

# Trustworthy Decision Management

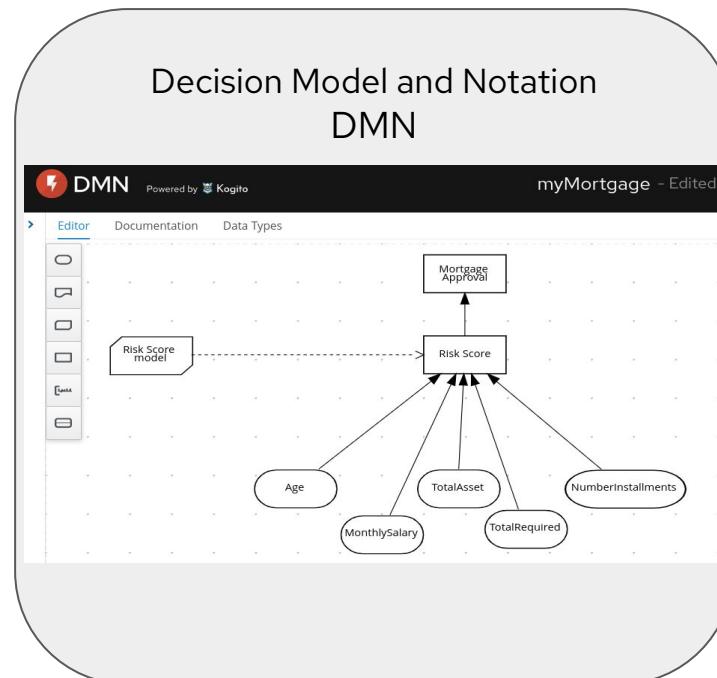
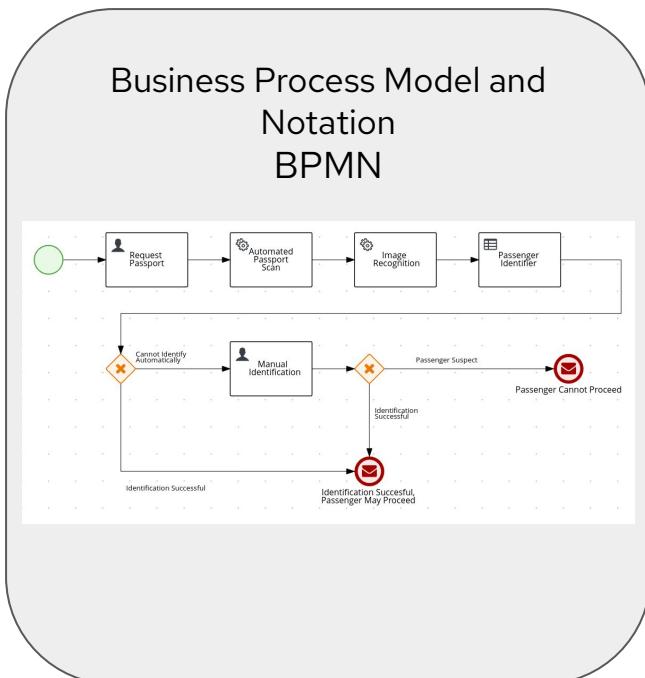
How explainable, predictive decision making can help us trust our Decision models

Jacopo Rota  
Senior Software Engineer



# Red Hat Process Automation Manager

It used to be that businesses simply needed to **automate processes**, and that was enough to **increase efficiency** and **control costs** across the organization. This was largely achieved through **BPM (Business Process Management)** and **BRM (Business Rules Management)**. BPM and BRM are great technologies, but on their own they are no longer enough.



**Rule automation (DRL)**

The screenshot displays the Red Hat Process Automation Manager interface. On the left, a code editor shows DRL (Decision Rule Language) code for a 'HelloWorld.drl' file. The code defines rules for 'Hello World' and 'Good Bye' based on message status. On the right, the 'Console' tab shows the activation of a rule named 'Good Bye' with the message 'Goodbye cruel world'.

```

rule "Hello World"
  dialect "mvel"
  when
    m : Message( status == Message.HELLO, message : message )
  then
    System.out.println( message );
    modify { m } { message = "Goodbye cruel world",
      status = Message.GOODBYE };
  end

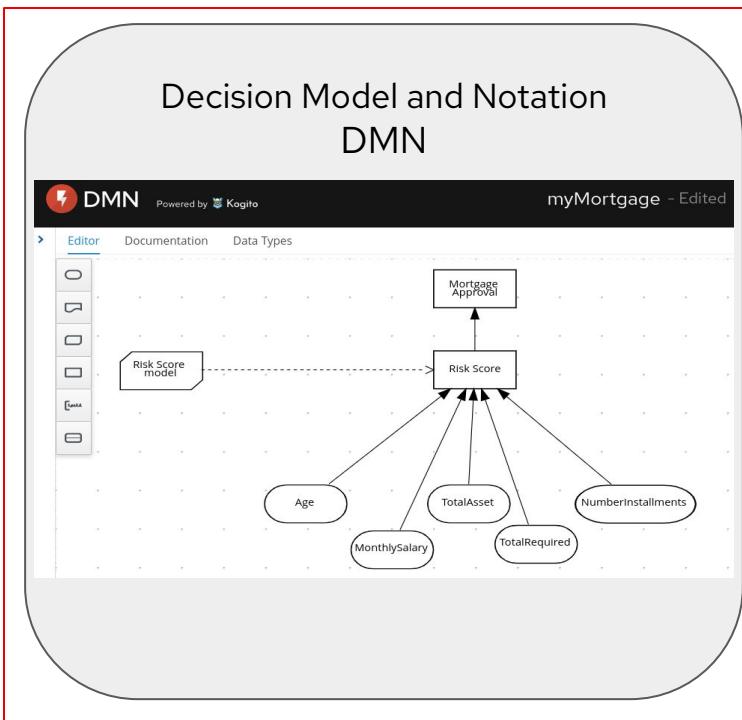
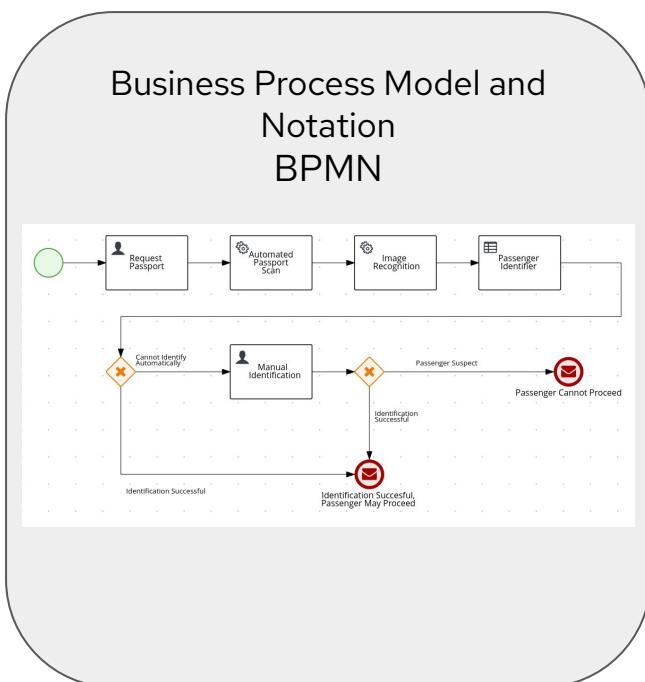
rule "Good Bye"
  dialect "java"
  when
    Message( status == Message.GOODBYE, message : message )
  then
    System.out.println( message );
  end
  
```

MAIN[focus]=BinaryHeapQueueAgendaGroup (id=1530)
 □ ▲ [0]=Activation
 □ ▲ ruleName= "Good Bye"
 □ ▲ message= "Goodbye cruel world"



# Red Hat Process Automation Manager

It used to be that businesses simply needed to **automate processes**, and that was enough to **increase efficiency** and **control costs** across the organization. This was largely achieved through **BPM (Business Process Management)** and **BRM (Business Rules Management)**. BPM and BRM are great technologies, but on their own they are no longer enough.



**Rule automation (DRL)**

The screenshot shows the Red Hat Process Automation Manager interface. In the top right, there is a text editor window titled 'HelloWorldExample.java' containing DRL code. Below it is a 'Rule Tree' window showing the hierarchical structure of the rules. At the bottom, there is a navigation bar with tabs like 'Console', 'Tasks', 'Agenda View', 'Audit View', 'Global Data View', 'Rules View', 'Working Memory View', and 'JUnit'.

```

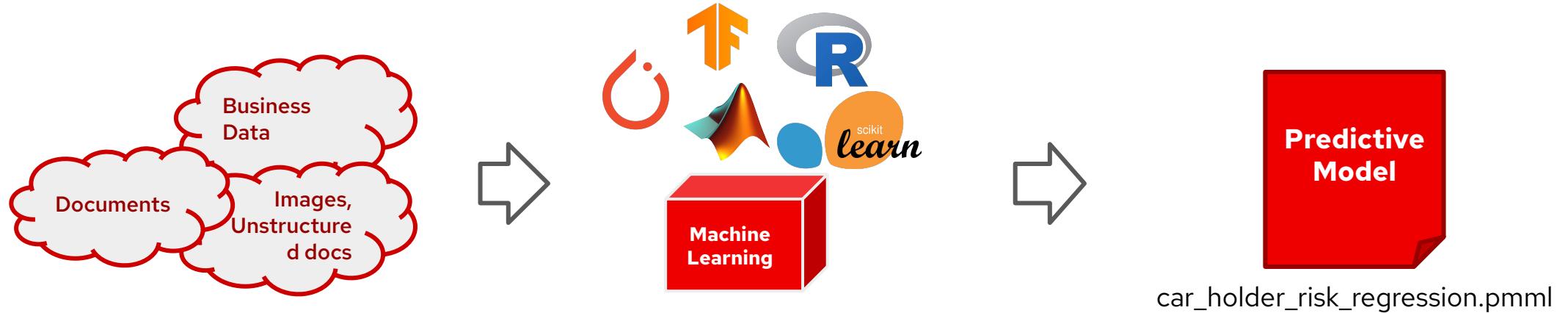
rule "Hello World"
  dialect "jrule"
  when
    m : Message( status == Message.HELLO, message : message )
  then
    System.out.println( message );
    modify { m } { message = "Goodbye cruel world",
      status = Message.GOODBYE };
  end

rule "Good Bye"
  dialect "java"
  when
    Message( status == Message.GOODBYE, message : message )
  then
    System.out.println( message );
  end
  
```

# Connecting Business Automation and Machine Learning



# DMN meets PMML



3. Editor automatically shows the parameters the model expects

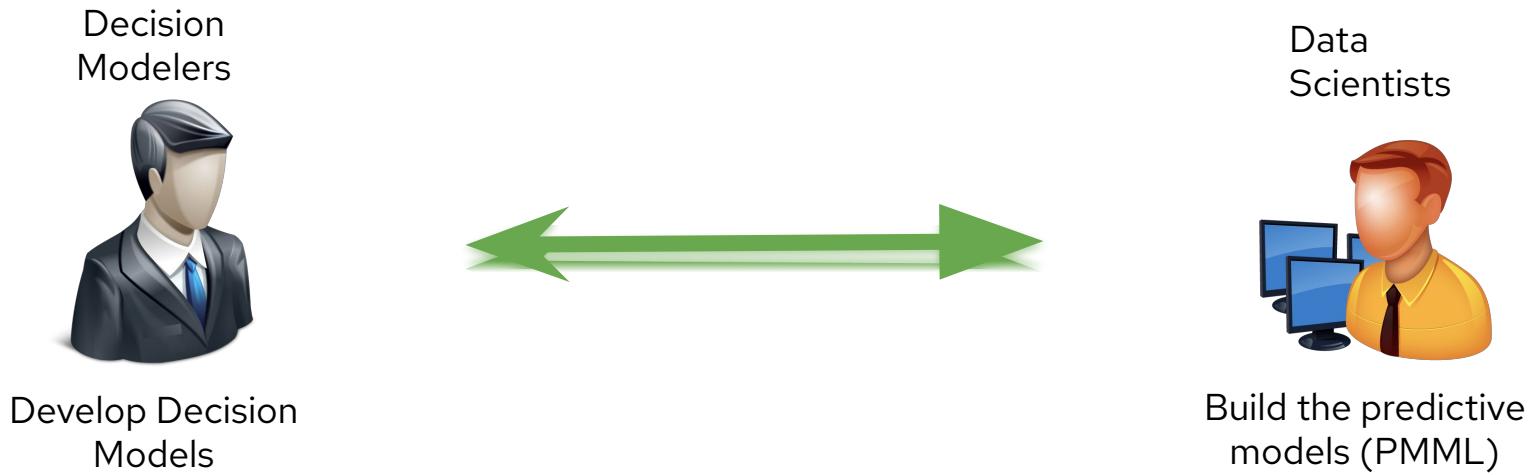
Cardholder Risk Model ( <i>Function</i> )		
Cardholder Risk Model (<Undefined>)		
	(age, holder_index, incidents)	
P	1	document ( <i>string</i> ) "card_holder_risk_linear_regression"
	2	model ( <i>string</i> ) "LinearRegression"

1. Choose the PMML file

2. Choose the model within the file

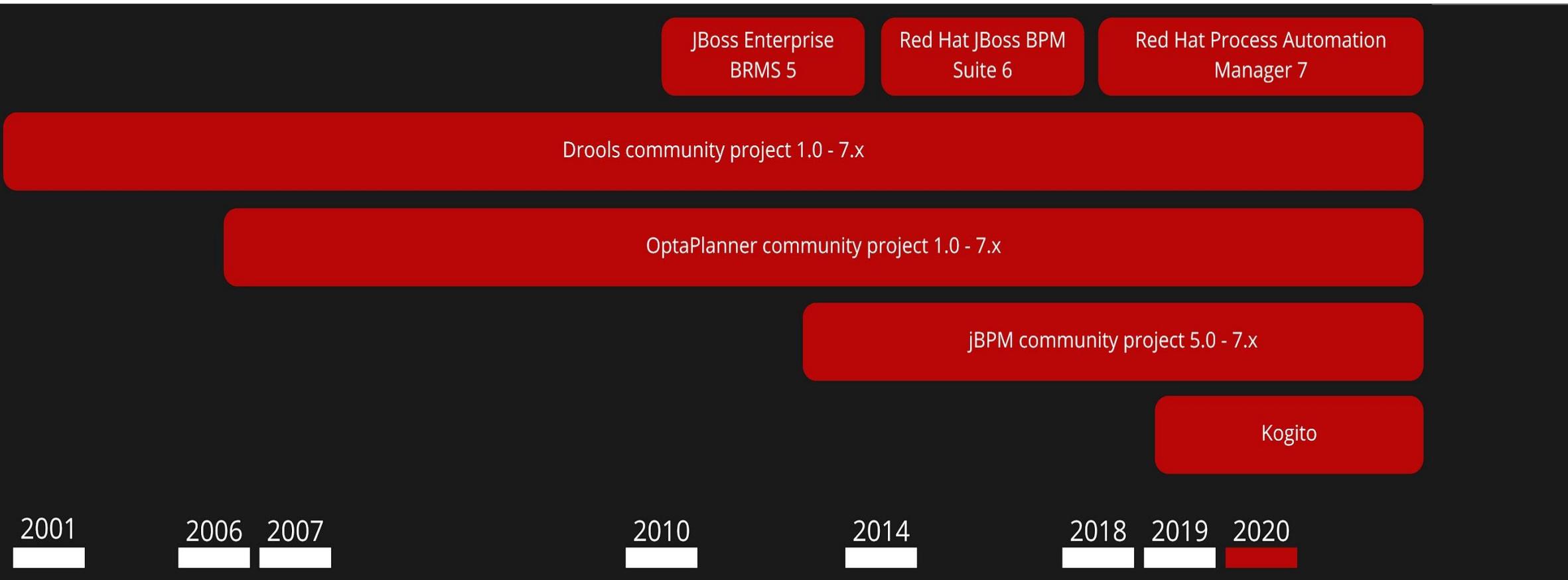


# Benefits of an integrated, standards-based solution



- Direct consumption of predictive models in decision models
  - No translation needed
- Supports all 19 executable models from PMML (Score cards, Neural nets, Regression, Random Forest, etc)
- Open the AI box - helps with transparency and explanation
- Direct collaboration between Data scientists and Decision Modelers
- Enables event correlation and consolidation - KPI monitoring

# Product History and Upstream Projects



# Next-gen Cloud-Native Business Automation

Cloud-Native Business Automation for building intelligent applications,  
backed by battle-tested capabilities



Kogito



kubernetes



OPENSHIFT



OPERATOR  
FRAMEWORK



QUARKUS



spring  
boot



Can you *trust* your decisions?

 REUTERS    World    Business    Markets    Politics    TV

Midterm Elections    Imprisoned In Myanmar    Sectors Up Close    Breakingviews    Investing    Future of Money    Charged: The Future of Aut

BUSINESS NEWS    OCTOBER 10, 2018 / 5:12 AM / A MONTH AGO

## Amazon scraps secret AI tool that showed bias against women

Jeffrey Dastin

SAN FRANCISCO (Reuters) - Amazon.com Inc's specialists uncovered a big problem: their

The group created 500 computer models focused on specific job functions and locations. They taught each to recognize some 50,000 terms that showed up on past candidates' resumes. The algorithms learned to assign little significance to skills that were common across IT applicants, such as the ability to write various computer codes, the people said.

Instead, the technology favored candidates who described themselves using verbs more commonly found on male engineers' resumes, such as "executed" and "captured," one person said.

**Amazon trained a sexism-fighting, resume-screening AI with sexist hiring data, so the bot became sexist**



**THE VERGE**    TECH ▾ SCIENCE ▾ C

TECH    AMAZON    ARTIFICIAL INTELLIGENCE

## Amazon reportedly scraps internal AI recruiting tool that was biased against women

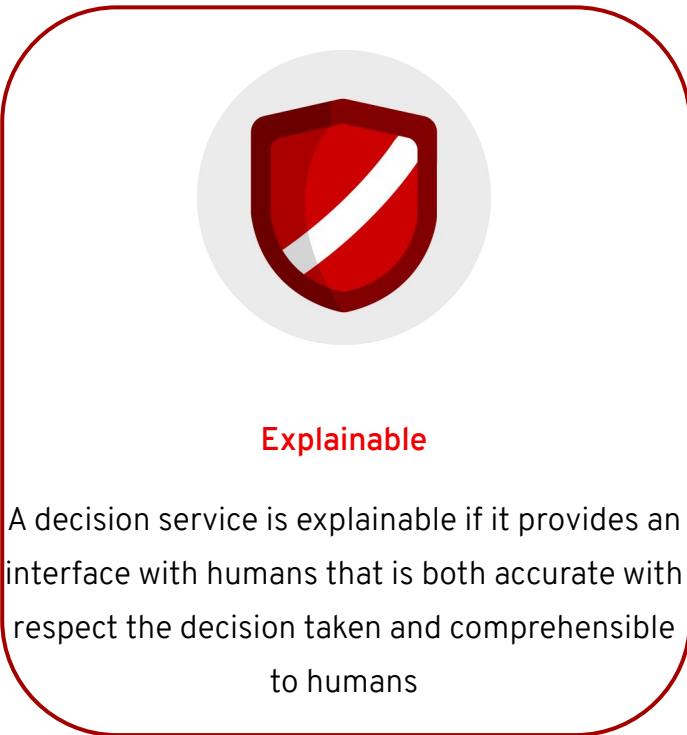
*The secret program penalized applications that contained the word "women's"*

# My decision service is...



## Monitorable

A decision service is considered monitorable if it is possible to check in real time the behaviour of the service



## Explainable

A decision service is explainable if it provides an interface with humans that is both accurate with respect the decision taken and comprehensible to humans



## Auditable

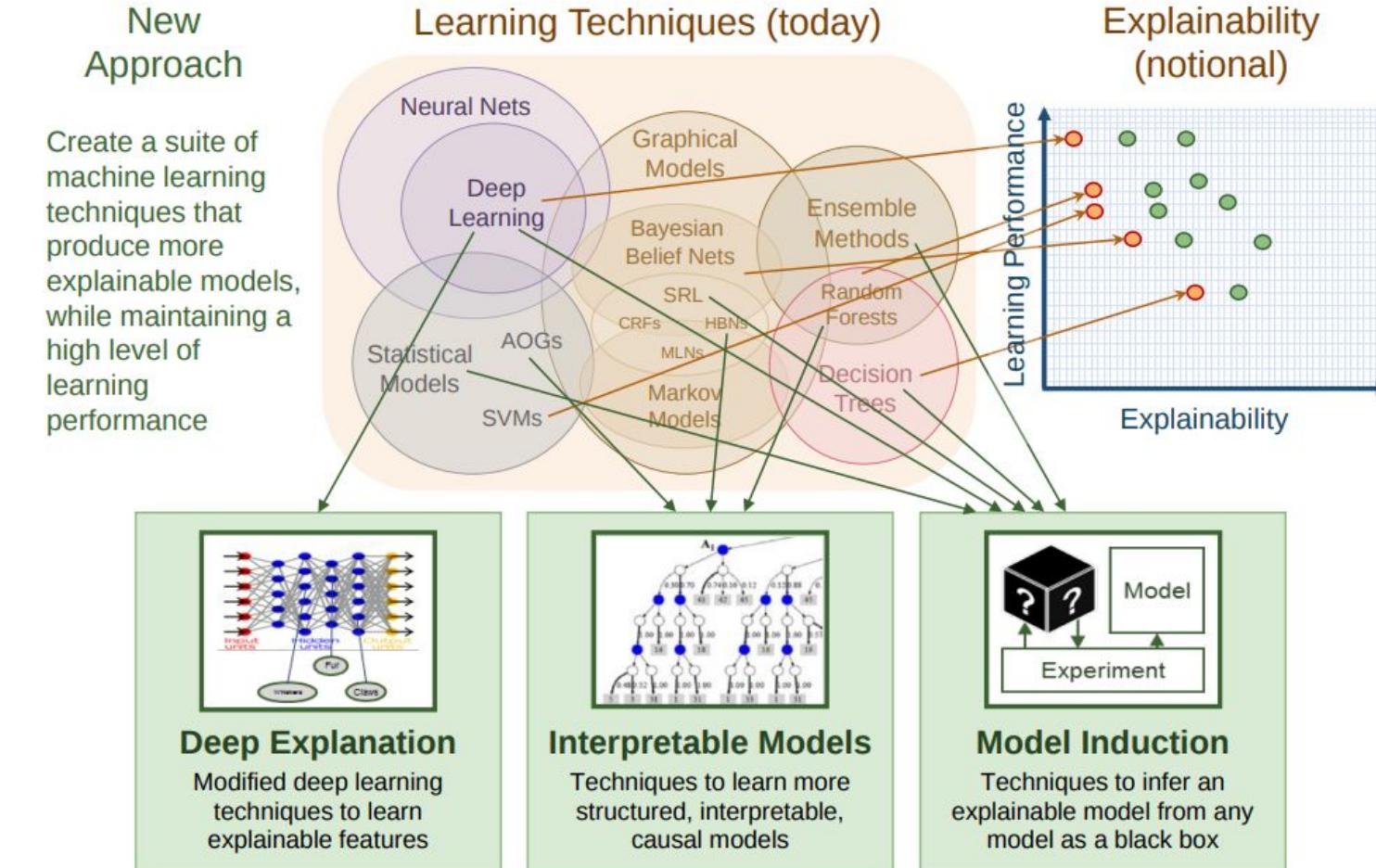
A decision service is auditable if it produces execution information, stores them and makes them accessible

# The right tool to the right stakeholder

- **Case worker**
  - Good domain knowledge, case by case
  - No technical knowledge
- **Compliance worker**
  - Good high level domain knowledge
  - No technical knowledge
- **Data scientist**
  - Low/limited domain knowledge
  - Good technical knowledge

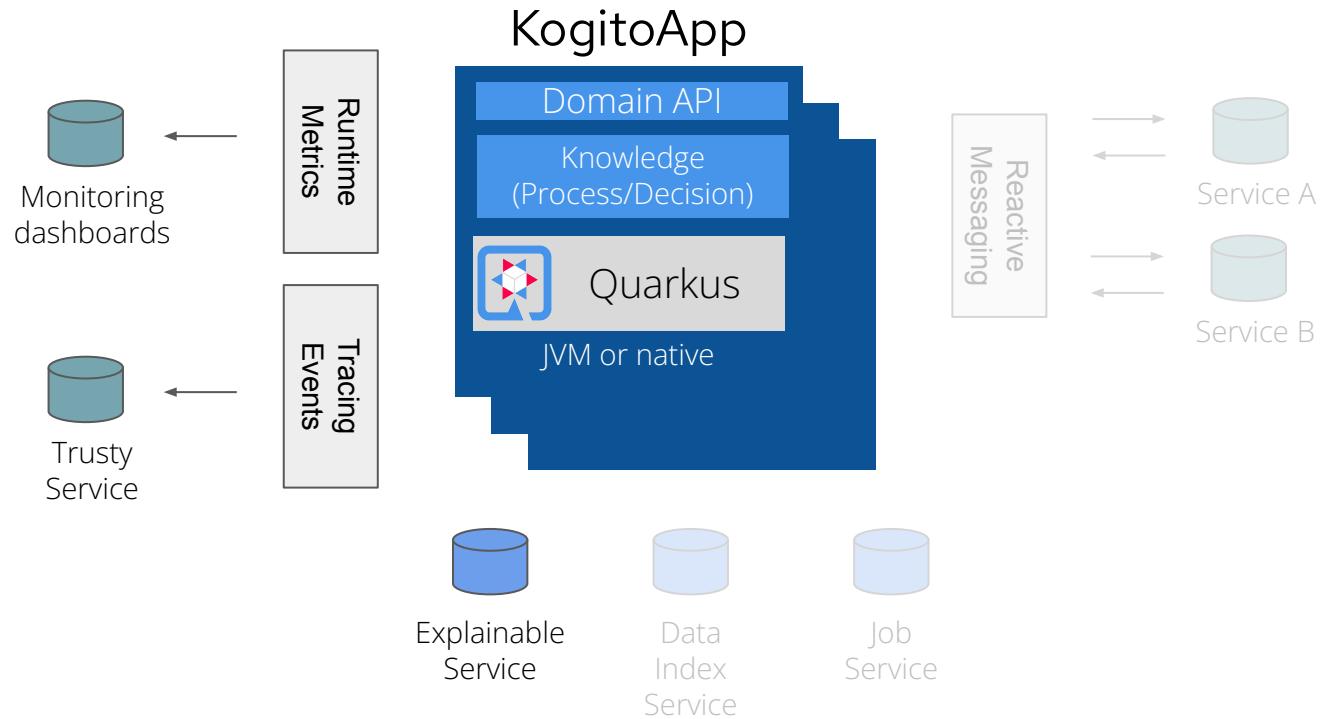


# Performance vs. Explainability



# TrustyAI

# TrustyAI Services



OpenShift

# TrustyAI

Offer value-added services for Business Automation.

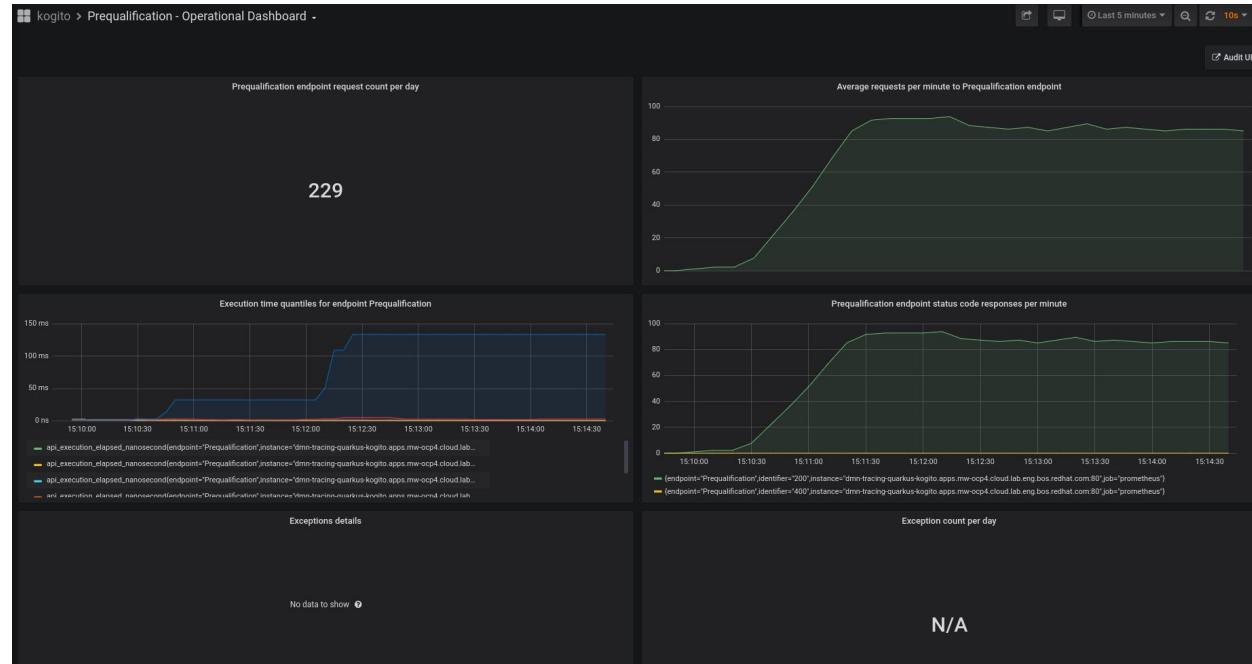
- **Runtime Monitoring Service**
  - dashboard for business runtime monitoring
- **Tracing and Accountability Service**
  - extract, collect and publish metadata for auditing and compliance
- **Explanation Service**
  - XAI algorithms to enrich model execution information

# Business Monitoring



- Real time business metrics.
- Monitors decision making to ensure it is correct.
- Displays metrics based on model decisions.
- Stakeholders can then monitor the system for business risk and optimization opportunities.

# Operational Monitoring



- Real time monitoring service for operational metrics.
- Provides execution monitoring for the decisions.
- Devops engineers can check for correct deployment and system health.

# Audit UI

The screenshot shows the Kogito Audit Investigation UI. The top navigation bar includes the Kogito logo and a user icon. The left sidebar has links for Audit Investigation, Business Monitoring, and Operational Monitoring. The main content area shows an audit investigation with ID #B2B0ED8D-C1E2-46B5-3AC54FF4BEAE-1006, which is completed. The 'Outcomes' tab is selected. It displays two outcome cards: one for 'Confidence' (value 0.8) and one for 'Approved' (value False). Each card has a 'View Details' button.

Outcome Type	Value	Status
Confidence	0.8	Approved
Approved	False	Approved

- Trace decision execution
- Provides ability to query historic decisions
- Introspection of each individual decision made within the system
- Details of decision outcomes
- Provides model metadata for auditing purposes

# LIME

- Saliency method
- LIME tests what happens to the prediction when you provide *perturbed* versions of the input to the black box model
- Trains an **interpretable** model (e.g. a linear classifier) to separate perturbed data points by label
- The *weights* of the linear model (one for each feature) are used as **feature importance** scores



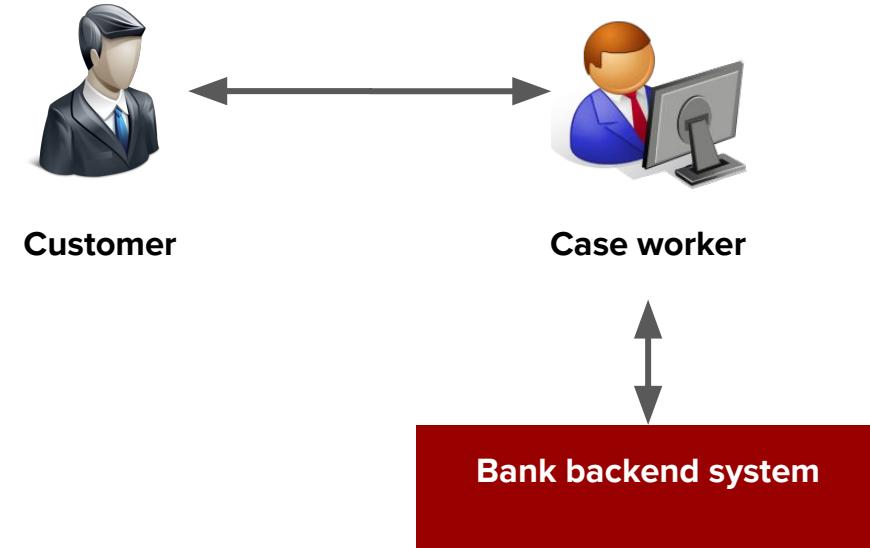
# Counterfactual explanations

- **Exemplar** explanations provide explanations for single predictions by means of **examples** (in the input space)
  - **Counterfactual explanations** provide examples that
    - Have a *desired* prediction, according to the black box model
    - Are as *close* as possible to the original input
  - How should the user change its inputs in order to get a formerly rejected credit card request granted?
- Usually work by **minimizing** two cost functions
  - **Input cost** : representing the distance between the original input and a new input
  - **Target cost** : representing the distance between the desired output and the output generated by querying the model with the new input

# Demo

# Demo use case: Mortgage approval

"As a case worker I want to be able to **explain** to customer **why** that mortgage request was rejected or accepted."



## Demo

OpenAPI URL: /q/openapi

Select POST endpoint: /myMortgage

Age\*: 47

MonthlySalary\*: 4000

TotalAsset\*: 250000

TotalRequired\*: 100000

NumberInstallments\*: 100

Submit

Form request payload:

```
{
  "Age": 47,
  "MonthlySalary": 4000,
  "TotalAsset": 250000,
  "TotalRequired": 100000,
  "NumberInstallments": 100
}
```

Override request payload manually

Response:

```
{
  "MonthlySalary": 4000,
  "TotalAsset": 250000,
  "Risk Score": 60.24883274455427,
  "TotalRequired": 100000,
  "NumberInstallments": 100,
  "Mortgage Approval": false,
  "Policy Compliance": true,
  "Age": 47,
  "Risk Score model": "function Risk Score model( Age, MonthlySalary, NumberInstallments, TotalAsset, TotalRequired )"
}
```

The screenshot shows a web browser window titled "Kogito - TrustyAI" with the URL "localhost:1338/audit". The page has a dark header with the Kogito logo and a user icon. A sidebar on the left is titled "Audit Investigation". The main content area is titled "Audit Investigation" and contains the message "Here you can retrieve all the available information about past cases.". Below this is a search bar with fields for "Search by ID" and "From" (set to August 10, 2021) and "To" (set to September 10, 2021). To the right of the search bar is a refresh icon. At the bottom of the search bar are pagination controls showing "1-1 of 1" and "1 of 1". The main table displays one row of data:

ID	Description	Executor	Date	Execution Status
#143e9fac	myMortgage		1 minute ago	✓ Completed

Below the table are two sets of pagination controls, both showing "1-1 of 1" and "1 of 1".

The screenshot shows a web browser window titled "Kogito - TrustyAI" displaying the "Input Data" section of an audit investigation. The URL in the address bar is `localhost:1338/audit/decision/143e9fac-e53e-4081-90c2-5802f217e767/input-data`. The page header includes the Kogito logo and a user profile icon. The breadcrumb navigation shows "Audit Investigation > Execution #143e9fac > Input Data". The execution status is "Completed". Below the navigation, there are tabs for "Outcomes", "Outcomes Details", "Input Data" (which is underlined), "Model Lookup", and "Counterfactual Analysis". The main content area is titled "Input Data" and features a "Browse Sections" button. A "Root" section is selected, highlighted with a blue background. The table below lists input data items with their values:

Input Data	Value
MonthlySalary ROOT	4000
TotalAsset ROOT	250000
TotalRequired ROOT	100000
NumberInstallments ROOT	100
Age ROOT	47

The screenshot shows a web browser window titled "Kogito - TrustyAI" displaying audit results. The URL in the address bar is `localhost:1338/audit/decision/143e9fac-e53e-4081-90c2-5802f217e767/outcomes`. The page header includes the Kogito logo and a user profile icon. The main navigation menu on the left lists "Audit Investigation". The current page title is "Execution #143e9fac" (status: Completed). Below the title, there are tabs for "Outcomes" (selected), "Outcomes Details", "Input Data", "Model Lookup", and "Counterfactual Analysis". The "Outcomes" section displays three items: "Policy Compliance" (True), "Risk Score" (60.24883274455427), and "Mortgage Approval" (False). Each item has a "View Details" link.

Audit Investigation > Execution #143e9fac > Outcomes

Execution #143e9fac ✓ Completed

Outcomes    Outcomes Details    Input Data    Model Lookup    Counterfactual Analysis

Outcomes

Policy Compliance	Risk Score	Mortgage Approval
True	60.24883274455427	False
<a href="#">View Details →</a>	<a href="#">View Details →</a>	<a href="#">View Details →</a>

Kogito

Execution #143e9fac Completed

Outcomes Outcomes Details Input Data Model Lookup Counterfactual Analysis

Risk Score

**Outcome Details**

Risk Score: 60.24883274455427

**Explanation**

**Features Score Chart**

Positive Impact

Negative Impact

TotalRequired

MonthlySalary

TotalAsset

Age

**Features Weight**

Positive Weight	Score
TotalRequired	3.49
Age	2.75
TotalAsset	2.75
MonthlySalary	2.75
NumberInstallments	2.58

**Negative Weight**

Negative Weight	Score
None	0.00

The screenshot shows a browser window titled "Kogito - TrustyAI" with the URL "localhost:1338/audit/decision/143e9fac-e53e-4081-90c2-5802f217e767/counterfactual-analysis". The page header includes the Kogito logo and a user profile icon. The main navigation bar shows "Audit Investigation > Execution #143e9fac > Counterfactual Analysis". The execution status is "Completed". Below the navigation, there are tabs: "Outcomes", "Outcomes Details", "Input Data", "Model Lookup", and "Counterfactual Analysis", with "Counterfactual Analysis" being the active tab.

The "Counterfactual Analysis" section displays the message: "Selected Outcomes: Policy Compliance: Any, Risk Score: Any, Mortgage Approval: true". A sub-section titled "Create a counterfactual" contains the instruction: "Select a desired counterfactual Outcome: one or more Data Types, and modify the input constraints." A tooltip for "Set the desired decision outcomes for a counterfactual analysis." points to this instruction. Below this are three buttons: "Run Counterfactual", "Set Up Outcomes" (which is highlighted with a blue border), and "Reset".

The main table lists six input parameters with their corresponding constraints and results:

Input	Constraint	Input Value	Counterfactual Result
MonthlySalary	+ Constraint	4000	No available results
TotalAsset	+ Constraint	250000	No available results
TotalRequired	+ Constraint	100000	No available results
NumberInstallments	+ Constraint	100	No available results
Age	+ Constraint	47	No available results

The screenshot shows a web browser window for 'Kogito - TrustyAI' at the URL `localhost:1338/audit/decision/143e9fac-e53e-4081-90c2-5802f217e767/counterfactual-analysis`. The main navigation bar includes 'Audit Investigation', 'Execution #143e9fac' (Completed), and 'Counterfactual Analysis'. The 'Counterfactual Analysis' tab is active, displaying a modal dialog titled 'Specify desired outcomes'. The dialog lists three original decision outcomes: 'Policy Compliance' (True), 'Risk Score' (60.24883274455427), and 'Mortgage Approval' (False). For each, there is a corresponding 'Desired counterfactual outcome' section. Under 'Policy Compliance', the toggle is set to 'True' with the checkbox 'Automatically adjust for counterfactual.' checked. Under 'Risk Score', a numeric slider is set to 60.24883274455427 with the checkbox 'Automatically adjust for counterfactual.' checked. Under 'Mortgage Approval', the toggle is set to 'True' with the checkbox 'Automatically adjust for counterfactual.' checked. At the bottom of the dialog are 'Confirm' and 'Cancel' buttons. Below the dialog, the main page shows three rows of constraints: 'TotalRequired' (Constraint: 100000, No available results), 'NumberInstallments' (Constraint: 100, No available results), and 'Age' (Constraint: 47, No available results).

The screenshot shows a web browser window for the Kogito - TrustyAI application. The URL is `localhost:1338/audit/decision/143e9fac-e53e-4081-90c2-5802f217e767/counterfactual-analysis`. The page title is "Audit Investigation > Execution #143e9fac > Counterfactual Analysis". The execution status is "Completed". The navigation tabs include Outcomes, Outcomes Details, Input Data, Model Lookup, and Counterfactual Analysis, with the latter being the active tab.

The main content area is titled "Counterfactual Analysis" and displays the message: "Selected Outcomes: Policy Compliance: Any, Risk Score: Any, Mortgage Approval: true". Below this, there is a section titled "Create a counterfactual" with the instruction: "Select a desired counterfactual Outcome; one or more Data Types, and modify the input constraints." It includes buttons for "Run Counterfactual", "Set Up Outcomes" (which is highlighted in blue), and "Reset".

A table lists input fields and their corresponding constraints:

Input	Constraint	Input Value
MonthlySalary	+ Constraint	4000
TotalAsset	+ Constraint	250000
TotalRequired	+ Constraint	100000
NumberInstallments	+ Constraint	100
Age	+ Constraint	47

To the right, a modal dialog is open for the "Age" field, titled "Age". It shows the "Data Type" as "number" and the "Original Value" as "47". In the "Constraint" section, the "Minimum Value" is set to "25" and the "Maximum Value" is set to "60". There are "Apply" and "Cancel" buttons at the bottom of the modal.

The screenshot shows a web browser window for the Kogito - TrustyAI application. The URL is `localhost:1338/audit/decision/143e9fac-e53e-4081-90c2-5802f217e767/counterfactual-analysis`. The page title is "Audit Investigation > Execution #143e9fac > Counterfactual Analysis". The execution status is "Completed". The navigation tabs include Outcomes, Outcomes Details, Input Data, Model Lookup, and Counterfactual Analysis, with the latter being the active tab.

The main content area is titled "Counterfactual Analysis" and displays the message "Selected Outcomes: Policy Compliance: Any, Risk Score: Any, Mortgage Approval: true". Below this, a sub-section titled "Create a counterfactual" contains the instruction "Select a desired counterfactual Outcome; one or more Data Types, and modify the input constraints. Run the counterfactual analysis based on selected Inputs and Outcomes." A tooltip-like callout points to the "Run Counterfactual" button.

Below the instructions are three buttons: "Run Counterfactual" (dark blue), "Set Up Outcomes" (light blue), and "Reset". The "Run Counterfactual" button is highlighted with a dark blue background and white text.

The main table lists inputs and their corresponding constraints and values:

Input ⓘ	Constraint ⓘ	Input Value ⓘ	Counterfactual Result
MonthlySalary	+ Constraint	4000	No available results
TotalAsset	+ Constraint	250000	No available results
TotalRequired	+ Constraint	100000	No available results
NumberInstallments	+ Constraint	100	No available results
Age	25-60	47	No available results

The screenshot shows a browser window for 'Kogito - TrustyAI' at the URL `localhost:1338/audit/decision/143e9fac-e53e-4081-90c2-5802f217e767/counterfactual-analysis`. The page title is 'Audit Investigation > Execution #143e9fac > Counterfactual Analysis'. The execution is marked as 'Completed'. The navigation tabs include 'Outcomes', 'Outcomes Details', 'Input Data', 'Model Lookup', and 'Counterfactual Analysis', with the latter being the active tab.

The main content area is titled 'Counterfactual Analysis' and displays the message 'Selected Outcomes: Policy Compliance: Any, Risk Score: Any, Mortgage Approval: true'. Below this, there are two status messages: 'Calculating...' and 'Wrapping up', accompanied by a progress bar.

A table lists the input variables and their corresponding values:

Input	Constraint	Input Value	Counterfactual Result
MonthlySalary		4000	4000
TotalAsset		250000	250000
TotalRequired		100000	100000
NumberInstallments		100	100
Age	25-60	47	35.49

The 'Counterfactual Result' column for the 'Age' row is highlighted with a light blue background.

The screenshot shows a browser window titled "Kogito - TrustyAI" displaying the "Counterfactual Analysis" page for an execution. The URL is `localhost:1338/audit/decision/143e9fac-e53e-4081-90c2-5802f217e767/counterfactual-analysis`. The navigation bar includes "Audit Investigation" and "Execution #143e9fac" (Completed). The main content area has tabs for "Outcomes", "Outcomes Details", "Input Data", "Model Lookup", and "Counterfactual Analysis". The "Counterfactual Analysis" tab is selected. A success message states "Counterfactual Analysis completed successfully". Below it, a note says "To run another analysis, either create a new counterfactual or edit the existing counterfactual. Note: the current results will be cleared when another counterfactual has been initiated." There are two buttons: "New Counterfactual" and "Edit Counterfactual". The main table displays input values and their corresponding counterfactual results for five variables: MonthlySalary, TotalAsset, TotalRequired, NumberInstallments, and Age.

Input	Constraint	Input Value	Counterfactual Result	Counterfactual Result	Counterfactual Result
MonthlySalary		4000	4000	4000	4000
TotalAsset		250000	250000	250000	250000
TotalRequired		100000	100000	100000	100000
NumberInstallments		100	100	100	100
Age	25-60	47	35.50	35.50	35.49

## Demo

The image displays two browser windows demonstrating the integration of OpenAPI and machine learning models.

**OpenAPI as Forms (Left Window):**

- OpenAPI URL:** /q/openapi
- Select POST endpoint:** /myMortgage
- Form request payload:**

```
{
  "Age": 35,
  "MonthlySalary": 4000,
  "TotalAsset": 250000,
  "TotalRequired": 100000,
  "NumberInstallments": 100
}
```
- Override request payload manually:** (with Apply and Reset buttons)
- Response:**

```
{
  "MonthlySalary": 4000,
  "TotalAsset": 250000,
  "Risk Score": 27.77629639783284,
  "TotalRequired": 100000,
  "NumberInstallments": 100,
  "Mortgage Approval": true,
  "Policy Compliance": true,
  "Age": 35,
  "Risk Score model": "function Risk Score model( Age, MontlySalary, TotalAsset, TotalRequired )"
}
```

**Kogito - TrustyAI (Right Window):**

- Audit Investigation > Execution #334dcb59 > Outcomes**
- Execution #334dcb59** (Completed)
- Outcomes** (selected tab)
- Outcomes Details** | **Input Data** | **Model Lookup** | **Counterfactual Analysis**
- Outcomes** table:

Policy Compliance	Risk Score	Mortgage Approval
True	27.77629639783284	True
- View Details →** (links for Policy Compliance, Risk Score, and Mortgage Approval)

# Resources



[https://youtu.be/mg\\_4UvQzC3w](https://youtu.be/mg_4UvQzC3w)



<https://youtu.be/7CQqVlk6LdM>

This presentation **demo** - <https://github.com/r00ta/trusty-ai-ODSC-West>

Drools featuring **DMN support** - <https://drools.org/learn/dmn.html>

**Learn DMN** - <http://learn-dmn-in-15-minutes.com>

**Kogito** - <http://kogito.kie.org/>

**TrustyAI** landing page: <https://kogito.kie.org/trustya/>

**TrustyAI** playlist: <https://www.youtube.com/playlist?list=PLo3ZScdD9hW5hW3-9vN9MWXyWehjlqTtv>

**Red Hat** Business Process Automation: <https://www.redhat.com/en/products/process-automation>

# Thank you

Red Hat is the world's leading provider of enterprise open source software solutions. Award-winning support, training, and consulting services make Red Hat a trusted adviser to the Fortune 500.

 [linkedin.com/company/red-hat](https://www.linkedin.com/company/red-hat)

 [youtube.com/user/RedHatVideos](https://www.youtube.com/user/RedHatVideos)

 [facebook.com/redhatinc](https://www.facebook.com/redhatinc)

 [twitter.com/RedHat](https://twitter.com/RedHat)