



The prompt lab

Bee

Aleš Kalfas

September 24, 2024 v3



The prompt lab

Outlines

1. Concept
 1. Test case
 2. Test
 3. Project
2. The development process
 1. Initial
 2. Current
 3. Future
3. Planning
 1. Goals and feature
 2. Roadmap



The prompt lab *concept*

Test case

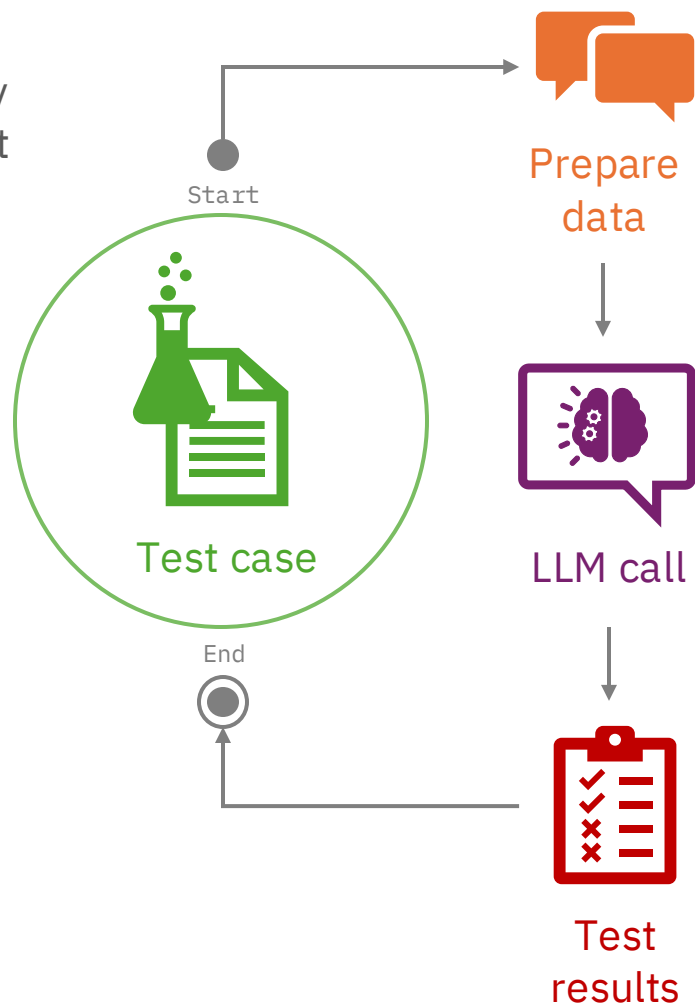


The prompt lab *concept*

Test case

Test case

It's the core functionality that we can logically split into three steps.





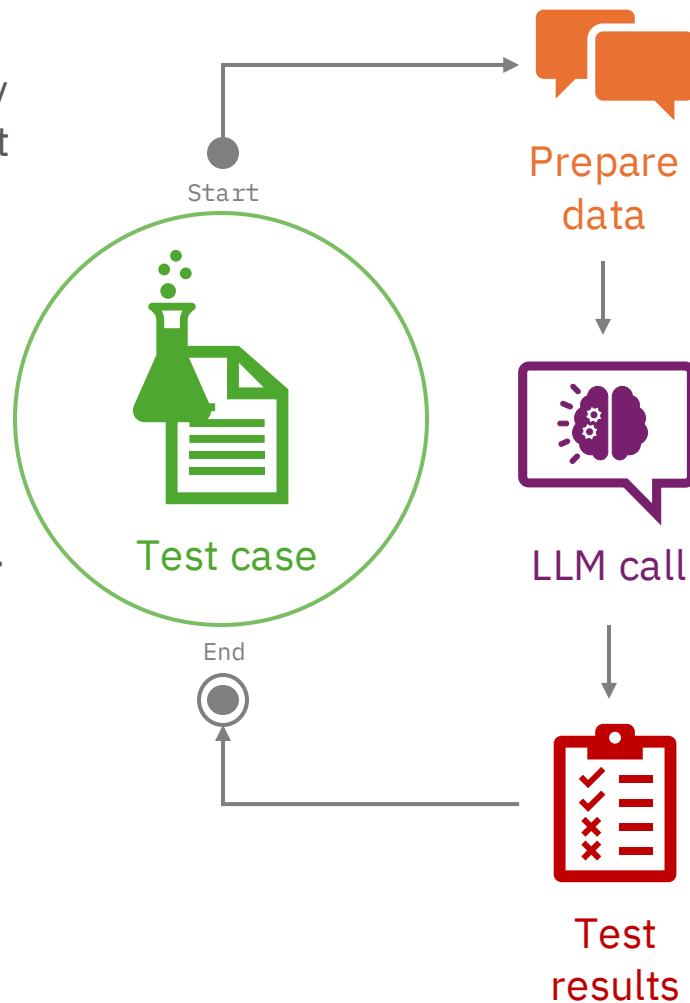
The prompt lab *concept*

Test case

Test case

It's the core functionality that we can logically split into three steps.

A test case is generic and can be run for different combinations of data and parameters. Each step takes own *>configs* or *>parameters*.





The prompt lab *concept*

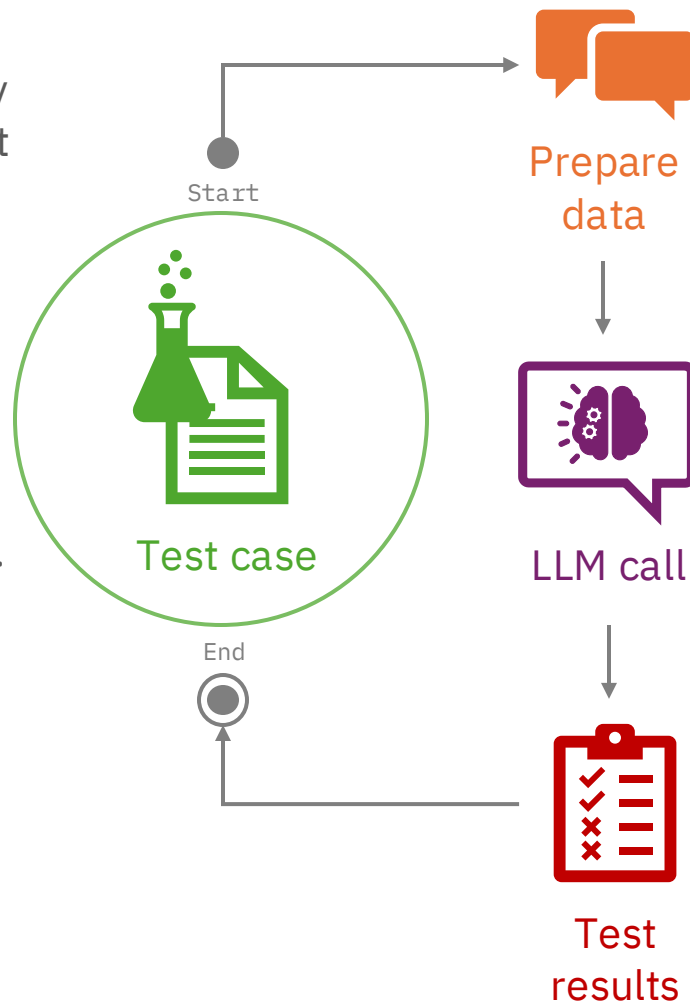
Test case

Test case

It's the core functionality that we can logically split into three steps.

A test case is generic and can be run for different combinations of data and parameters. Each step takes own *>configs* or *>parameters*.

The output of one step is *the input* to the next.





The prompt lab *concept*

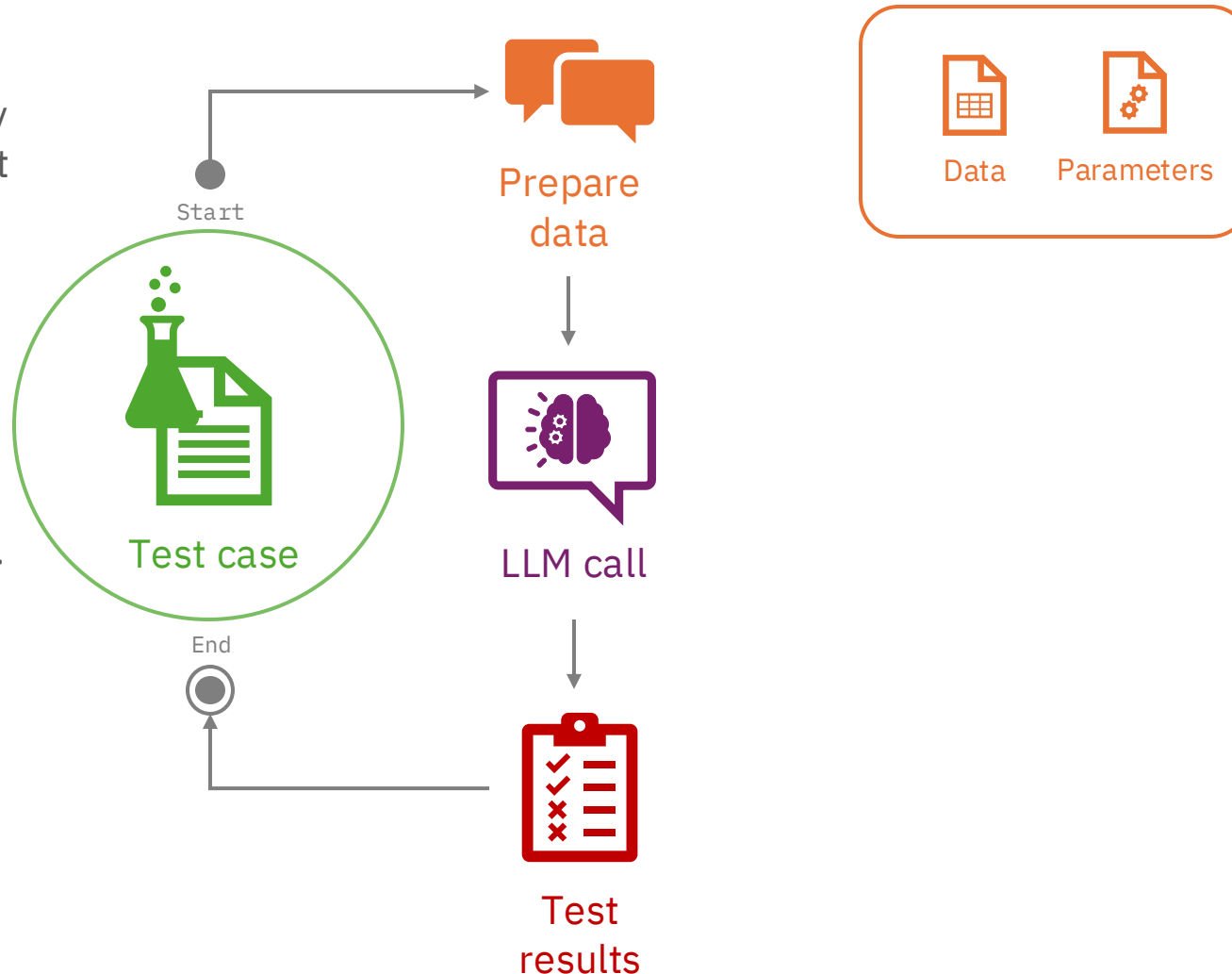
Test case

Test case

It's the core functionality that we can logically split into three steps.

A test case is generic and can be run for different combinations of data and parameters. Each step takes own *>configs* or *>parameters*.

The output of one step is *the input* to the next.



Prepare data

Takes *>test data* and prepare *the LLM prompt*.



The prompt lab *concept*

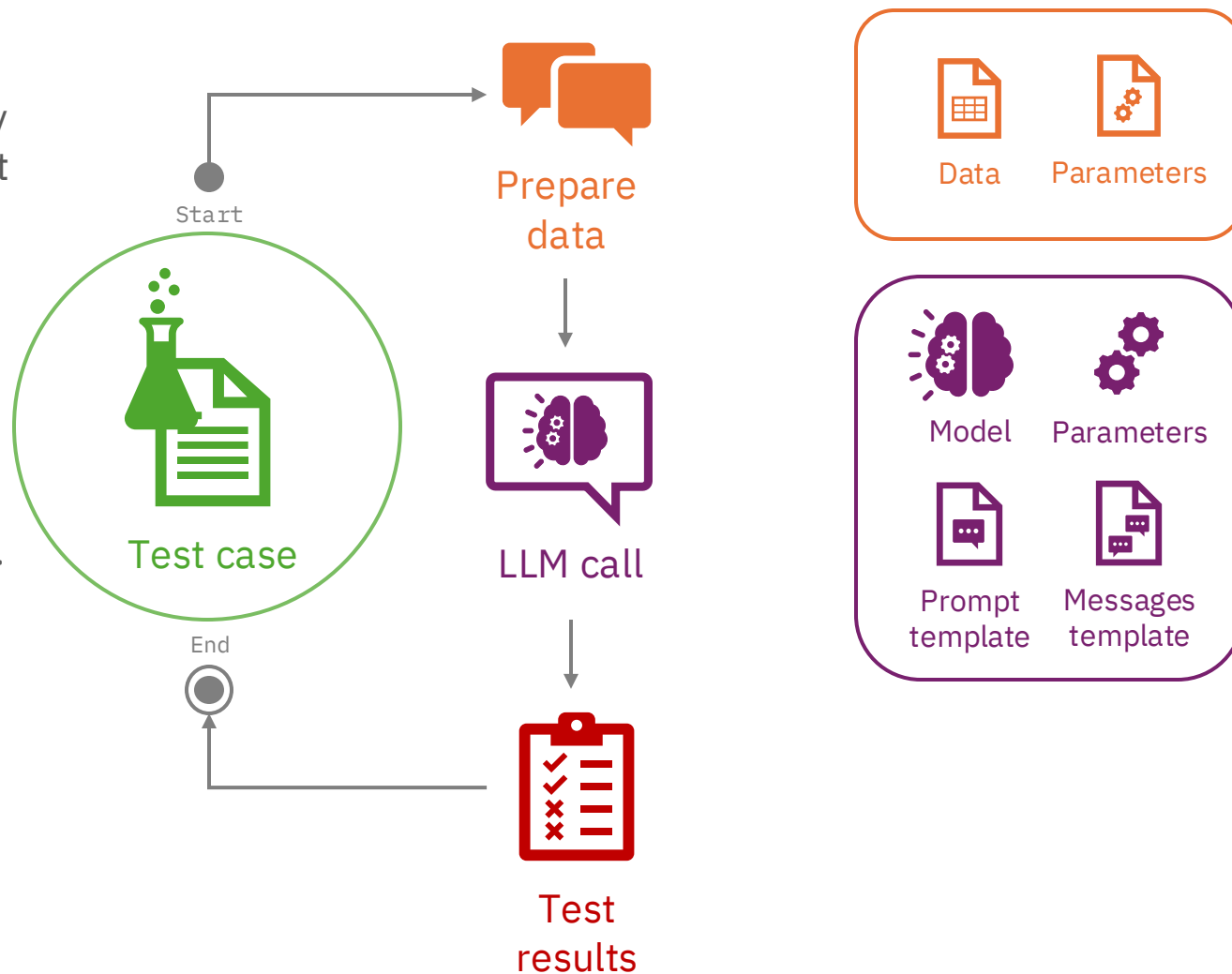
Test case

Test case

It's the core functionality that we can logically split into three steps.

A test case is generic and can be run for different combinations of data and parameters. Each step takes own *>configs* or *>parameters*.

The output of one step is *>the input* to the next.



Prepare data

Takes *>test data* and prepare *the LLM prompt*.

LLM call

With *>the LLM prompt* and *>the LLM config* calls LLM service. Parse *the LLM output*.



The prompt lab *concept*

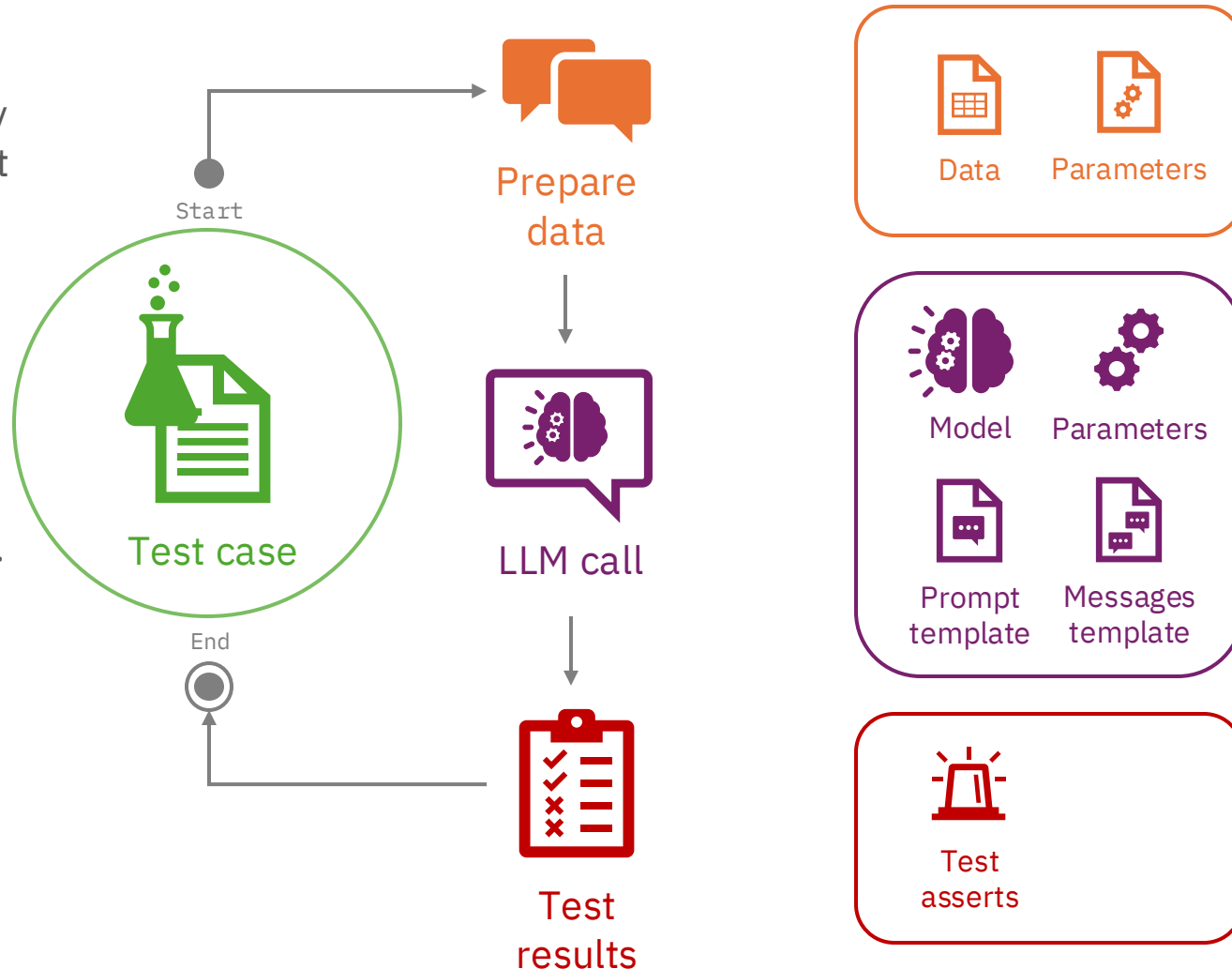
Test case

Test case

It's the core functionality that we can logically split into three steps.

A test case is generic and can be run for different combinations of data and parameters. Each step takes own *>configs* or *>parameters*.

The output of one step is *>the input* to the next.



Prepare data

Takes *>test data* and prepare *the LLM prompt*.

LLM call

With *>the LLM prompt* and *>the LLM config* calls LLM service. Parse *the LLM output*.

Test results

Takes *>the LLM output* against *>the test asserts*.



The prompt lab *concept*

Test case



Test input

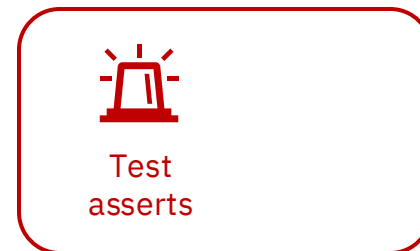
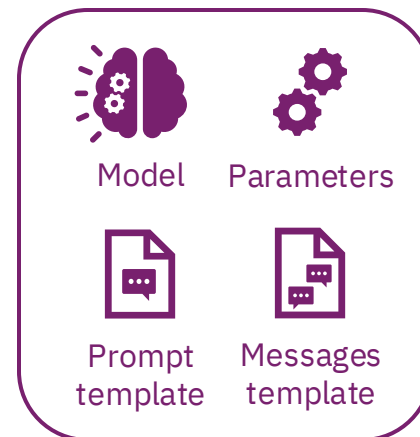
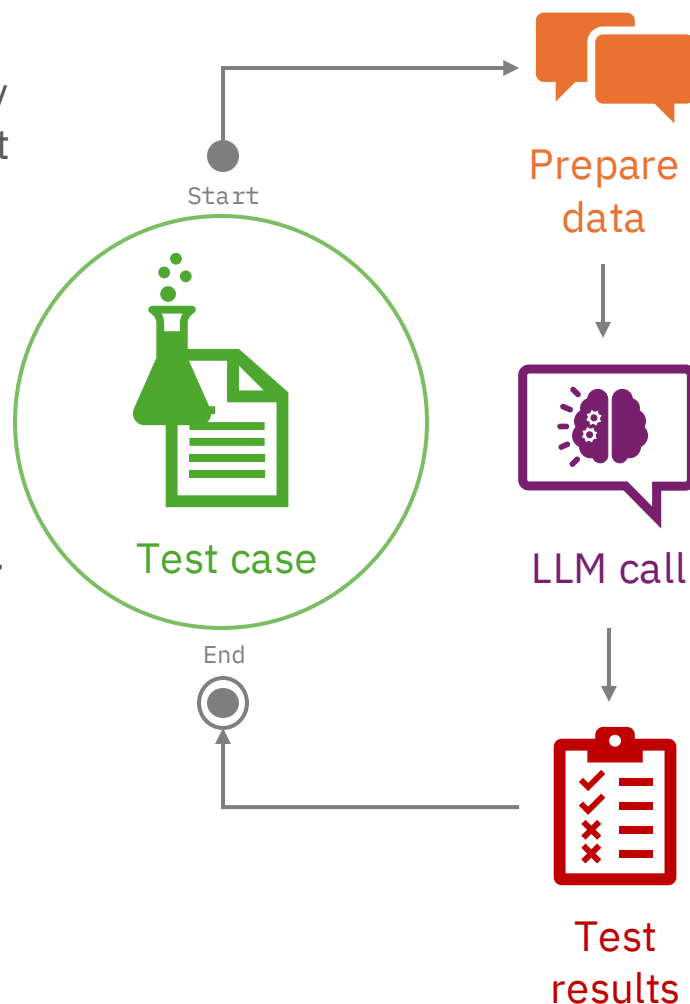
Test case

It's the core functionality that we can logically split into three steps.

A test case is generic and can be run for different combinations of data and parameters. Each step takes own *>configs* or *>parameters*.

The output> of one step is *>the input* to the next.

The combination of test data and parameters of each step is called **test input**.



Prepare data

Takes *>test data* and prepare *the LLM prompt>*.

LLM call

With *>the LLM prompt* and *>the LLM config* calls LLM service. Parse *the LLM output>*.

Test results

Takes *>the LLM output* against *>the test asserts*.



The prompt lab *concept*

Test



The prompt lab *concept*

Test

Test input

Because we want to run the same test case with different combinations of *>configs* and *>parameters*, we have abstracted them.



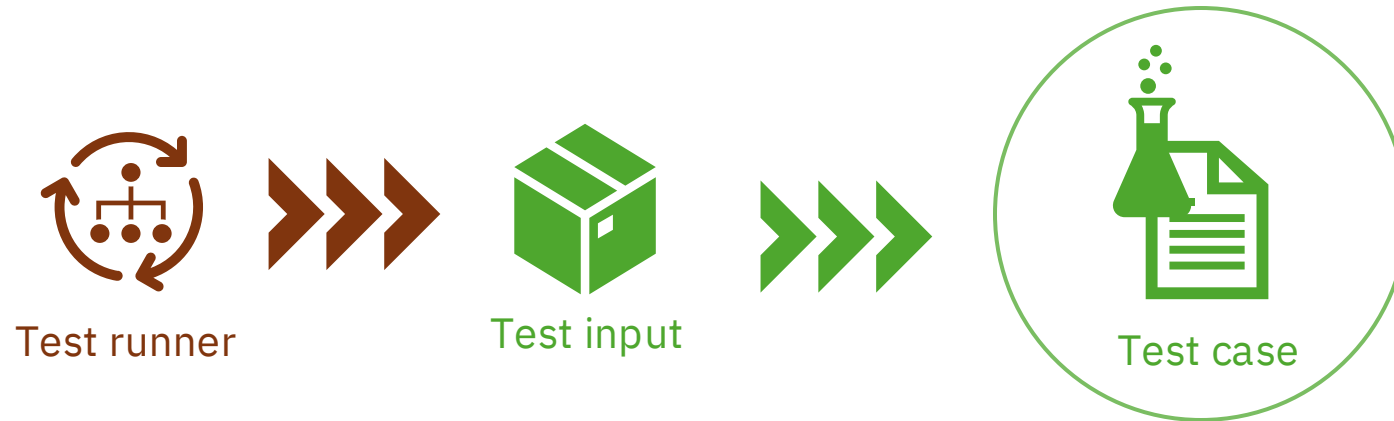


The prompt lab *concept*

Test

Test runner

It generates test input and runs a specific test case.





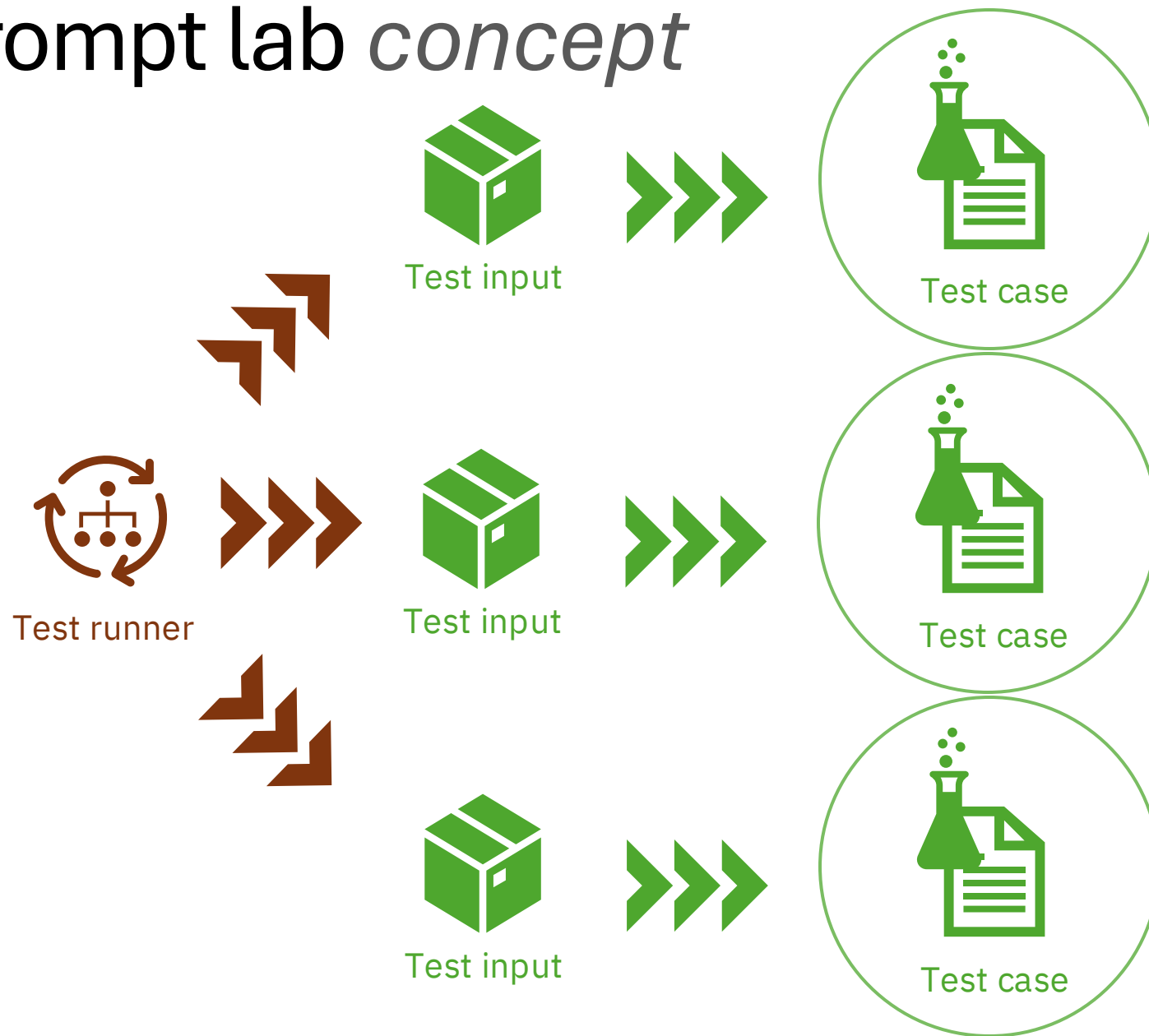
The prompt lab *concept*

Test

Test runner

It generates test input and runs a specific test case.

It generates multiple test inputs and runs them over the same test case.





The prompt lab *concept*

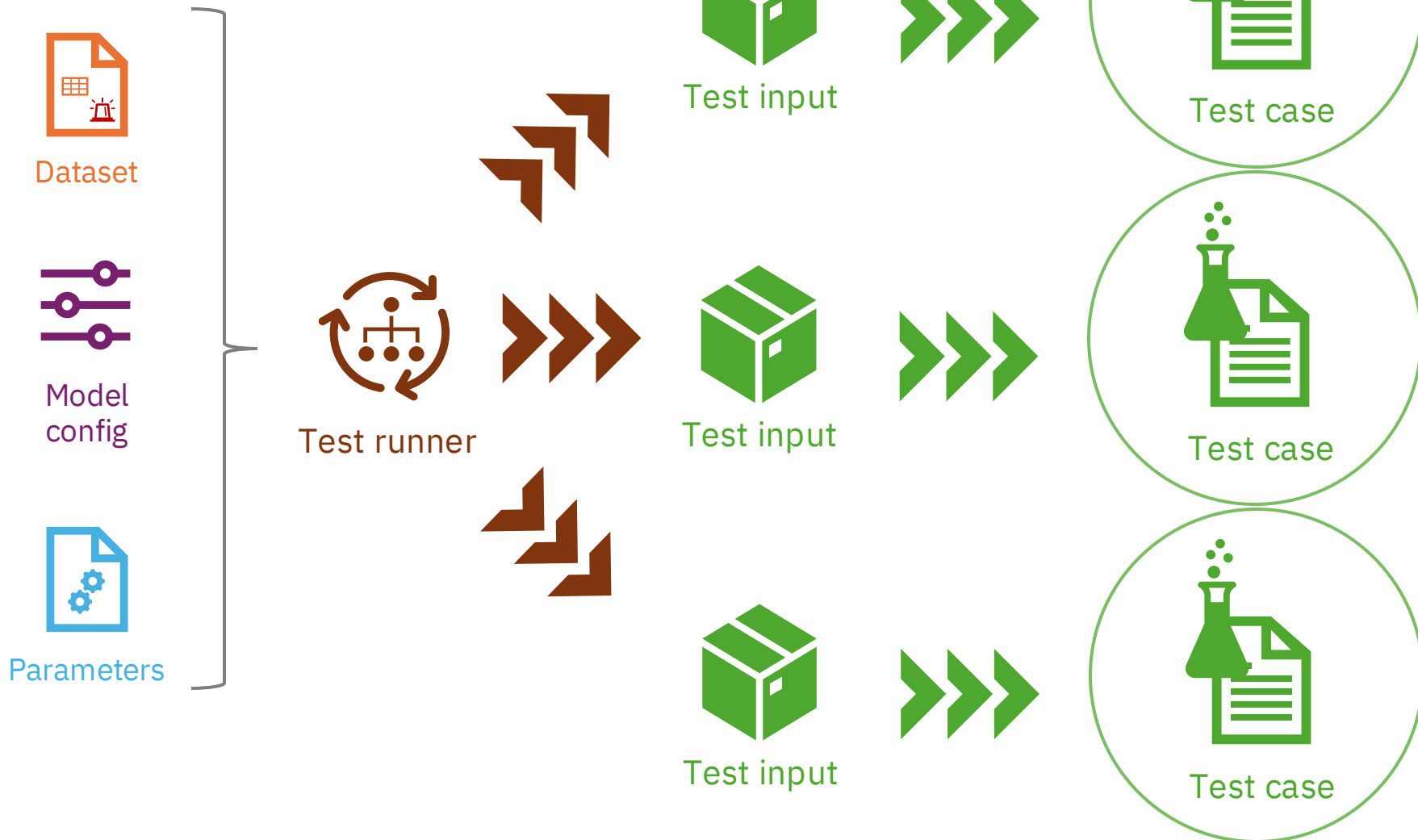
Test

Test runner

It generates test input and runs a specific test case.

It generates multiple test inputs and runs them over the same test case.

It receives three input objects *>dataset*, *>model config* and *>parameters*.





The prompt lab *concept*

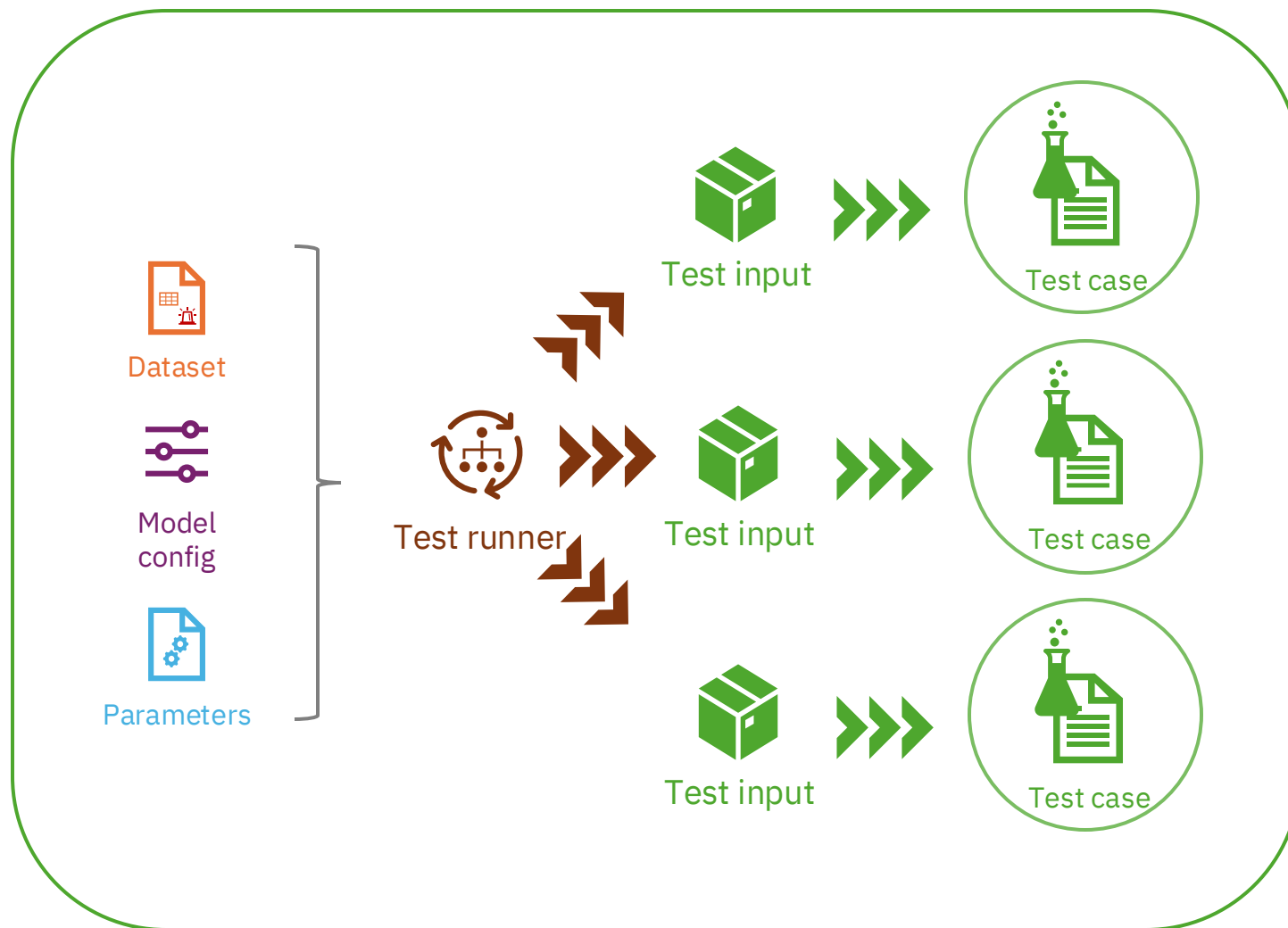
Test

Test

It is a code that takes *>dataset*,
>model config, *>parameters*,
>test case and runs them in a
test runner.



Test





The prompt lab *concept*

Project



The prompt lab *concept*

Project

Project

It's a structure where is well-defined contract between LLM and tests and wehere tests live.



Project



The prompt lab *concept*

Project

Project

It's a structure where is well-defined contract between LLM and tests and wehere tests live.



Project



Base

Base

It's the place where are defined the main code contract and the general reusable parts of the project.



The prompt lab *concept*

Project

Project

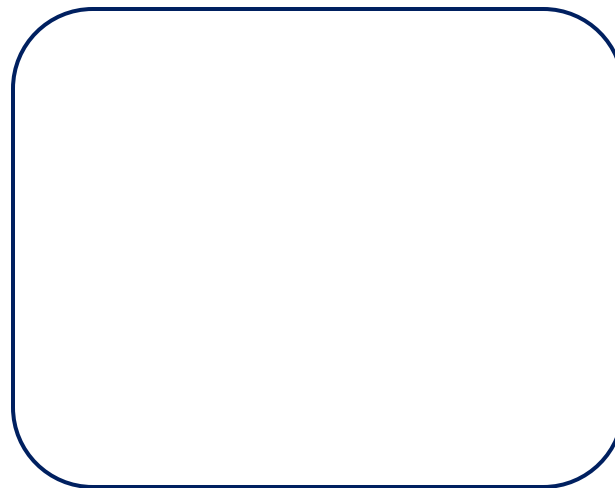
It's a structure where is well-defined contract between LLM and tests and wehere tests live.



Project



Base



Base

It's the place where are defined the main code contract and the general reusable parts of the project.



The prompt lab *concept*

Project

Project

It's a structure where is well-defined contract between LLM and tests and wehere tests live.

Base

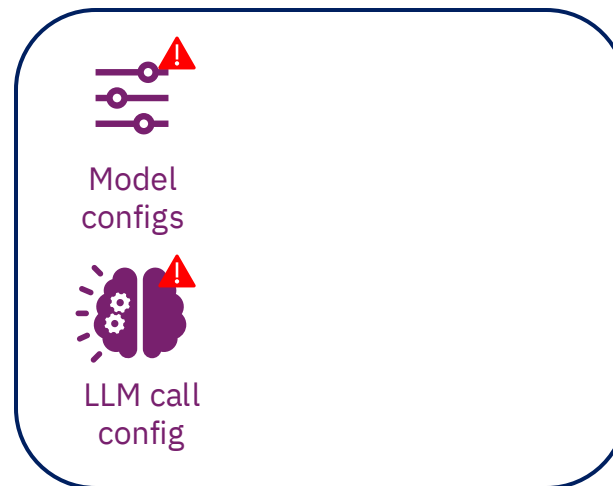
It's the place where are defined the main code contract and the general reusable parts of the project.



Project



Base



Crucial

LLM config defines the core communication concept. The prompt templates and the LLM output format are the most important contracts because test inputs and outputs depend on them.



The prompt lab *concept*

Project

Project

It's a structure where is well-defined contract between LLM and tests and wehere tests live.

Base

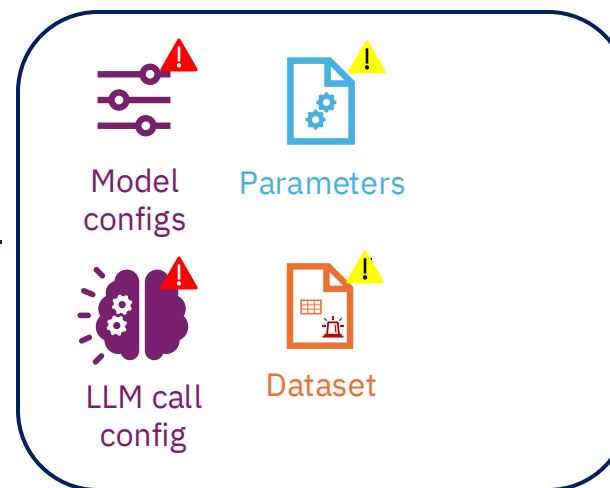
It's the place where are defined the main code contract and the general reusable parts of the project.



Project



Base



Crucial

LLM config defines the core communication concept. The prompt templates and the LLM output format are the most important contracts because test inputs and outputs depend on them.



Important

Dataset and **parameters** depend on **LLM config**. Define possibilities for test parameterization.



The prompt lab *concept*

Project

Project

It's a structure where is well-defined contract between LLM and tests and wehere tests live.

Base

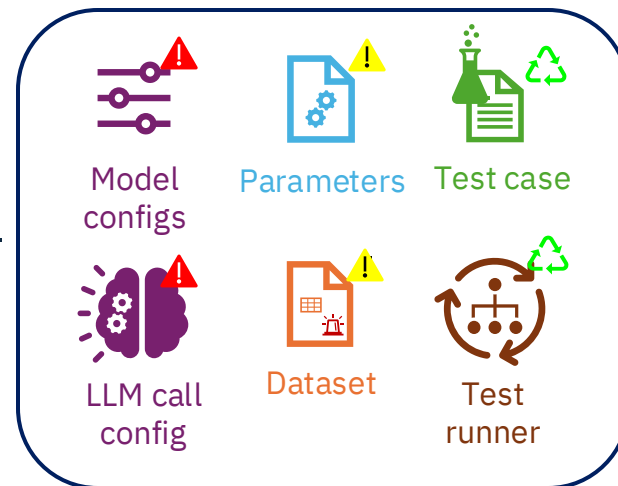
It's the place where are defined the main code contract and the general reusable parts of the project.



Project



Base



Crucial

LLM config defines the core communication concept. The prompt templates and the LLM output format are the most important contracts because test inputs and outputs depend on them.



Important

Dataset and **parameters** depend on **LLM config**. Define possibilities for test parameterization.



Reusable

Reusable parts of code.



The prompt lab *concept*

Project

Project

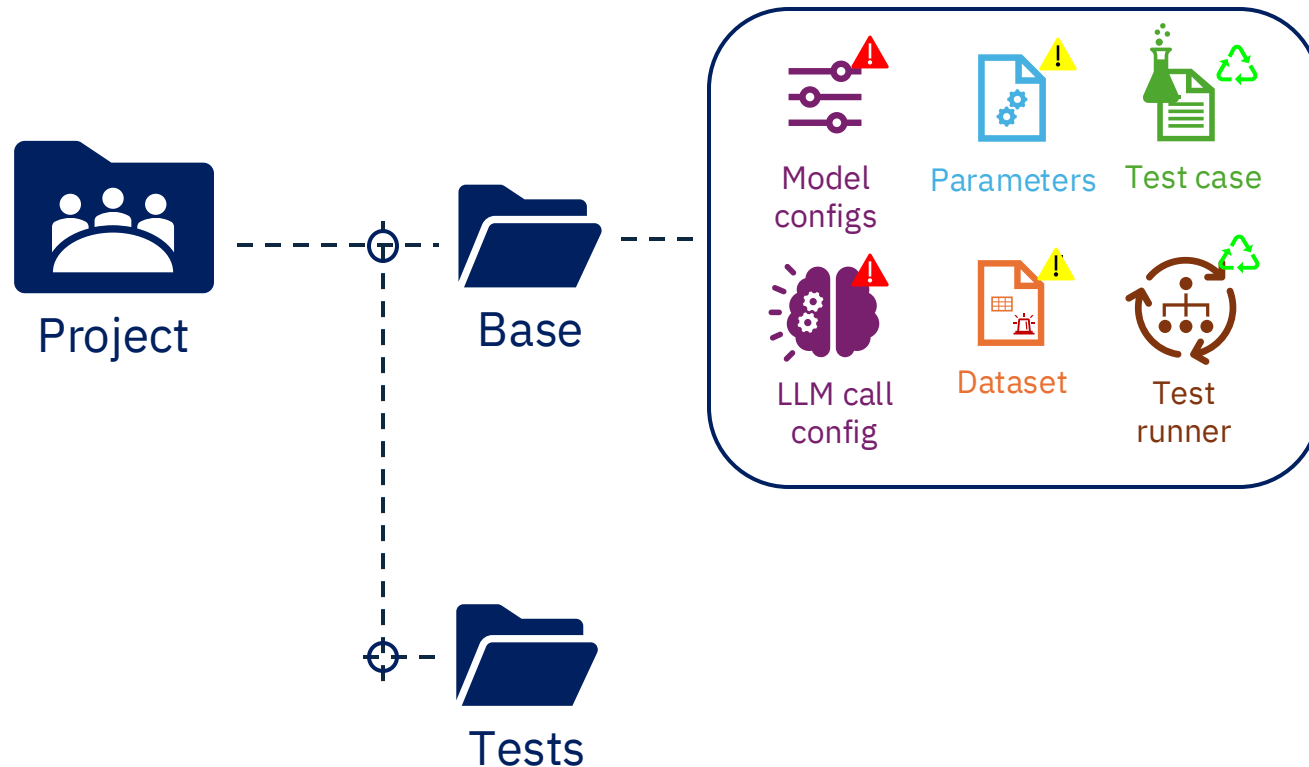
It's a structure where is well-defined contract between LLM and tests and wehere tests live.

Base

It's the place where are defined the main code contract and the general reusable parts of the project.

Tests

It's a container for the all project tests.



Crucial

LLM config defines the core communication concept. The prompt templates and the LLM output format are the most important contracts because test inputs and outputs depend on them.



Important

Dataset and **parameters** depend on **LLM config**. Define possibilities for test parameterization.



Reusable

Reusable parts of code.



The prompt lab *concept*

Project

Project

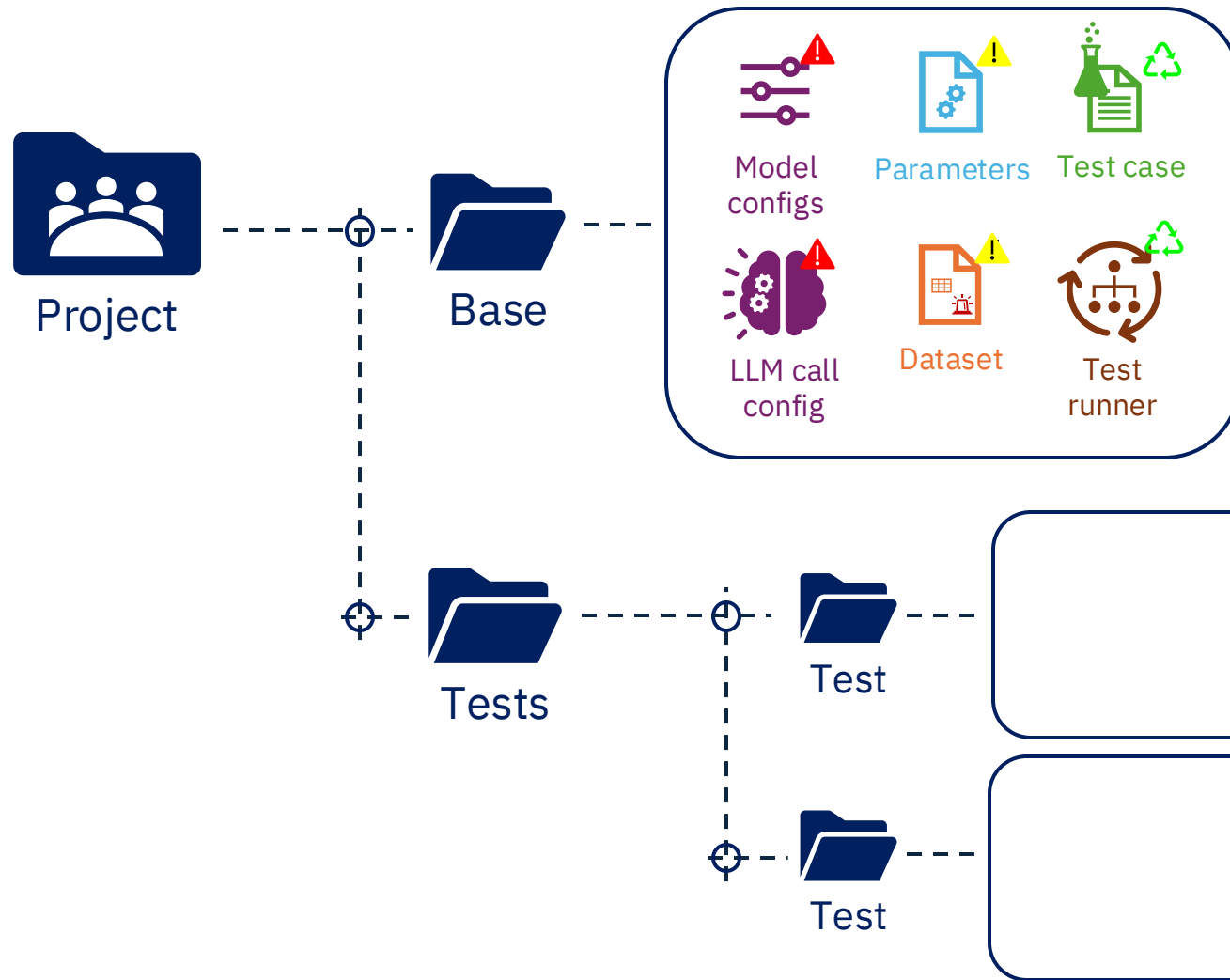
It's a structure where is well-defined contract between LLM and tests and wehere tests live.

Base

It's the place where are defined the main code contract and the general reusable parts of the project.

Tests

It's a container for the all project tests.



! Crucial

LLM config defines the core communication concept. The prompt templates and the LLM output format are the most important contracts because test inputs and outputs depend on them.

! Important

Dataset and **parameters** depend on **LLM config**. Define possibilities for test parameterization.

🔄 Reusable

Reusable parts of code.



The prompt lab *concept*

Project

Project

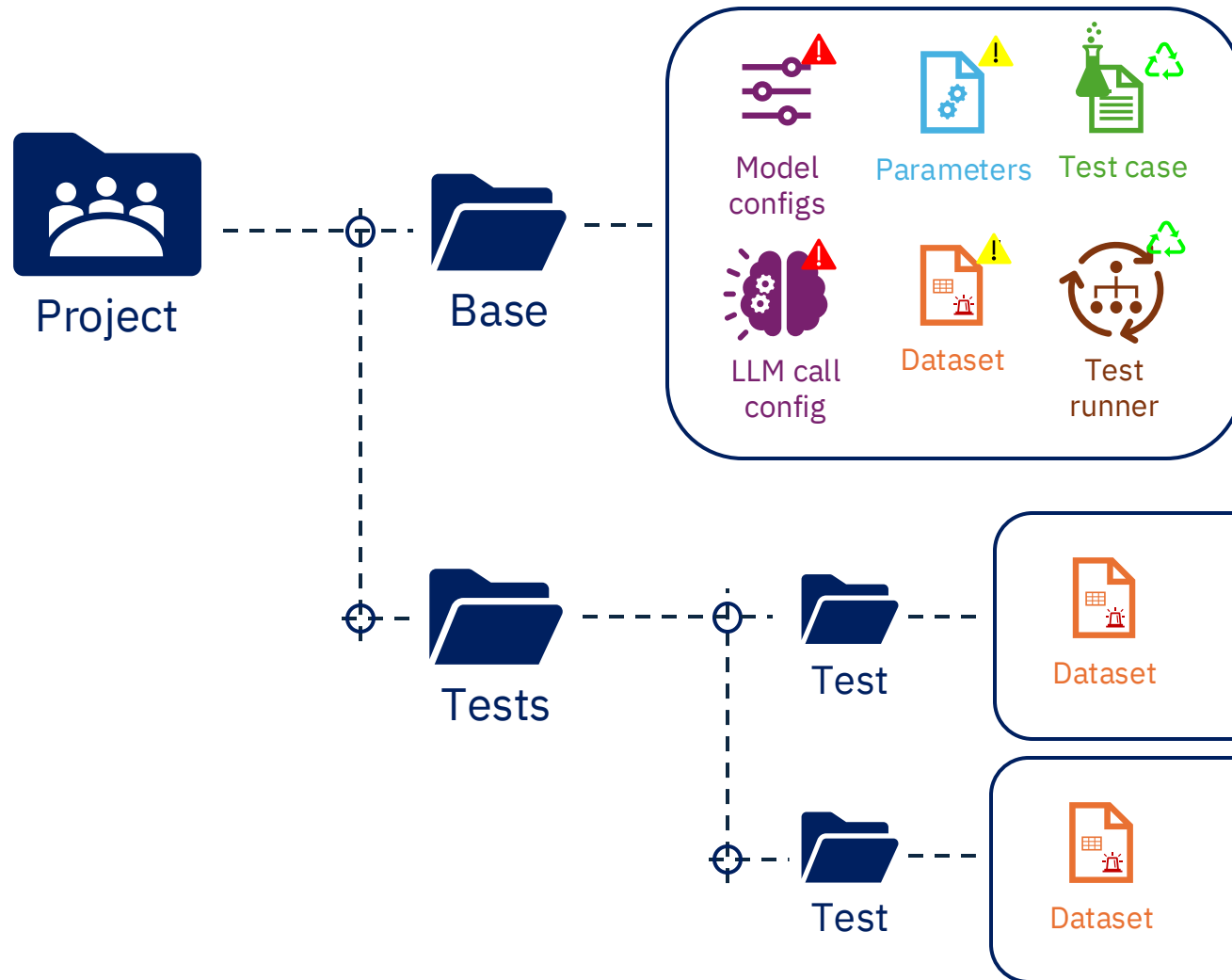
It's a structure where is well-defined contract between LLM and tests and wehere tests live.

Base

It's the place where are defined the main code contract and the general reusable parts of the project.

Tests

It's a container for the all project tests.



Crucial

LLM config defines the core communication concept. The prompt templates and the LLM output format are the most important contracts because test inputs and outputs depend on them.



Important

Dataset and **parameters** depend on **LLM config**. Define possibilities for test parameterization.



Reusable

Reusable parts of code.



The prompt lab *concept*

Project

Project

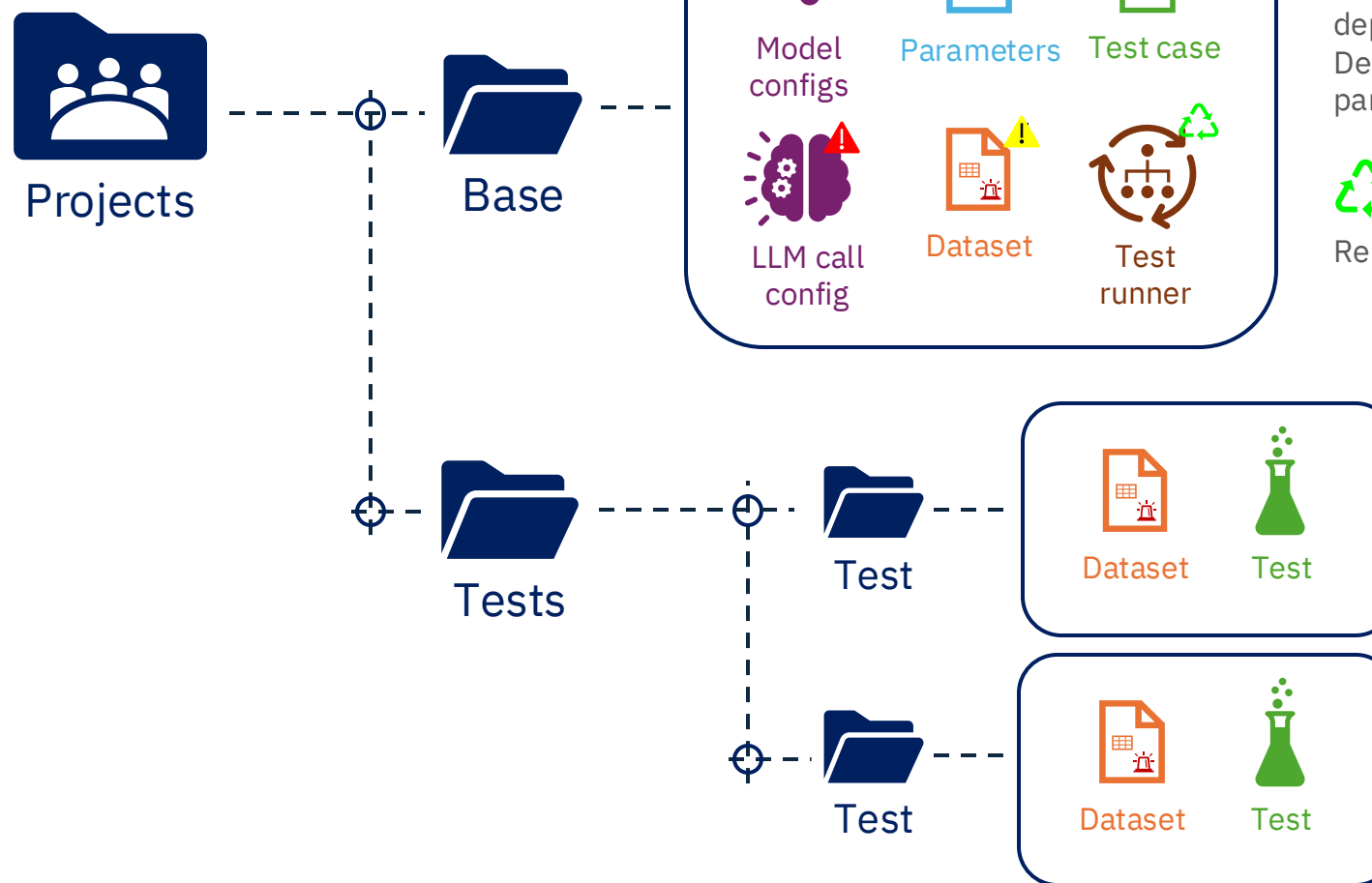
It's a structure where is well-defined contract between LLM and tests and wehere tests live.

Base

It's the place where are defined the main code contract and the general reusable parts of the project.

Tests

It's a container for the all project tests.



! Crucial

LLM config defines the core communication concept. The prompt templates and the LLM output format are the most important contracts because test inputs and outputs depend on them.

! Important

Dataset and **parameters** depend on **LLM config**. Define possibilities for test parameterization.

♻️ Reusable

Reusable parts of code.

The End