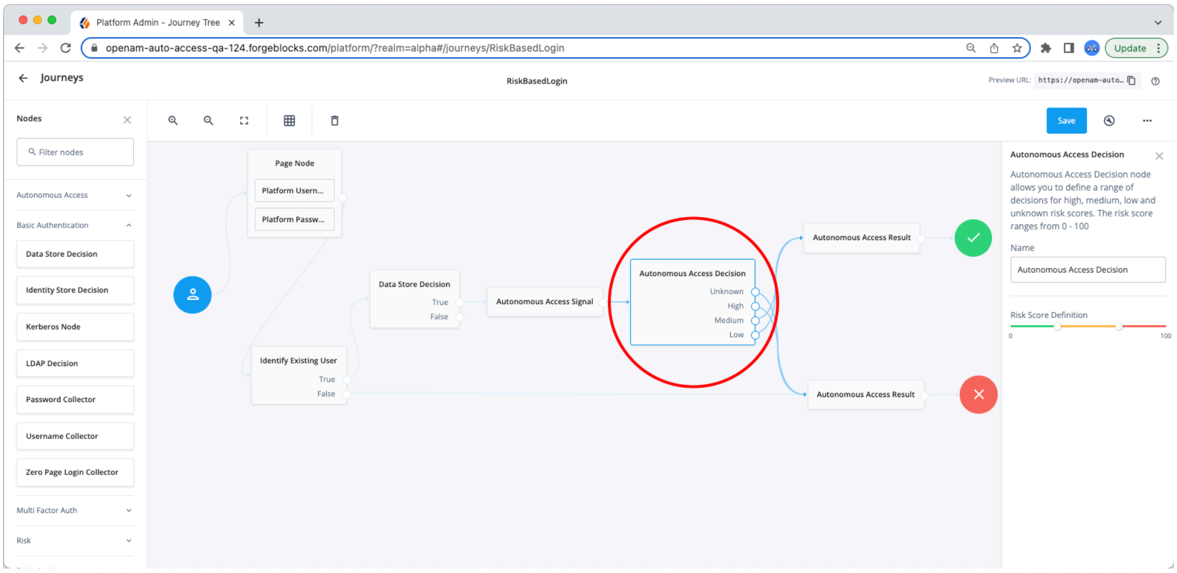
Heuristic	Usage
Anomaly Detection	Detects behavioral anomalies based on user city, country, day of week, time of day, operating system (OS), OS version, device, device type, and browser type.
Credential Stuffing	Detects if an IP address is trying to access a number of users over a period of time.
Suspicious IP	Detects if an IP is making many authentication attempts.
Automated User Agent Filter	Detects automated bots in the user-agent string.
Impossible Traveler	Detects if a user is moving between two locations at an impossible speed.
Brute Force	Detects direct users failing multiple authentication attempts.

Heuristic	Usage
Prevent Double Jeopardy	Ensures that if a user has been flagged as anomalous due to some explainable risk by the machine learning system and subsequently passed a multifactor authentication (MFA), the user will not be asked again for multifactor authentication for the same explainable risk within the preconfigured time window in the Risk Configuration UI. For example, if a user in New York flies to Toronto and authenticates to the system, the system flags the user's authentication event as anomalous ("Unusual city"). The system then issues a multifactor authentication challenge. If the user passes the challenge, the user will not be flagged again for the same explainable risk for any subsequent logins in Toronto as long as it is within the preconfigured time window of 60 minutes (default). You must also check the MFA Enabled feature on the Result node whenever multifactor authentication is configured. Refer to Result node.
Allow/Block IP Addresses	Validates if an incoming request event and its IP address is in the Allow/Block IP address lists as defined in the Rick Configuration UI. • If the IP address is on the Allow IP address list, Autonomous Access assigns a risk score of 0, passes the IP through and excludes it from heuristics and machine learning. • If the IP address is on the Block IP address list, Autonomous Access runs heuristics and machine learning on the IP, computes a risk score, and then overrides it by assigning it a score of 100, indicating high risk.
Distributed Attack	Flags an event as a distributed attack when there are more than the preconfigured threshold of unique IP addresses attempting to log on using the same user ID in the time window specified in the configuration.

Decision node

The Decision node takes the data sent by the Signal node and directs the flow to actionable paths depending on where the risk score falls within the range of high, medium, low, and unknown scores. The full range of scores is from 0 to 100, where 0 indicates no risk and 100 indicates the highest risk. You can adjust the dual range threshold sliders to set the ranges.

The node takes its input from the signal node and outputs to a corresponding path depending on the journey's configuration.



Autonomous Access Decision

Autonomous Access Decision node allows you to define a range of decisions for high, medium, low and unknown risk scores. The risk score ranges from 0 - 100

Name: Autonomous Access Decision

NOTE

If you want more granularity for the risk score ranges, use the Scripted Decision node in place of the out-of-the-box decision node. Contact Ping Identity and refer to Scripted decision node API.

The following list describes the decision node configuration for each heuristic:

Table 2. Decision node configuration

Property	Usage
Unknown	*Unknown* risk scores occur when a risk score could not be calculated during the AI/ML job runs for the following reasons: • There are not enough data points for the AI/ML analytics. • The service is down. • There is a timeout.

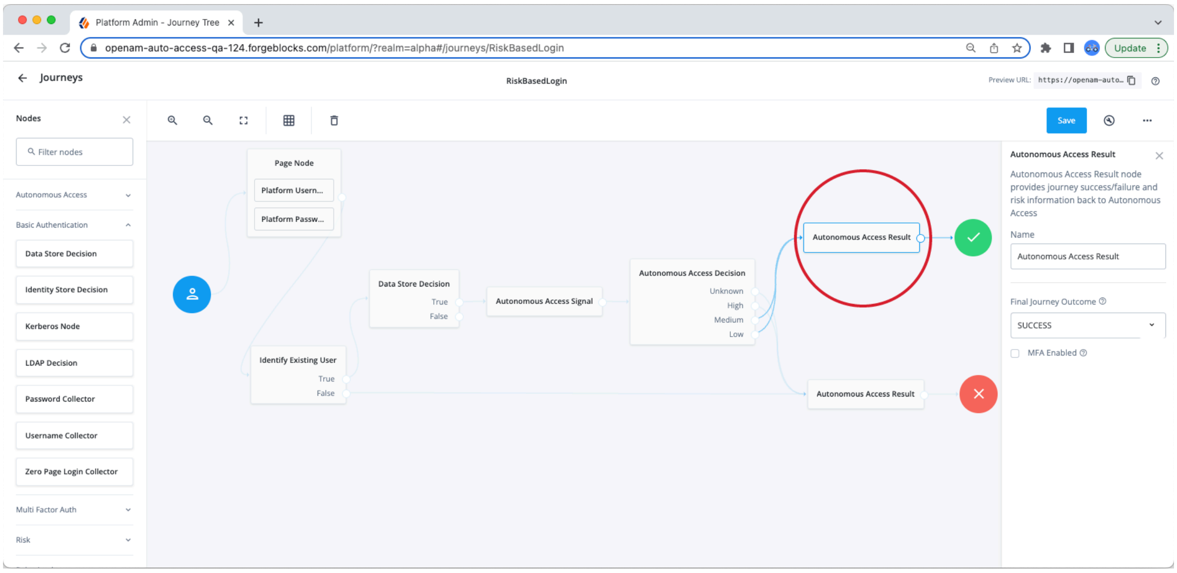
Property	Usage
High	Evaluation continues along its outcome path if the risk node is within the *high* range of scores. High risk scores indicate that the login attempt is highly likely to be a threat. The high range is determined by how you set it on the medium risk threshold slider.
Medium	Evaluation continues along its outcome path if the risk node is within the *medium* range of scores. Medium risk scores indicate that the login attempt is unlikely to be a high-risk threat event, but there is enough risk to warrant flagging the user for additional information to complete the authentication process. Depending on the application, you can use this range in your multifactor authentication journeys, or send these scores together with high risk scores. The medium range is determined by how you set it on the medium risk threshold slider.
Low	Evaluation continues along its outcome path if the risk node is within the *low* range of scores. Low risk scores indicate that the login attempt is unlikely to pose as a threat and does not require additional information to complete the authentication process. The range is determined by how you set it on the low risk threshold slider.
Low risk threshold slider	Sets the maximum (inclusive) value for low risk score threshold. For example, if the low risk threshold is 30; then, the low risk range is from 0 to 30. The low risk range is displayed in green on the slider.
Medium risk threshold slider	Sets the maximum (inclusive) value for the medium range of scores. For example, if the medium risk threshold is 70, values between 31 (that is, the low risk threshold) to 70 specify the range for medium risk scores, displayed on the slider in yellow. Values between 71 and 100 are defined as the high risk score range, displayed on the slider as a red line.

Result node

The Result node communicates the final state (e.g., SUCCESS, FAILURE, MFA ENABLED) and risk prediction results from the AI/ML analytics to the AAI server. The AAI server needs to know if the event and subsequent transaction has completed successfully or not to ensure up-to-date models.

> IMPORTANT
>
> The Result node must be present **before** the Success node and Failure node for data collection.

> NOTE
>
> You will need two or three result nodes for each journey, depending on application. One or two before the Success node and one before the Failure node.



Autonomous Access Result

Autonomous Access Result node provides journey success/failure and risk information back to Autonomous Access

Name: Autonomous Access Result

Final Journey Outcome: SUCCESS

□ MFA Enabled

Table 3. Result node properties

Property	Usage
Success	Indicates a successful journey outcome.
Failure	Indicates a failed journey outcome.
MFA Enabled	Check this box if step up or multifactor authentication (MFA) is enabled for the journey. This node is required for the AAI server to know that MFA completed successfully or not.

Learn about some configuration tips

This section presents tips on any aspect of Autonomous Access and will be updated on an ongoing basis.

GDPR Requests

PingOne Advanced Identity Cloud’s Autonomous Access capability stores user records with a 6-month archival rule. End users under specific privacy laws, such as General Data Protection Regulation (GDPR), can enter a self-service request to access or remove their data processed by Autonomous Access by filling out an online form. The general instructions are presented in the following section.

Handling GDPR requests

1. End users can request to access or remove their data by filling out a questionnaire on the ForgeRock Privacy Hub.
2. End user fills out and completes the form and sends it to the ForgeRock Privacy Team via privacy@forgerock.com.
3. ForgeRock Privacy Team enters a ticket with the completed form to the ForgeRock Operations team. The ForgeRock Operations team also copies the Autonomous Access Product team to action the request.
4. Autonomous Access Product team provides the Operations team with instructions to carry out the request.
5. The ForgeRock Operations team does the following:
    - Right to access: Downloads the data into a JSON or CSV file and deposits it into the end user’s preferred BackStage account.
    - Right to remove: Removes the data and advises the end user of the deletions via their BackStage account.

Allow and block lists

Autonomous Access supports allow and block IP lists to accept certain IP addresses and block known risky IPs.

Allow and block lists have the following general rules, Autonomous Access:

- Supports IPv4 and IPv6 formats.
- Takes single IP addresses, like 10.0.48.0, or IP subnets, like 10.0.48.0/24.
- Cannot use regular expressions or wildcards with the IP addresses.
- Assigns a risk score of 0 to any IPs on the *allow list* and excludes them from heuristics and machine learning.
- Subjects IPs on the *block list* to heuristics and machine learning and overrides their computed risk score with a score of 100.

IMPORTANT

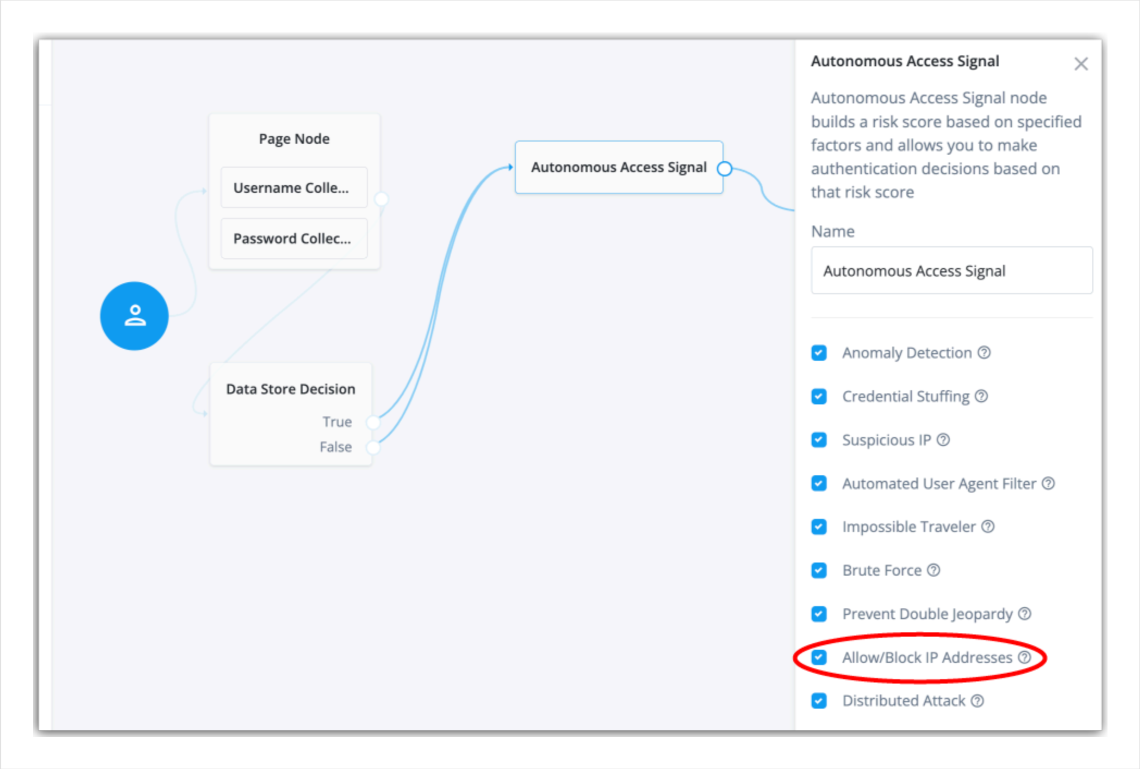
IMPORTANT

Autonomous Access is not a firewall. You must consume the output risk score in a succeeding node in the journey for actionable outcomes. Autonomous Access cannot allow or block any IP address by itself.

Configure allow and block IP lists

1. Set up a journey or edit an existing one.
2. On your journey, click the Autonomous Access signal node, and then click **Allow/Block IP Addresses**. This checkbox enables the `Allow/Block IP Addresses` heuristic in the machine learning process.

▼ *Click to display the signal node*



Autonomous Access Signal ✕

Autonomous Access Signal node builds a risk score based on specified factors and allows you to make authentication decisions based on that risk score

Name

Autonomous Access Signal

☑ Anomaly Detection ⓘ

☑ Credential Stuffing ⓘ

☑ Suspicious IP ⓘ

☑ Automated User Agent Filter ⓘ

☑ Impossible Traveler ⓘ

☑ Brute Force ⓘ

☑ Prevent Double Jeopardy ⓘ

☑ Allow/Block IP Addresses ⓘ

☑ Distributed Attack ⓘ

3. Click **Risk Administration > Risk Config**.
4. On the Risk Config page, enter the IP addresses that you want to include in the `BLOCK_LIST` and `ALLOW_LIST`. For example:

```
block_and_allow_list:
  BLOCK_LIST: [120.18.1.10, 120.22.8.2]
  ALLOW_LIST: [10.0.48.0/24]
```

5. Click **Save**.
6. On the Preview Risk Evaluation popup window, do the following:
   a. Click Bucket Search to select your data source location or type the name of the data source location.

b. Optional. Enter an object prefix to filter your search results.

c. Next to your desired object, click the trailing dots, and then click **Preview Object** to display your data source change(s).

d. Click **Preview Risk Evaluation** to review a simulated risk evaluation for the first event.

e. If you are satisfied with your change(s), click **Save Config**.

> NOTE
>
> If you get a preview error, ignore the message and click **Save Config**.

Was this helpful? 👍 👎