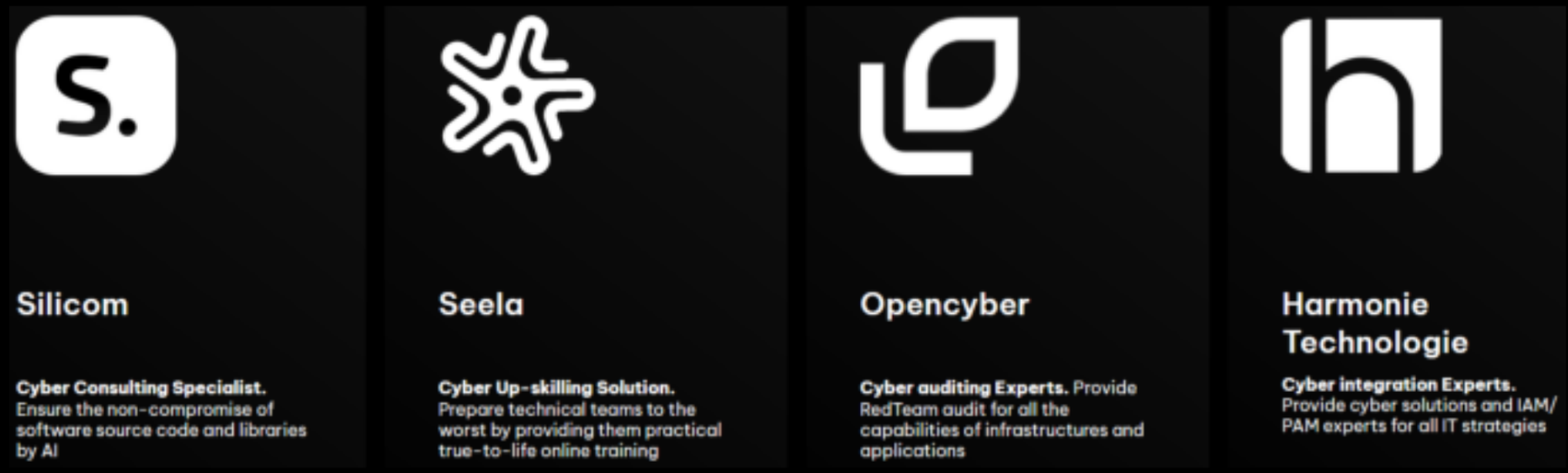
PR0PH3CY

https://www.pr0ph3cy.com/

C4PTION: Why Characterise the Authors of code PorTIONs?

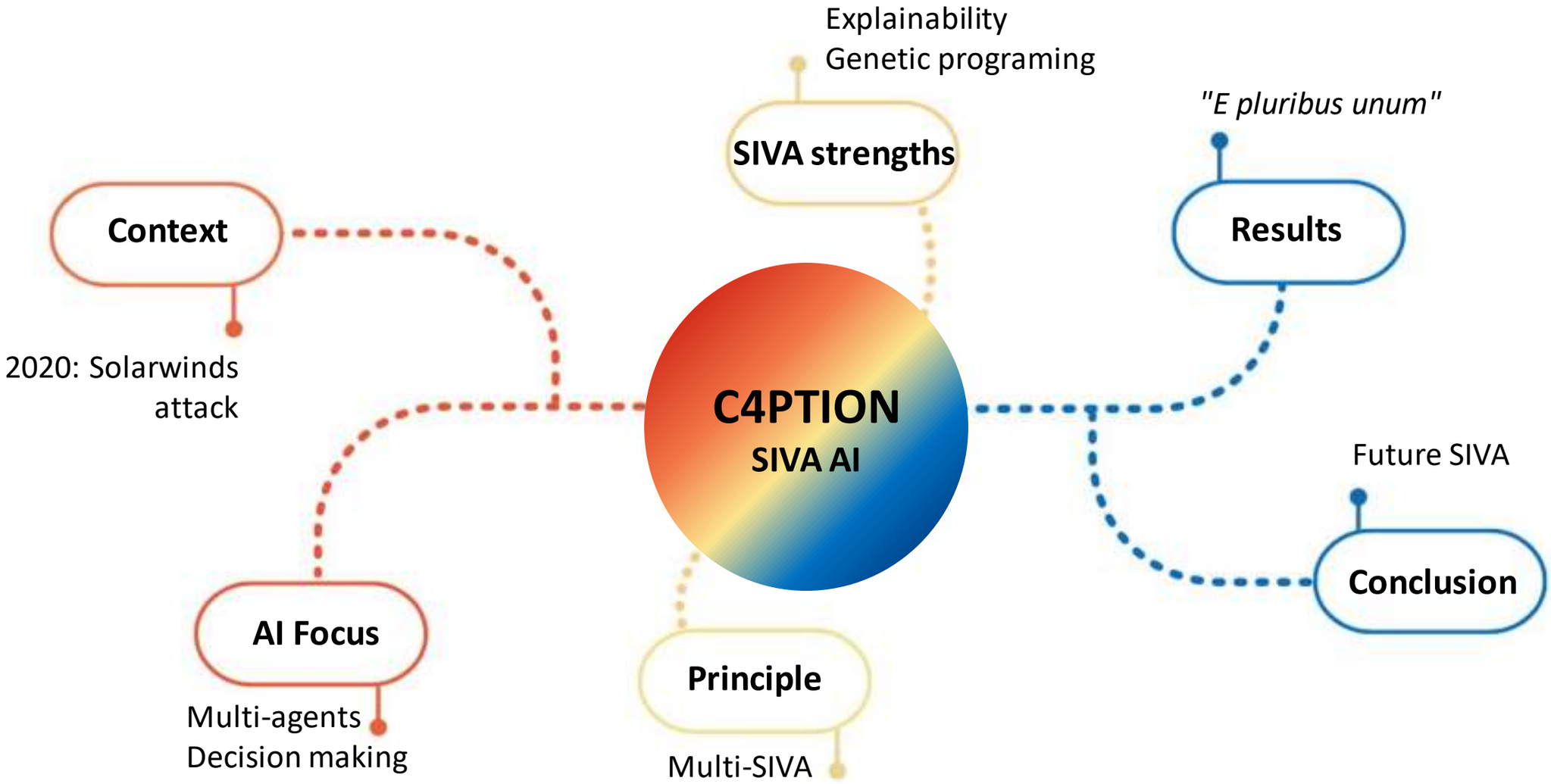
#Evolutionnist AI, #Explainable AI, #zero trust, #git source code, #git metadata, #CI/CD



Olivier Gesny - Head of Innovation Pr0ph3cy – ogesny@silicom.fr
Pierre-Marie SATRE – R&D Team – pmsatre@silicom.fr

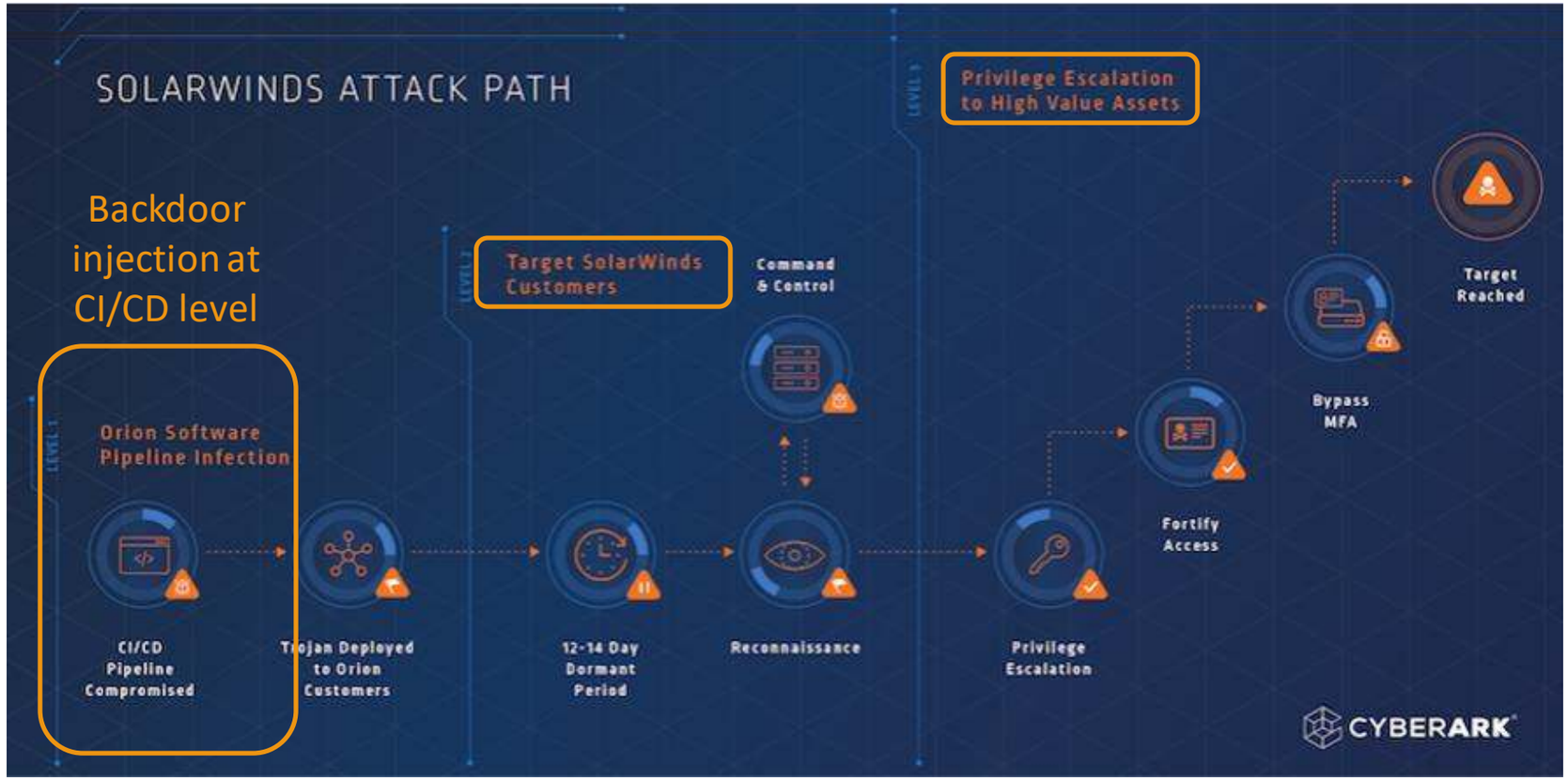
European Cyber Week Rennes 2022
Stands 21, 22 and 23

Table of contents



Context

Backdoor injection: Solarwinds attack



Source: https://www.cyberark.com/resources/blog/the-anatomy-of-the-solarwinds-attack-chain

Extreme sensitivity of the supply chain

2020-2021: Increase of supply chain attack by 650%

Serious impact on customer side:

- Data spying (Confidentiality)
- Data corruption (Integrity)
- Data destruction (Availability)

Serious impact on editor side:

- Loss of customer confidence
- Loss of corporate identity
- **Lawsuits**
- **Financial loss** ($12 millions = 11% of their annual revenue)

Context

Needs

- Innovative authentication systems in continuous integration pipelines (CI)
- Alerting system to avoid supply chain corruption
- Low false positive rate of alert

Context

State of the art

- *https://arxiv.org/pdf/2202.06043.pdf*: Deep learning and data augmentation to reduce the rate of attack success by obfuscating authors of code
- *https://arxiv.org/pdf/2008.13768.pdf*: Use of TF-IDF to attribute Android code to their author
- *https://arxiv.org/pdf/2007.00772.pdf*: Improving performances of classification and speeding up learning

» No known tool to authenticate commit of a developer
in a
gitlab Continuous Integration pipeline (CI)

Context

Contribution of C4PTION

- Contribution of a **multi agent** verdict on a combined analysis of source code and git meta data
- Learns **syntactic, lexical** and **behavioural** habits of developers
- **Automation** in a continuous integration gitlab pipeline
- Increased supply chain's security by alerting project manager in case of **suspect commit's authorship**

Context

C4PTION Solution

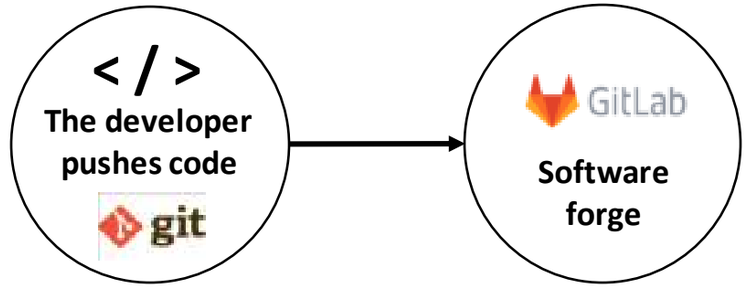
Context

C4PTION Solution

< / >
The developer pushes code

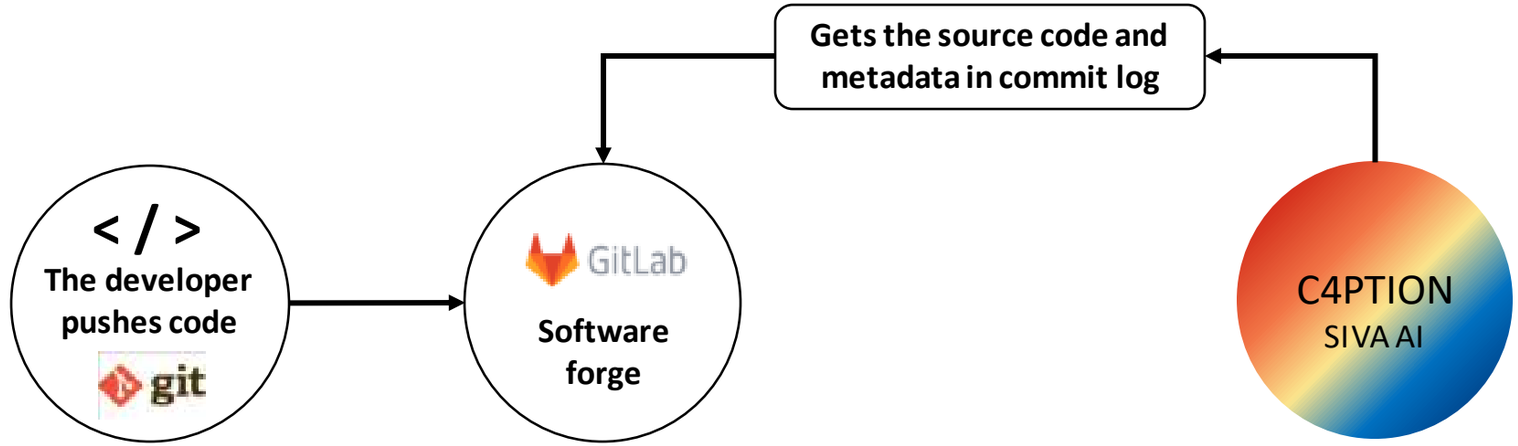
Context

C4PTION Solution



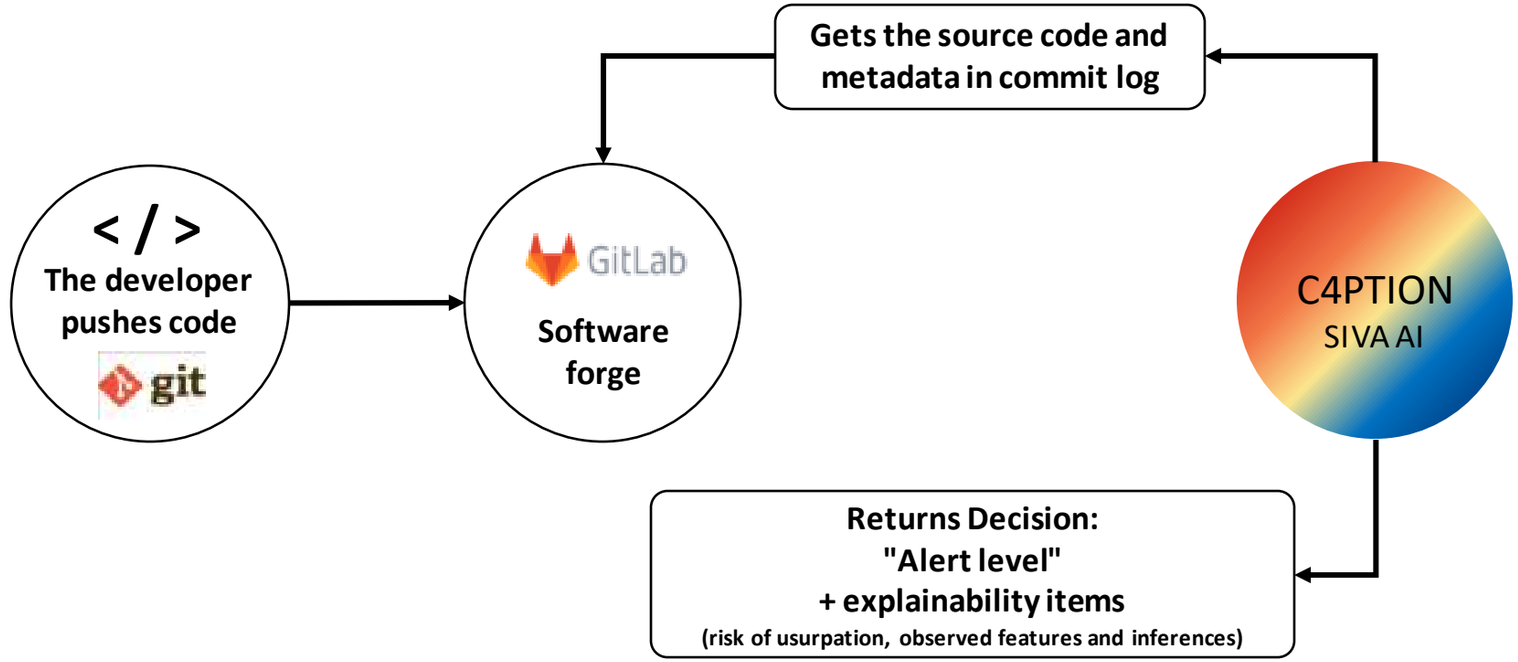
Context

C4PTION Solution



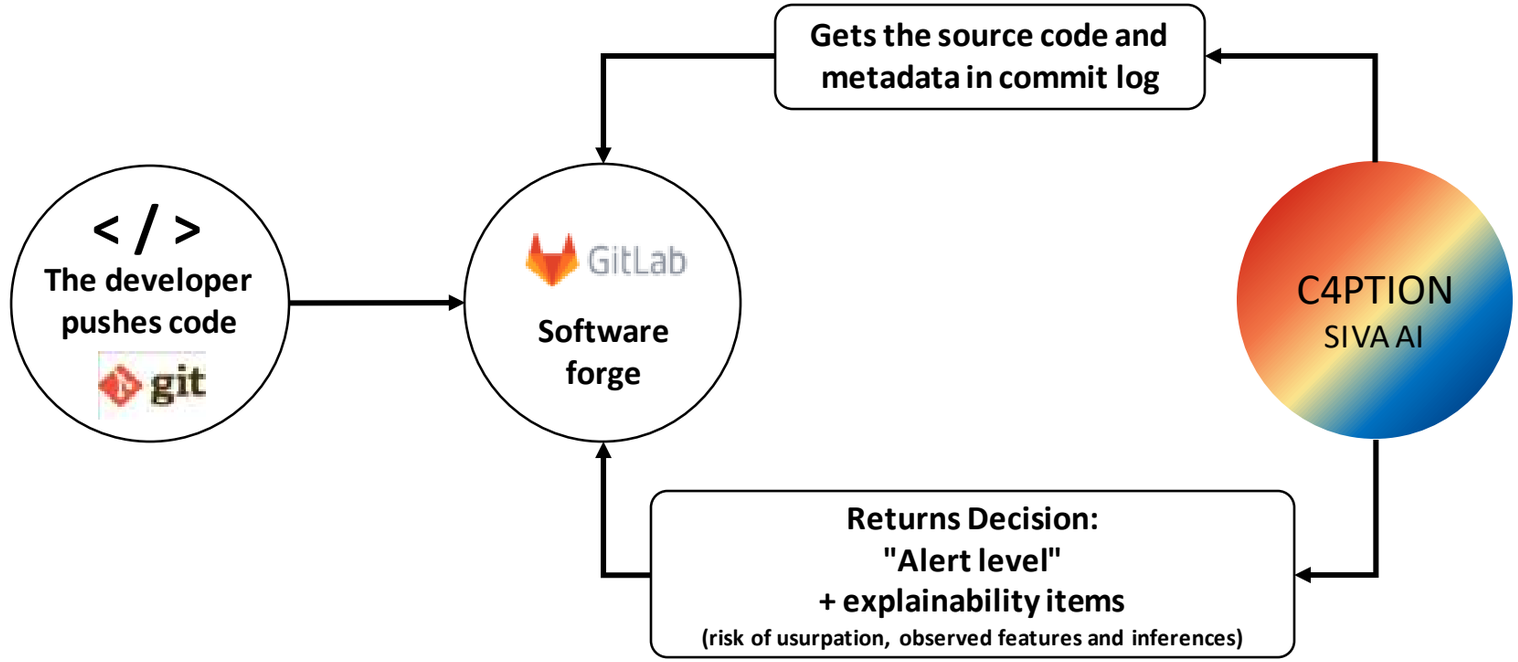
Context

C4PTION Solution



Context

C4PTION Solution

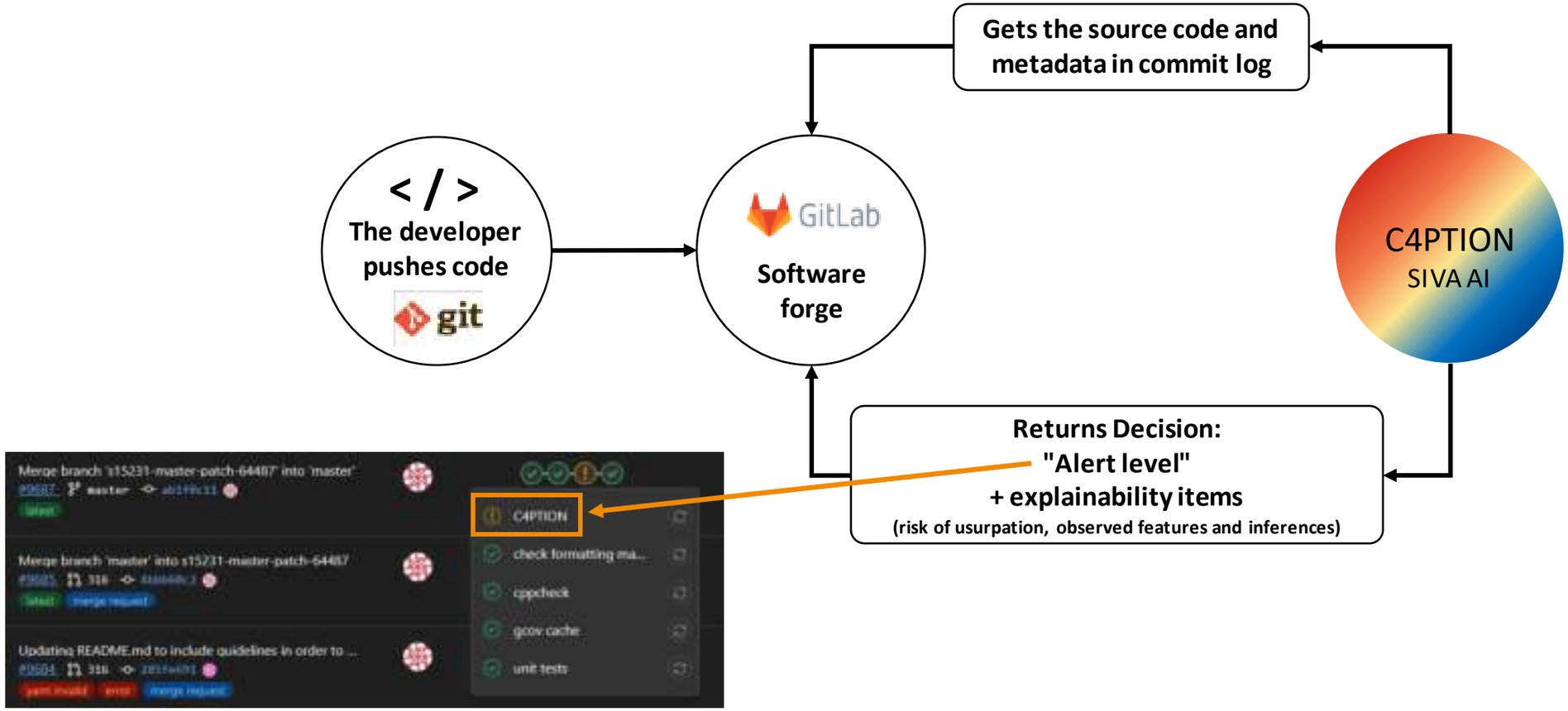


Configurable confidence thresholds of alerts :

- High risk
- Medium risk
- Neutral risk

Context

C4PTION Solution

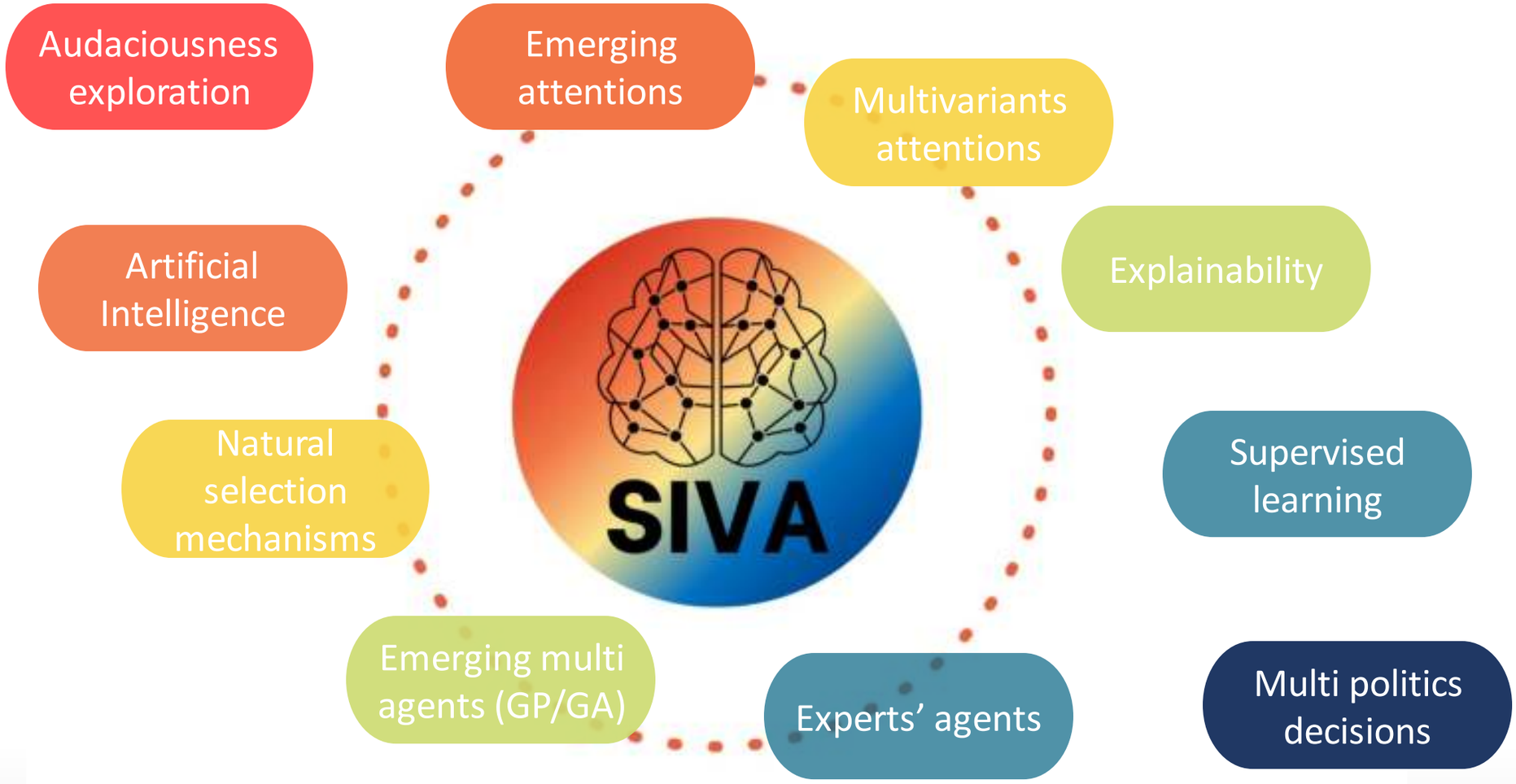


Configurable confidence thresholds of alerts :

- High risk
- Medium risk
- Neutral risk

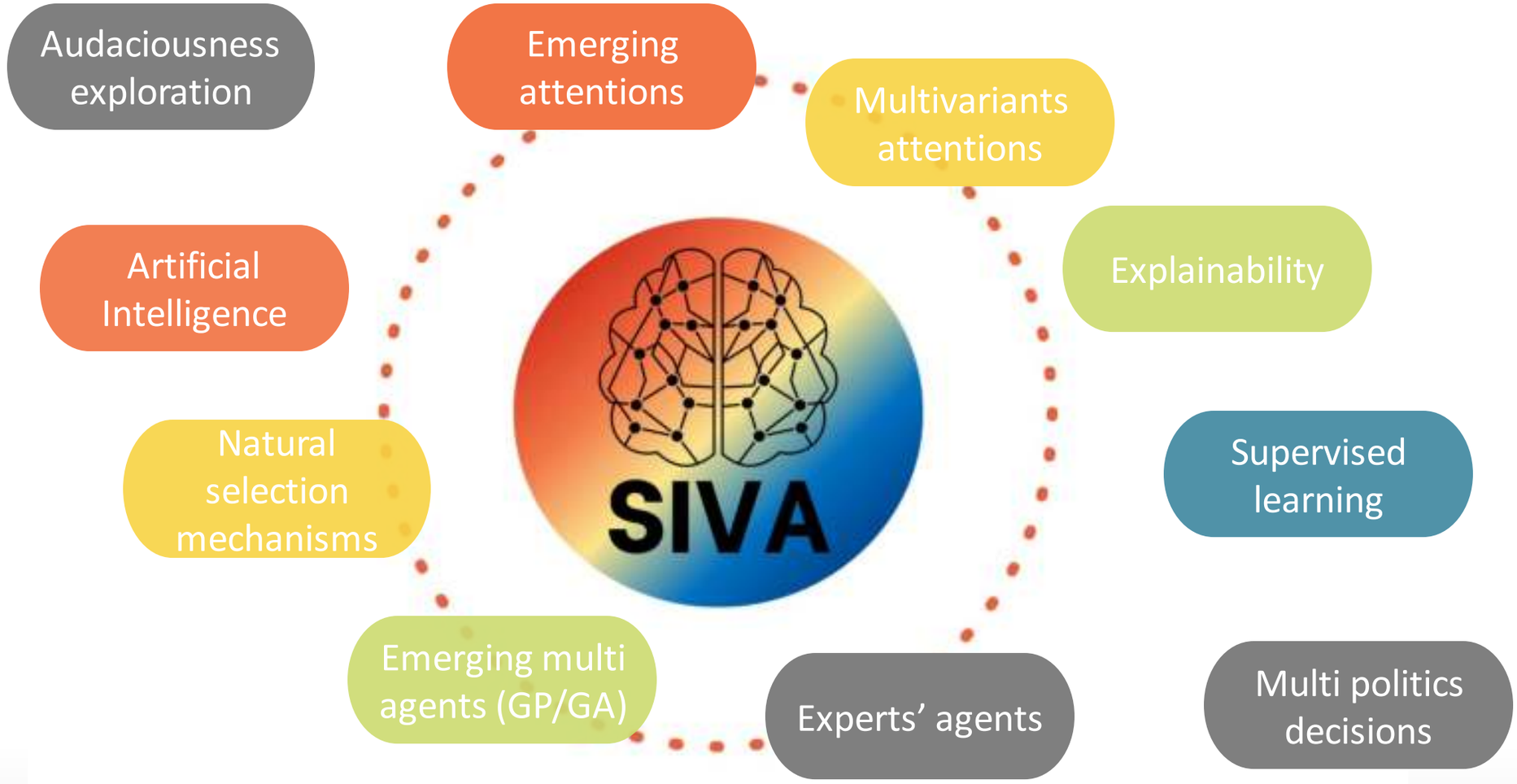
AI Focus

SIVA AI in a few words



AI Focus

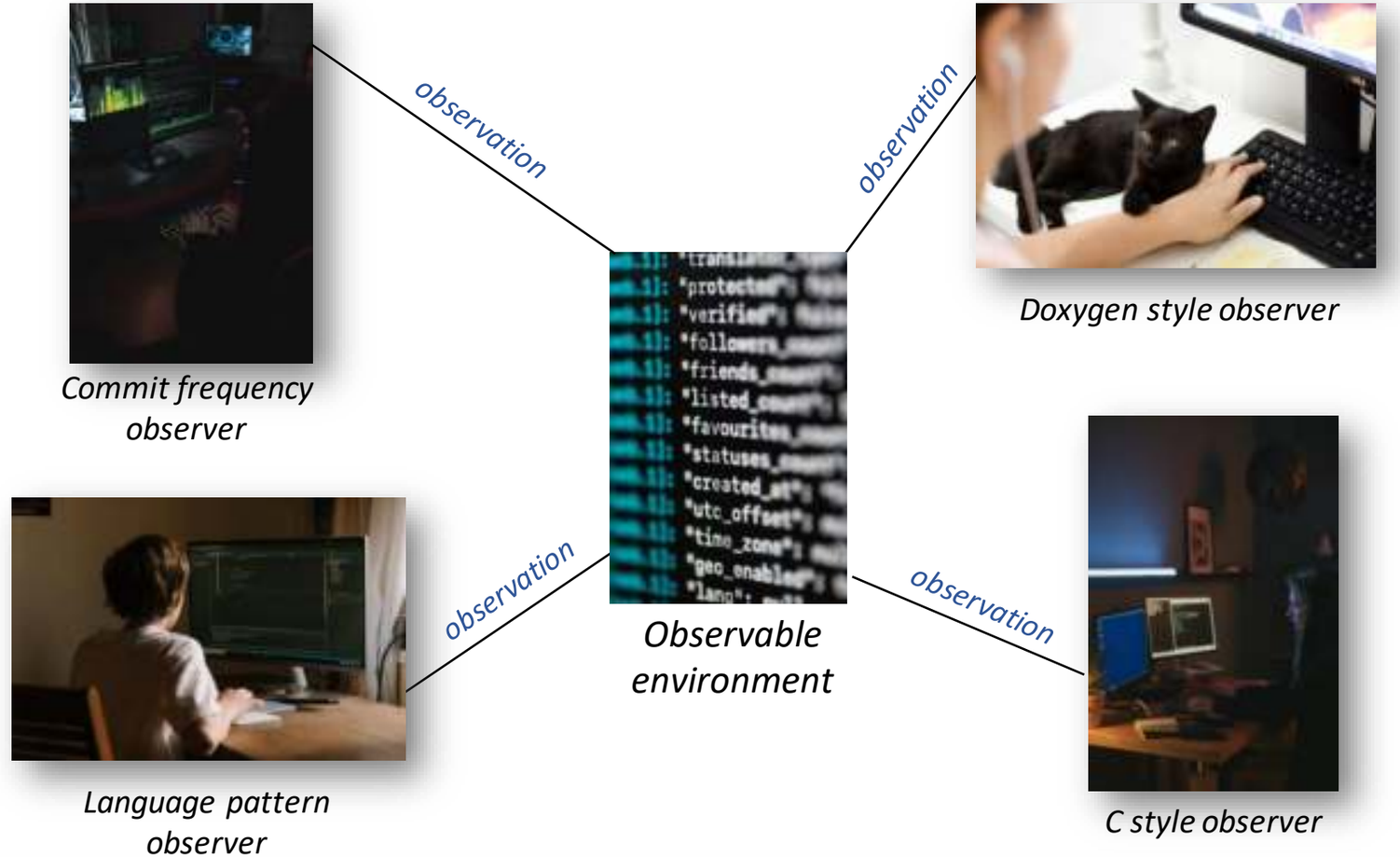
SIVA AI in a few words



AI Focus

General description of SIVA

- Ensemble of emerging specialist observers
- Agents give their **opinion** only on the base of their observation specialty, like a sensor
- Every agent's opinion is taken into account, weighted by their **respective** experience

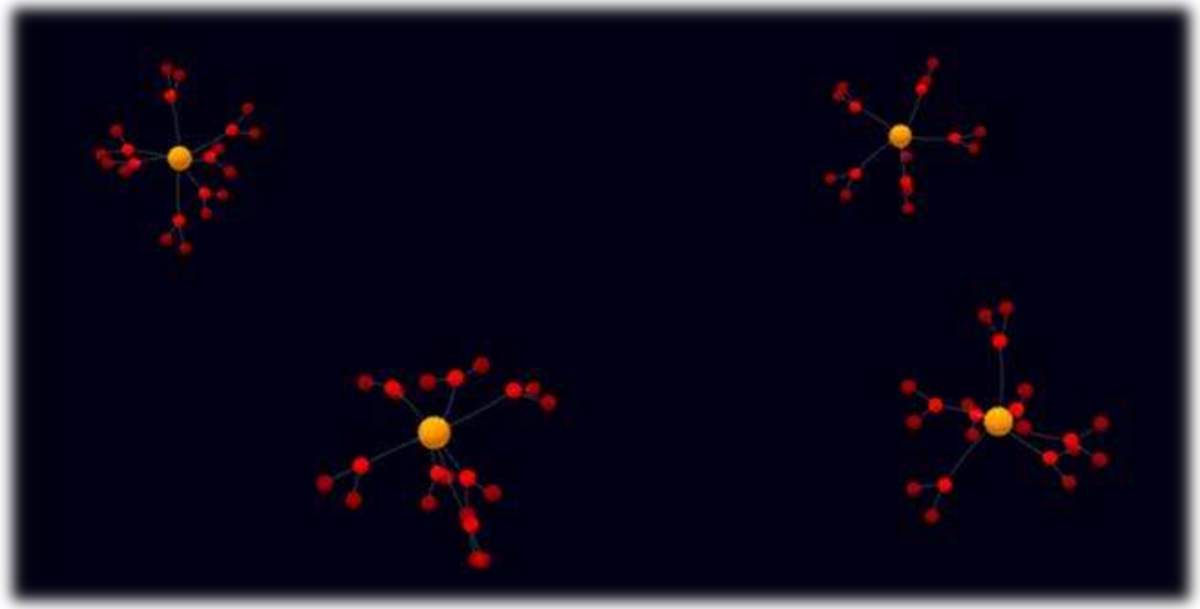


AI Focus

Agent creation from scratch

Emergence through agent creation:

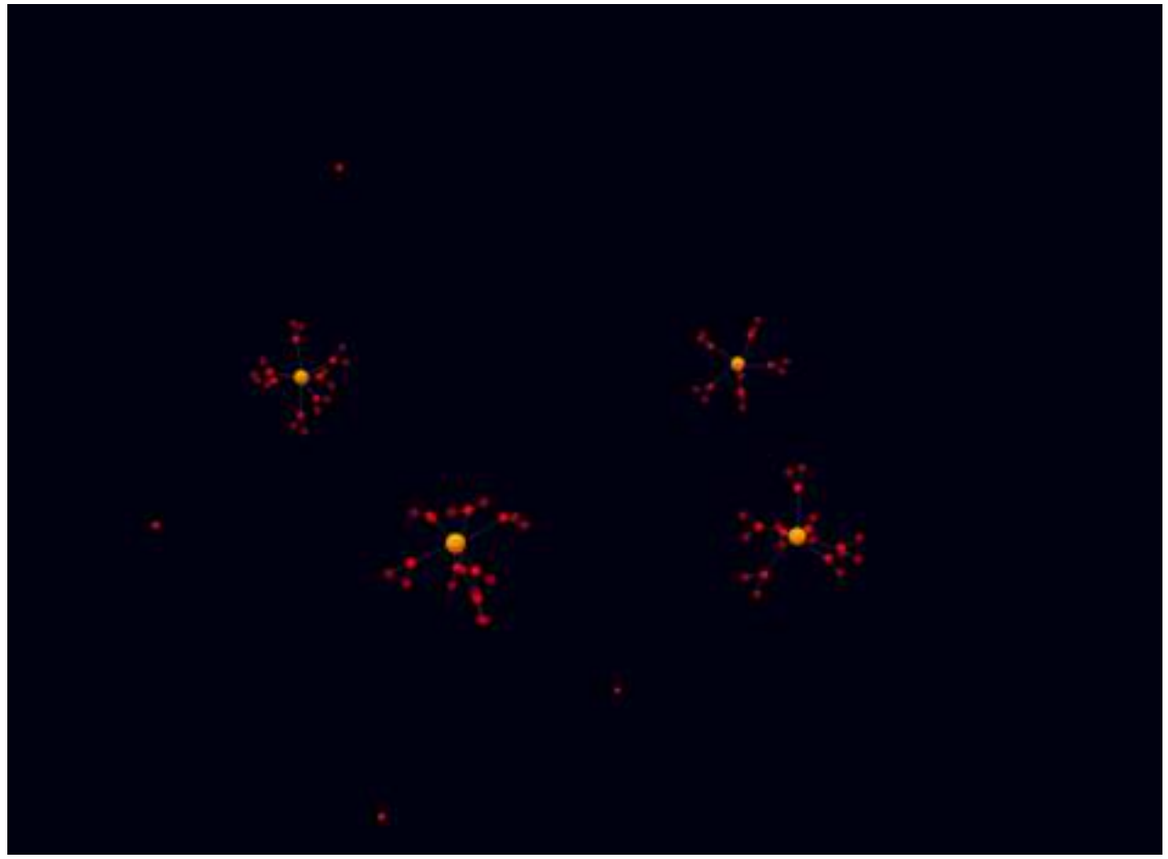
- Random choice of type of agent
- Property of observation emerges depending on the type of the agent
- Life experiment is empty at agent creation. Experiment is acquired through time



Agents as training starts

AI Focus

Genetic creation / mutation of groups of agents during training



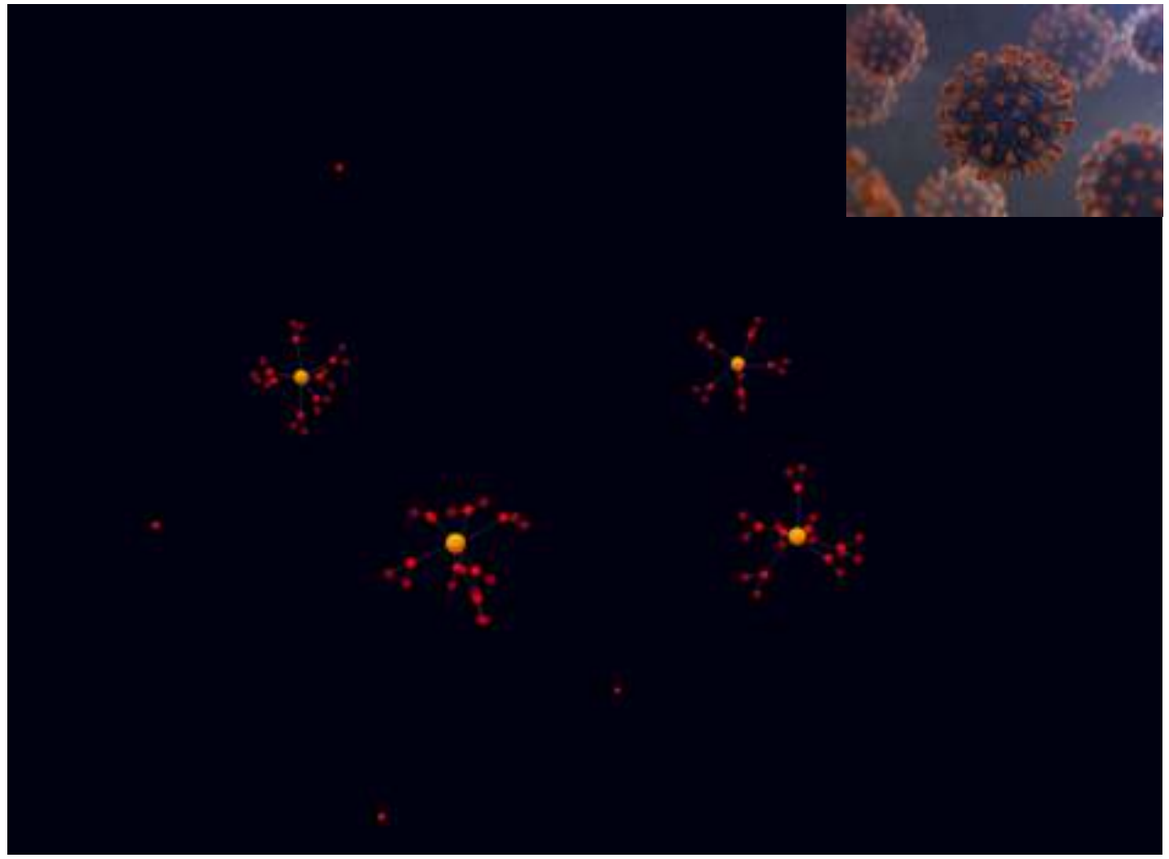
Expansion through training

SIVA possesses genetic components close to those of viruses which are exploited in its exploratory system:

- Capacity of erasing populations of least performing agents
- Capacity of cloning and mutating best populations of agents

AI Focus

Genetic creation / mutation of groups of agents during training



https://www.franceculture.fr/sciences/comprendre-les-variants-du-coronavirus-en-5-questions

Expansion through training

SIVA possesses genetic components close to those of viruses which are exploited in its exploratory system:

- Capacity of erasing populations of least performing agents
- Capacity of cloning and mutating best populations of agents

AI Focus

Agent creation

Several correlated criterias define a SIVA's agent:

- **Observed property** in the environment
- Life experience:
  - **Action$_i$**
  - **Reward$_i$**
  - **Experience$_i$**

AI Focus

Agent creation

Several correlated criterias define a SIVA's agent:

- **Observed property** in the environment
- Life experience:
  - **Action$_i$**
  - **Reward$_i$**
  - **Experience$_i$**

Agent example

Theoretical component	Concrete example
Observation of property	Specific pattern of comment: *"as if"*

AI Focus

Agent creation

Several correlated criterias define a SIVA's agent:

- **Observed property** in the environment
- Life experience:
  - **Action$_i$**
  - **Reward$_i$**
  - **Experience$_i$**

Agent example

Theoretical component	Concrete example
Observation of property	Specific pattern of comment: *"as if"*
Correlation: Action / reward/ experience n°1	Action Édouard/ reward 1 / **57**

AI Focus

Agent creation

Several correlated criterias define a SIVA's agent:

- **Observed property** in the environment
- Life experience:
  - **Action$_i$**
  - **Reward$_i$**
  - **Experience$_i$**

Agent example

Theoretical component	Concrete example
Observation of property	Specific pattern of comment: *"as if"*
Correlation: Action / reward/ experience n°1	Action Édouard/ reward 1 / **57**
Correlation: Action / reward/ experience n°2	Action Louise/ reward -1 / **12**

Agent creation

Several correlated criterias define a SIVA's agent:

- **Observed property** in the environment
- Life experience:
  - **Action$_i$**
  - **Reward$_i$**
  - **Experience$_i$**

Agent example

Theoretical component	Concrete example
Observation of property	Specific pattern of comment: *"as if"*
Correlation: Action / reward/ experience n°1	Action Édouard/ reward 1 / **57**
Correlation: Action / reward/ experience n°2	Action Louise/ reward -1 / **12**
Correlation: Action / reward/ experience n°3	Action Mathieu / reward -1 / **1**

AI Focus

Types of agent

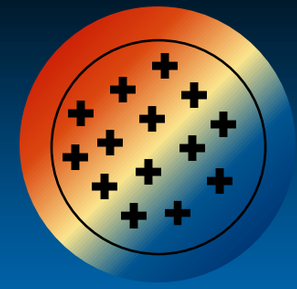
The type of an agent corresponds to the type of property it observes:

- Classification
  - Lexical style: patterns, vocabulary, comments
  - Syntactic style: doxygen style, indentation
  - Developers habits: files modified, hours of commit
- Normality

Lexical style

Vocabulary and lexical field specific to a developer
(ex: languages primitives [int, uint32_t, ==], used vocabulary)

Syntactic style



Form of code specific to a developer (ex: number of spaces and tab, position of brackets, camel case or snake case, variables size, size of comments)

Developers habits

Usual behaviours of a developer (ex: hours of commit, usual days of commits, usual time zone)

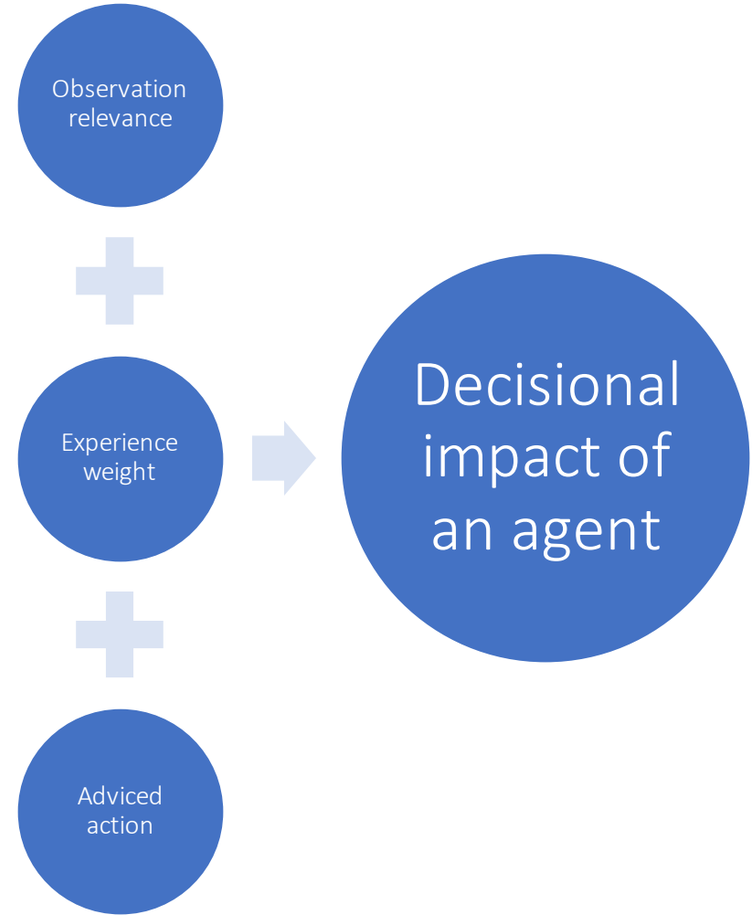
AI Focus

Decision making

Siva decision making is the resulting of its agents contributions:

- Listens only the **relevant** agents in their **observation**
- Various **weights** are granted to the agents depending on their **experience**
- Decisions are taken in accordance with advices of all the executed and **relevant** agents

→ **Random** decision if there is default of confidence in the agents in order to promote exploration



AI Focus

Explainability

Undeniable advantage of SIVA: **explainability**

The SIVA decisions are explainable thanks to:

- Concrete and suggestive **observations** highlighting decision's discriminants
- **Partial activation** of the graph for the decision
- Explicit level of **confidence** for each decision

AI Focus

Explainability

Agents

Property: goto
Observed
Author: **Édouard;** Reward **1;** Experience **: 36**
Author: **Louise;** Reward **-1;** Experience **: 36**
Author: **Mathieu;** Reward **-1;** Experience **: 36**

1

AI Focus

Explainability

Agents

Property: goto
Observed
Author: **Édouard;** Reward **1;** Experience **: 36**
Author: **Louise;** Reward **-1;** Experience **: 36**
Author: **Mathieu;** Reward **-1;** Experience **: 36**

1

Property: doxygen comment style
Not observed
Author: **Louise;** Reward **1;** Experience **: 50**
Author: **Louise;** Reward **-1;** Experience **: 150**
Author: **Mathieu;** Reward **-1;** Experience **: 200**
Author: **Édouard;** Reward **1;** Experience **: 150**
Author: **Édouard;** Reward **-1;** Experience **: 50**

2

AI Focus

Explainability

Agents

Property: goto
Observed
Author: **Édouard;** Reward **1;** Experience **: 36**
Author: **Louise;** Reward **-1;** Experience **: 36**
Author: **Mathieu;** Reward **-1;** Experience **: 36**
1

Property: doxygen comment style
Not observed
Author: **Louise;** Reward **1;** Experience **: 50**
Author: **Louise;** Reward **-1;** Experience **: 150**
Author: **Mathieu;** Reward **-1;** Experience **: 200**
Author: **Édouard;** Reward **1;** Experience **: 150**
Author: **Édouard;** Reward **-1;** Experience **: 50**
2

Property: 4 spaces indent before '}'
Observed
Author: **Mathieu;** Reward **1;** Experience **: 500**
Author: **Louise;** Reward **-1;** Experience **: 500**
Author: **Édouard;** Reward **-1;** Experience **: 500**
3

AI Focus

Explainability

Agents

Property: goto
Observed
Author: **Édouard;** Reward **1;** Experience : **36**
Author: **Louise;** Reward **-1;** Experience : **36**
Author: **Mathieu;** Reward **-1;** Experience : **36**
1

Property: doxygen comment style
Not observed
Author: **Louise;** Reward **1;** Experience : **50**
Author: **Louise;** Reward **-1;** Experience : **150**
Author: **Mathieu;** Reward **-1;** Experience : **200**
Author: **Édouard;** Reward **1;** Experience : **150**
Author: **Édouard;** Reward **-1;** Experience : **50**
2

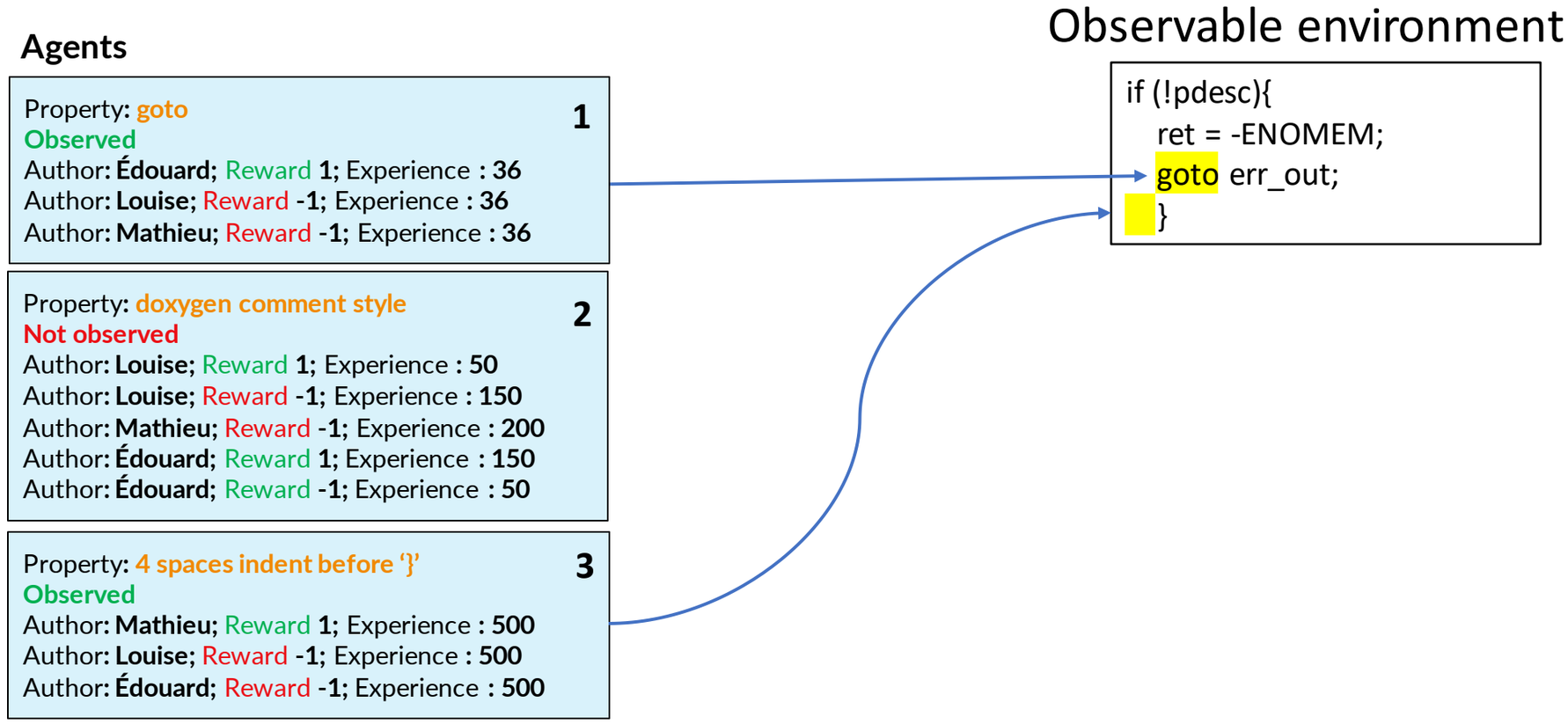
Property: 4 spaces indent before '}'
Observed
Author: **Mathieu;** Reward **1;** Experience : **500**
Author: **Louise;** Reward **-1;** Experience : **500**
Author: **Édouard;** Reward **-1;** Experience : **500**
3

Observable environment

```
if (!pdesc){
    ret = -ENOMEM;
    goto err_out;
    }
```

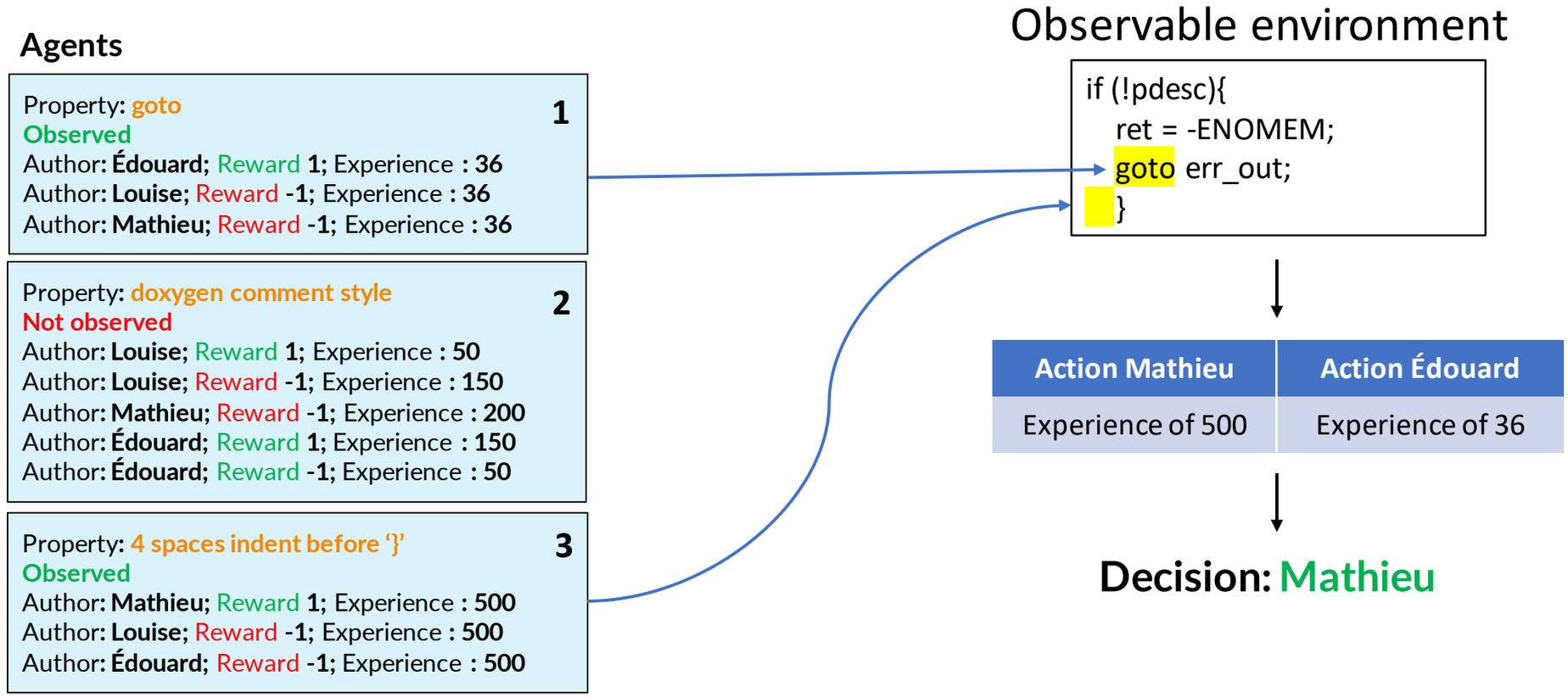
AI Focus

Explainability

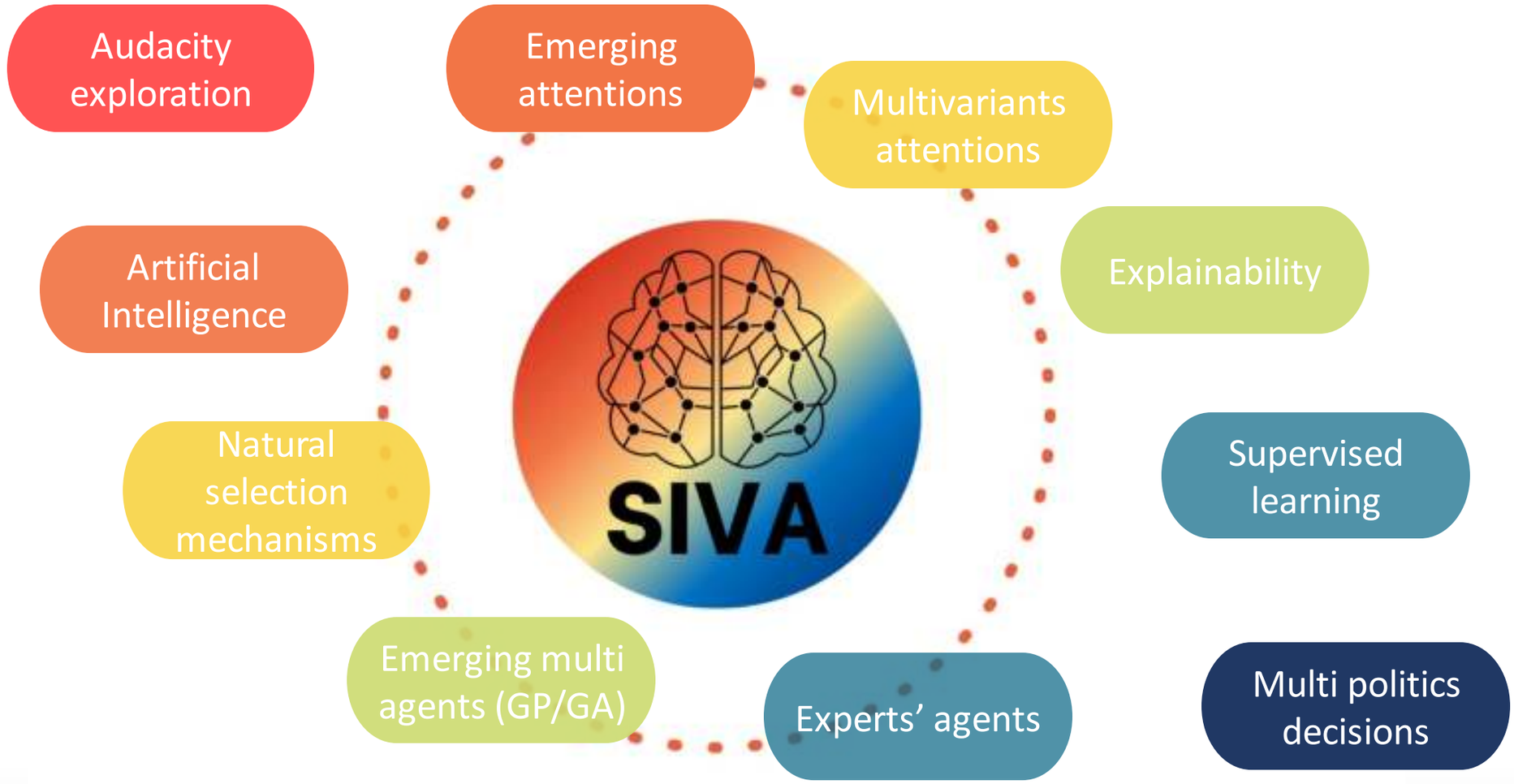


AI Focus

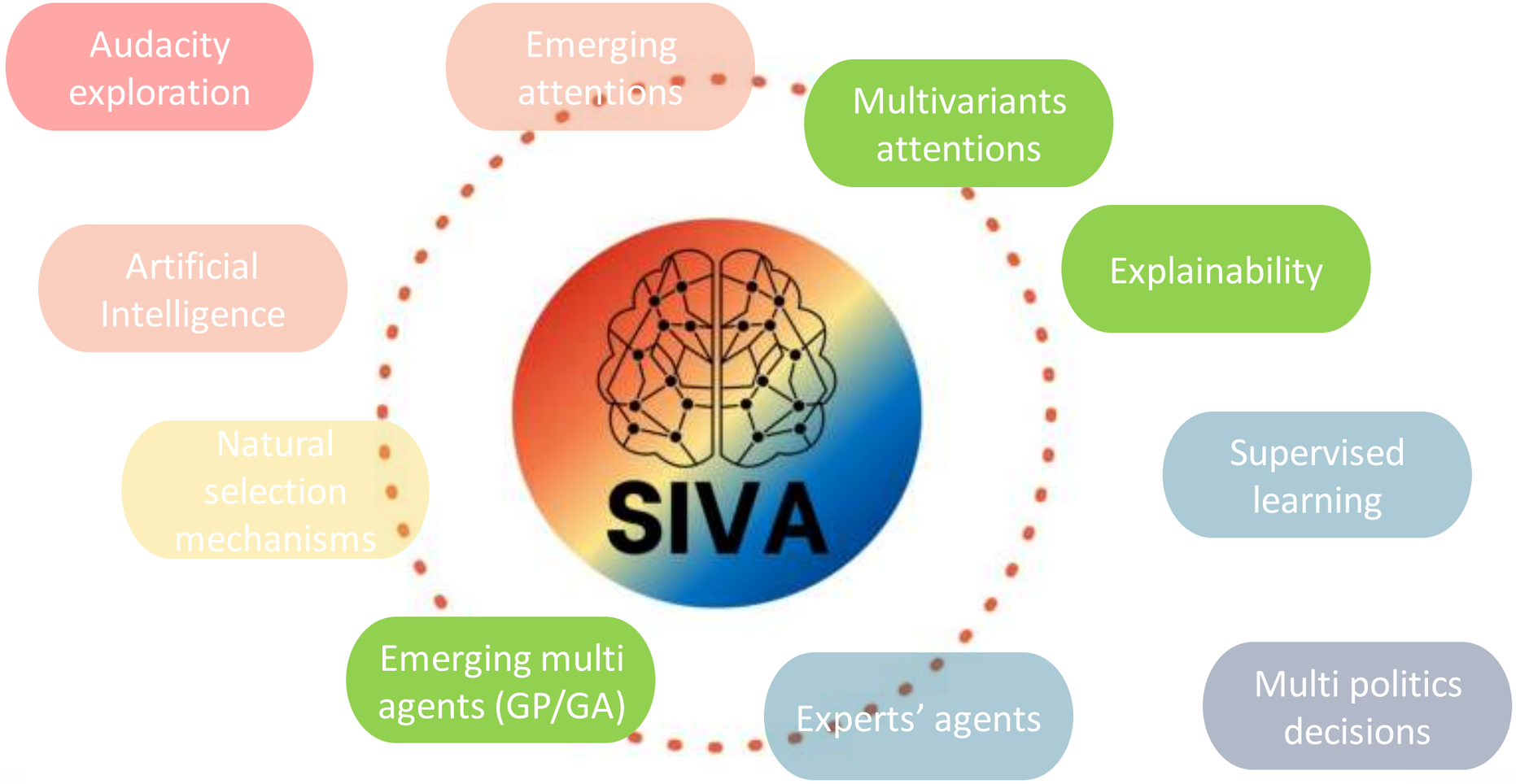
Explainability



Strengths


Audacity exploration
Emerging attentions
Multivariants attentions
Explainability
Supervised learning
Multi politics decisions
Experts' agents
Emerging multi agents (GP/GA)
Natural selection mechanisms
Artificial Intelligence
SIVA

Strengths


Audacity exploration
Emerging attentions
Multivariants attentions
Artificial Intelligence
Explainability
Natural selection mechanisms
SIVA
Supervised learning
Emerging multi agents (GP/GA)
Experts' agents
Multi politics decisions

Principle

C4PTION: The collaboration of 6 SIVA

6 different data types for 6 instances of SIVA:

- Log commit git
- Added code
- Raw code

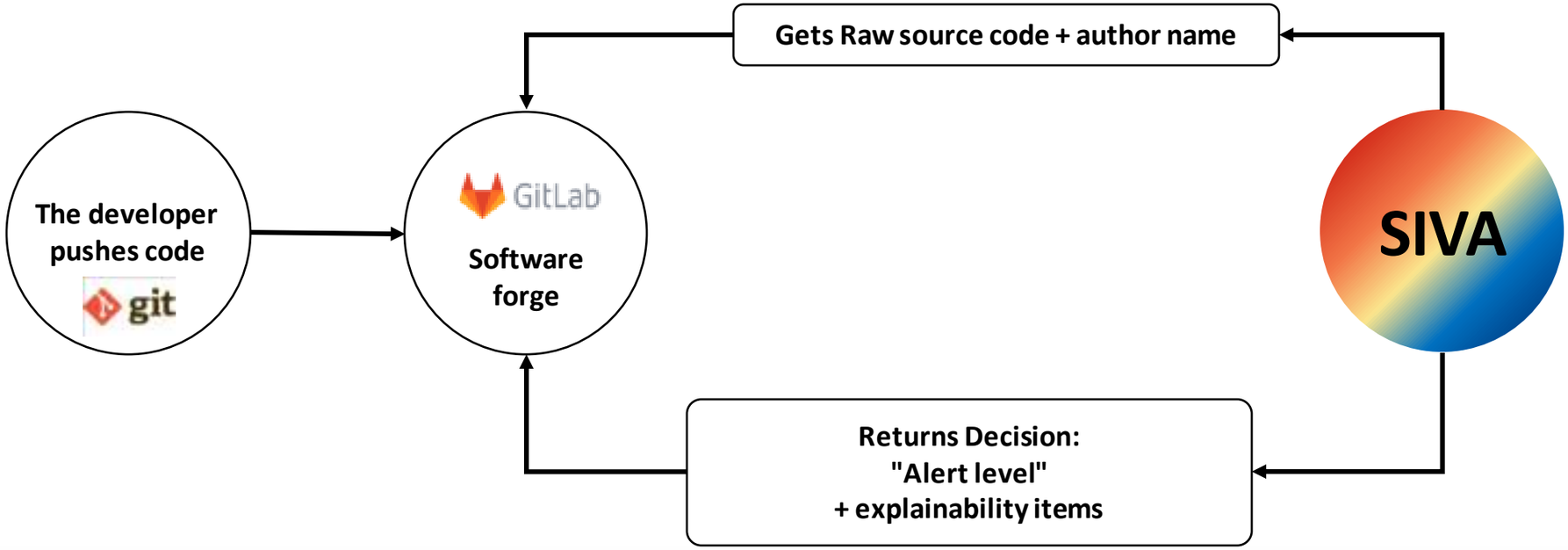
} 3 Used for **classification** of author

- Log commit git and author name
- Added code and author name
- Raw code and author name

} 3 Used for **abnormality** detection

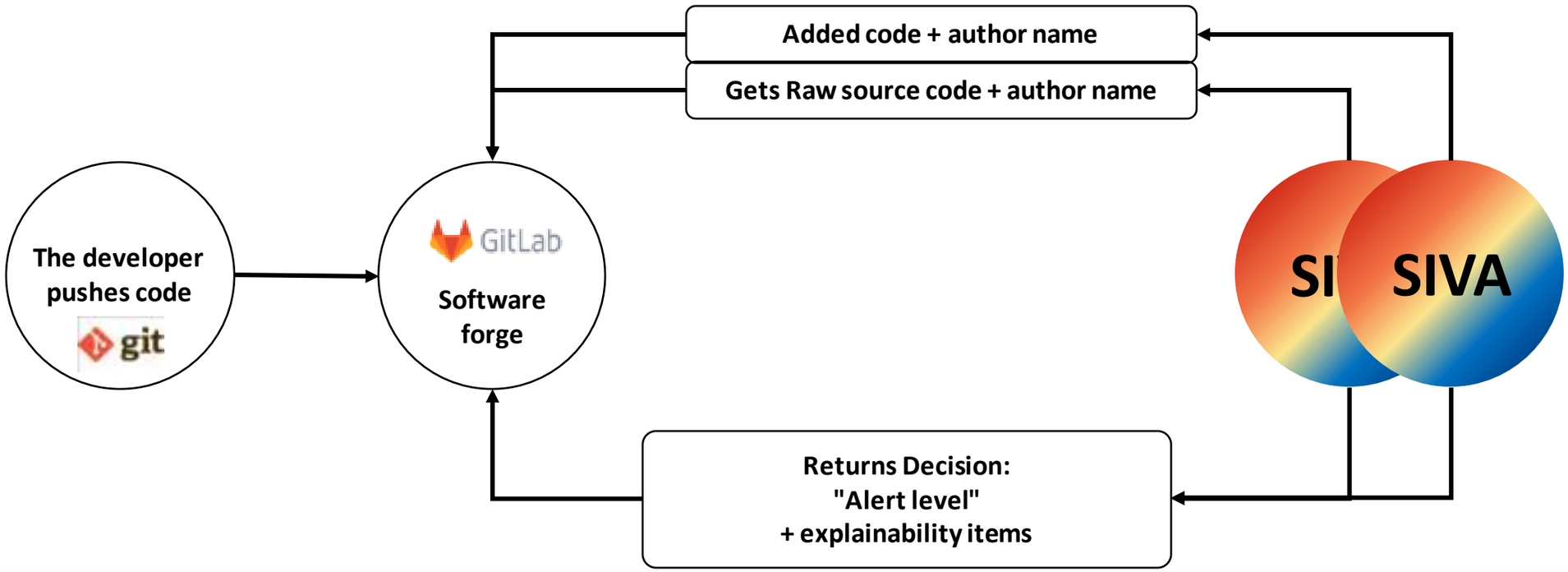
Principle

C4PTION: The collaboration of 6 SIVA



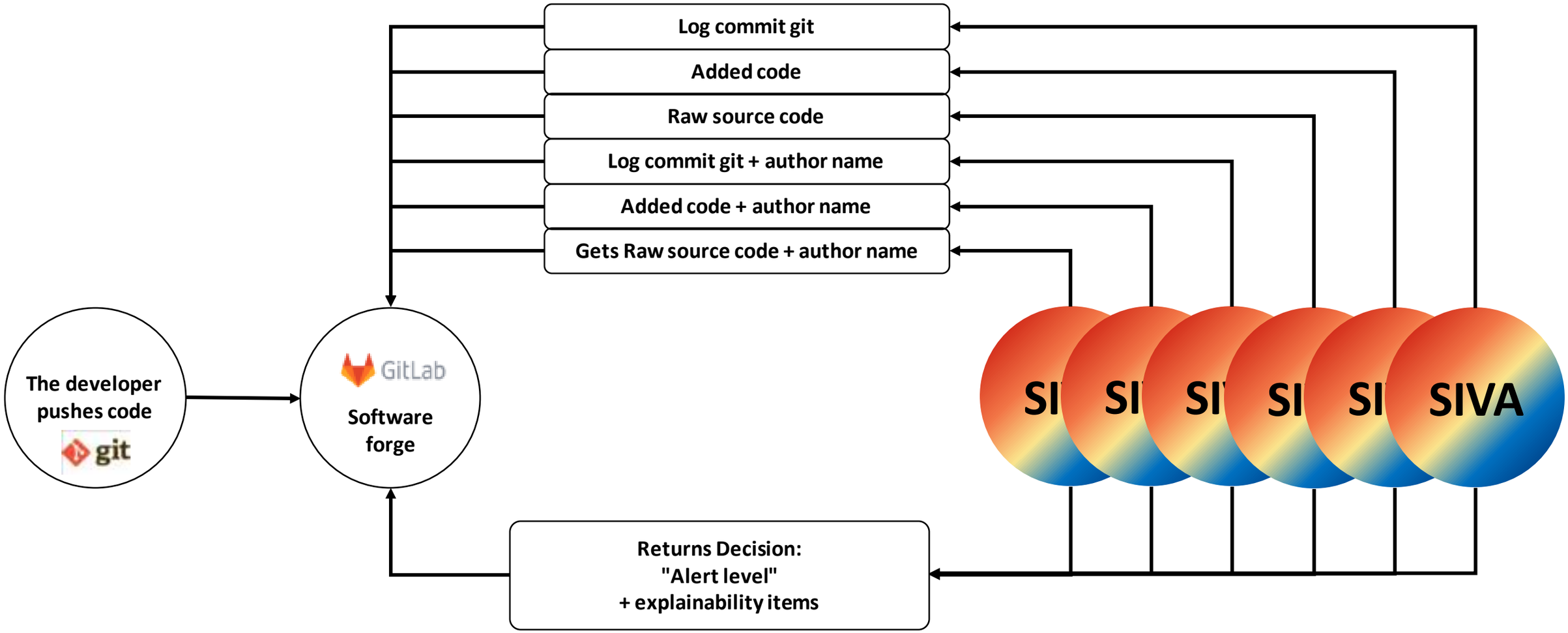
Principle

C4PTION: The collaboration of 6 SIVA



Principle

C4PTION: The collaboration of 6 SIVA



Results

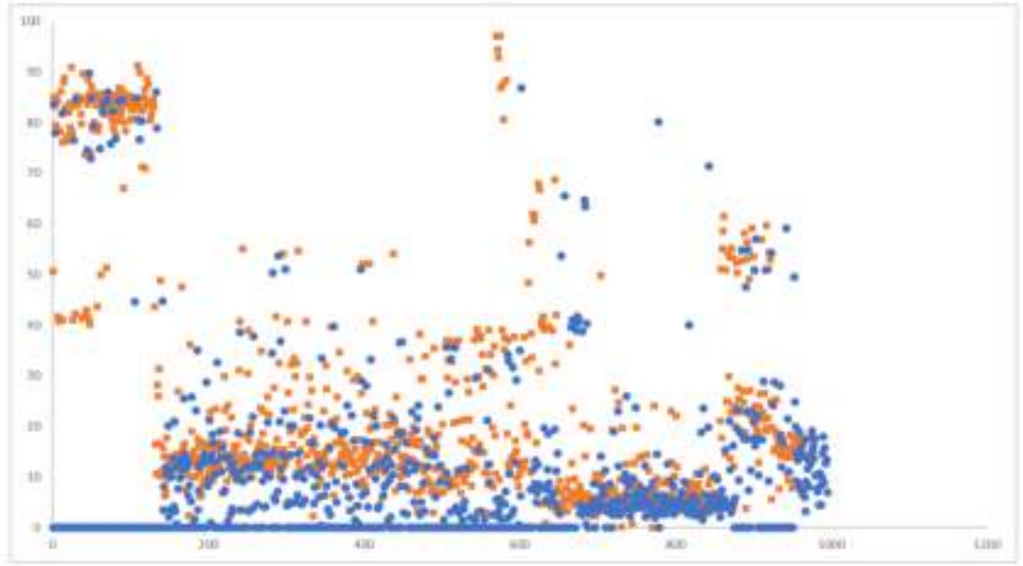
Tests configuration

- Trainings and validations on 13 repositories
- 16 CPU / 32Go RAM
- Ratio training/validation: 8 000 / 2 000 commits
- 50% of commits are malicious
- $Accuracy = \frac{TP + TN}{P + N}$

Github Url (prefix https://github.com/)	Number of authors
bfgroup/b2.git	179
yamashi/CyberEngineTweaks.git	28
DFIR-ORC/dfir-orc.git	6
ehids/ecapture.git	4
ros/kdl_parser.git	56
huggingface/neuralcoref.git	20
CPPAlliance/NuDB.git	11
preesm/preesm-apps.git	12
juliencombattelli/ProjectRomero.git	6
s-matyukevich/raspberry-pi-os.git	56
brchiu/raytracing.git	4
seL4/seL4.git	119
wkhtmltopdf/wkhtmltopdf.git	76

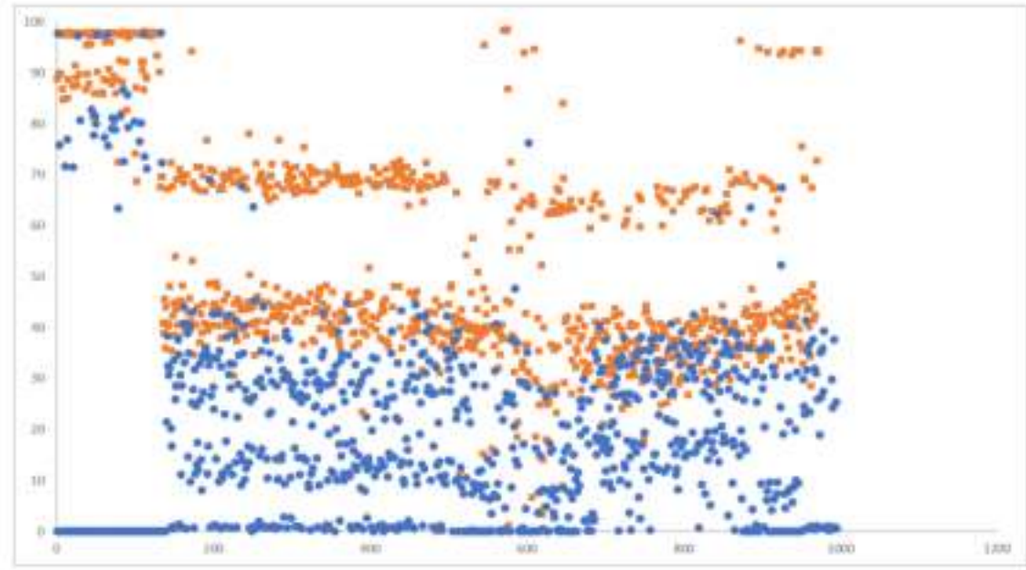
Results

Percentage of risk of alert for every commit



*Classification of raw source code (mono SIVA) with an accuracy of **72.08%***

- Benign commit
- Malicious commit



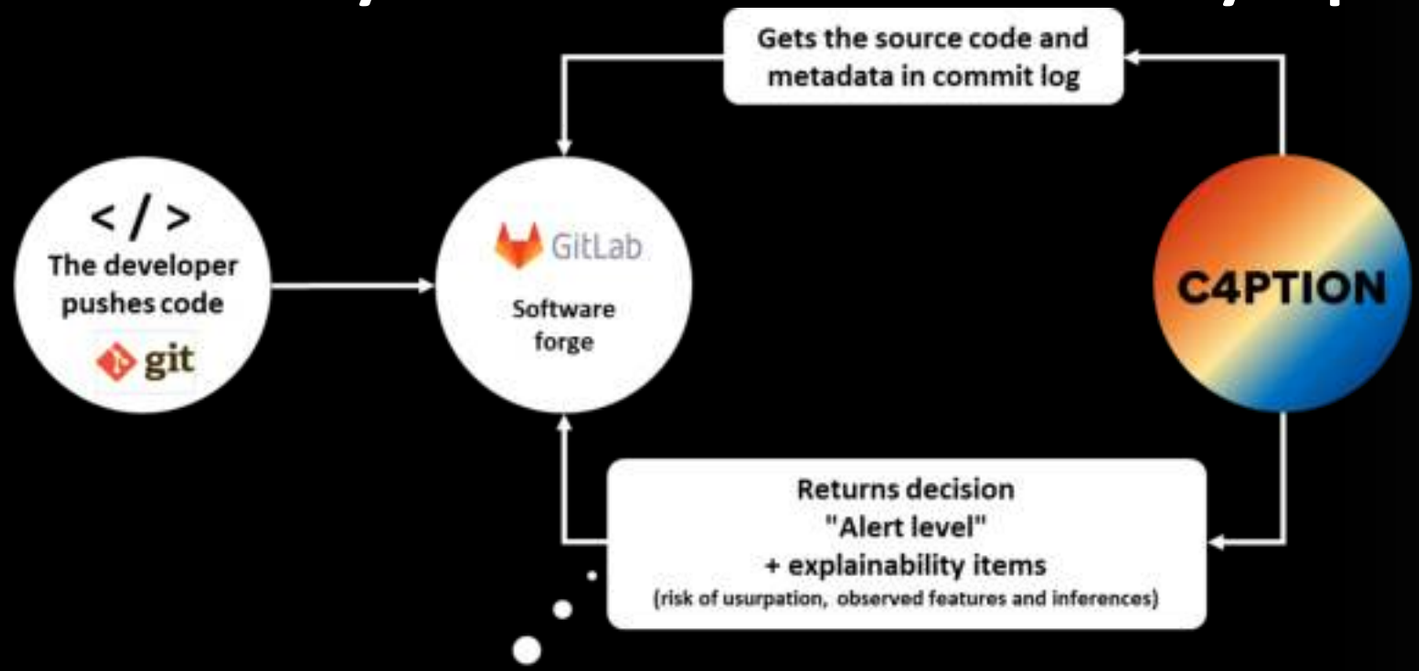
*Classification with 6 collaborative SIVA with an accuracy of **86.52%***

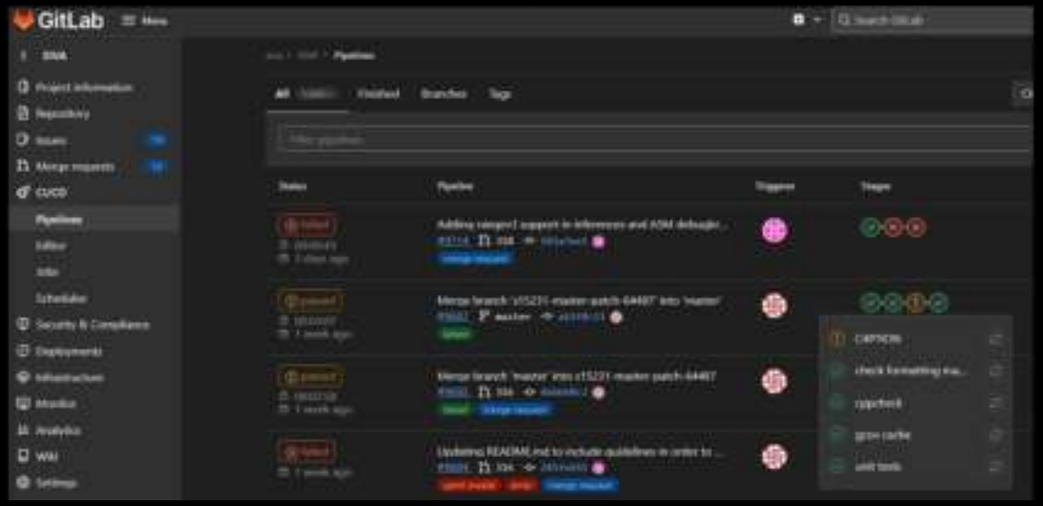
Conclusion

- C4PTION, the explainable artificial intelligence-based commits characterization solution to make supply chain more secure
- Performance of accuracy of 86.52%
- Contribution of the multi data multi agents system

- Diversification of agent types
- Improvement of user interface for explainability

Thanks for your attention. Any questions ?





Alerts are classified according to a configurable confidence threshold in 3 levels:

- High risk
- Medium risk (Questions identity of developer at a medium level)
- Neutral risk

Appendix

Intuitions on SIVA: systems 1 and 2

Inspiration of psychologist D. Kahneman, 2 systems of decision are used:

- A first one **rapid** and **instinctive** corresponding to the execution of the graph memorizing only the causal rules

- Another one **slower** and **thoughtful** corresponding to the execution of the graph paired with the base of expert agents

→ The base of experts is only used if the graph lacks of **confidence** within its decision