|  |  |
| --- | --- |
|  | Moving Picture, Audio and Data Coding by Artificial Intelligence  www.mpai.community |

|  |  |
| --- | --- |
| **Public document** | |
| **N155** | 2021/02/17 |
| **Source** | Valeria Lazzaroli, Guido Perboli, Mariangela Rosano |
| **Title** | Proposal for MPAI-CUI Use Cases and Functional Requirements |
| **Target** | MPAI Members |

# Introduction

Moving Picture, Audio and Data Coding by Artificial Intelligence (MPAI) is an [international association](http://mpai.community/) with the mission to develop *AI-enabled data coding standards*. Research has shown that data coding with AI-based technologies is *more efficient* than with existing technologies. Compression is a notable example of coding as is extraction of features.

The MPAI approach to developing AI data coding standards is based on the definition of *standard interfaces* of *AI Modules (AIM)* that operate on input and output data with standard formats. AIMs can be *combined* and *executed* in an MPAI-specified *AI-Framework* according to the emerging MPAI-AIF standard being developed based on the responses to [Call for MPAI-AIF Technologies](https://mpai.community/standards/mpai-aif/) [1] with associated [Use Cases and Functional Requirements](https://mpai.community/standards/mpai-aif/#Requirements) [2].

While AIMs must expose standard interfaces to be able to operate in an MPAI AI Framework, their performance may differ depending on the technologies used to implement them. MPAI believes that *competing* developers striving to provide more performing *proprietary* and *inter­operable* AIMs will promote *horizontal markets* of *AI solutions* that build on and further promote AI *innovation*.

The title of this document is *Compression and understanding of industrial data* (MPAI-CUI). It contains the “AI-based Performance Prediction” Use Case and associated Functional Requirem­ents. The MPAI-CUI standard uses AI substantially to reduce the amount of data and to extract the most relevant information from the industrial data, with the aim of assessing the performance of a company and predicting the risk of bankruptcy long before it may happen.

It should be noted that the AI-based Performance Prediction Use Case will be *non-normative*. The internals of the AIMs will also be *non-normative*. However, the input and output interfaces of the AIMs whose requirements have been derived to support the Use Cases will be *normative*.

The content of this document includes this Introduction and

|  |  |
| --- | --- |
| Chapter 2 | briefly introduces the AI Framework Reference Model and its six Components |
| Chapter 3 | briefly introduces the Use Case. |
| Chapter 4 | presents the MPAI-CUI Use Case with the following structure   1. Reference architecture 2. Description of AI Modules and their I/O data 3. Technologies and Functional Requirements 4. Interfaces of AIM I/O Data |
| Chapter 5 | gives a basic list of relevant terms and their definition |
| Chapter 6 | gives suggested references |

# The MPAI AI Framework (MPAI-AIF)

Most MPAI applications considered so far can be implemented as a set of AIMs – AI, ML and even traditional DP-based units with standard interfaces assembled in suitable topol­ogies to achieve the specific goal of an application and executed in an MPAI-defined AI Frame­work. MPAI is making all efforts to identify processing modules that are re-usable and upgradable without necessarily changing the inside logic. MPAI plans on completing the development of a 1st gener­ation MPAI-AIF AI Framework in July 2021.

The MPAI-AIF Architecture is given by *Figure 1*.



*Figure 1 – The MPAI-AIF Architecture*

MPAI-AIF is made up of 6 Components:

1. *Management and Control* manages and controls the AIMs, so that they execute in the correct order and at the time when they are needed.
2. *Execution* is the environment in which combinations of AIMs operate. It receives external inputs and produces the requested outputs, both of which are Use Case specific, activates the AIMs, exposes interfaces with Management and Control and interfaces with Communic­ation, Storage and Access.
3. *AI Modules* (AIM) are the basic processing elements receiving processing specific inputs and producing processing specific outputs.
4. *Communication* is the basic infrastructure used to connect possibly remote Components and AIMs. It can be implemented, e.g., by means of a service bus.
5. *Storage* encompasses traditional storage and is used to e.g., store the inputs and outputs of the individual AIMs, intermediary results data from the AIM states and data shared by AIMs.
6. *Access* represents the access to static or slowly changing data that are required by the application such as domain knowledge data, data models, etc.

# Use Cases

## AI-based Performance Prediction

A company may need to access the flow of internal (i.e., financial and governance data) and exter­nal data to assess and mon­itor its financial and organizational performance, as well as the impact of vertical risks (e.g., cyber, seismic, etc.), according to the current standards (e.g., ISO 31000 on risk assessment and management). In this phase only cyber and seismic risks that have an impact on financial performance are taken into account. Other risks will be considered in future version of the standard.

The company generating the data flow may need to perform compression and understanding for its own needs (e.g., to identify core and non-core data). Indeed, the company itself can analyse its financial performance, identifying possible clues to a crisis or risk of bankruptcy years in advance. It may help the board of directors and decision-makers to make the proper decisions to avoid these situations, conduct what-if analysis, and devise efficient strategies.

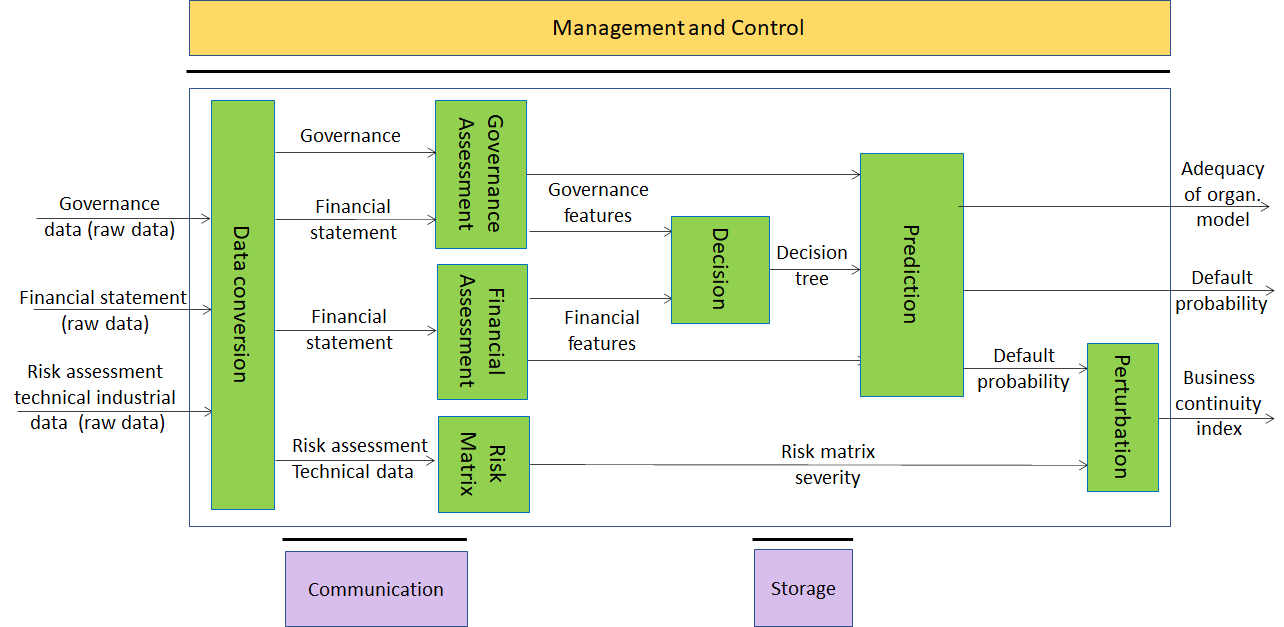
At the same time, a financial institution that receives a request for financial help from a troubled company, can access its financial and organizational data and make an AI-based assessment of that company, as well as a prediction of future performance. This helps the financial institution to make the right decision in funding or not a company, having a better insight of its situation.

# Functional Requirements

## AI-based Performance Prediction

### Reference architecture

This Use Case can be implemented as in *Figure 2*.



*Figure 2 – Compression and understanding of Industrial Data*

### AI Modules and their I/O data

The AI Modules of *Figure 2* perform the functions described in Table 1 – AI Modules .

*Table 1 – AI Modules of Industrial Data Compression and Understanding*

|  |  |
| --- | --- |
| **AIM** | **Function** |
| **Data Conversion** | Gathers data needed for the assessment from several sources (internal and external), in different formats and covert it in a unique format (e.g., json). |
| **Financial assessment** | Analyses the data generated by a company (i.e., financial statements) to assess the preliminary financial performances in the form of indexes.  Builds and extracts the financial features for the Decision tree and Pred­iction AIMs. |
| **Governance assessment** | Builds and extracts the features related to the adequacy of the governance asset for the Decision tree and Pred­iction AIMs. |
| **Risk matrix** | Builds the risk matrix to assess the impact of vertical risks (i.e., in this Use Case cyber and seismic). |
| **Decision** | Creates the decision trees for making decisions. |
| **Prediction** | Predicts values of the probability of company default in a time horizon of 36 months and of the adequacy of the organizational model. |
| **Perturbation** | Perturbs the probability value of company crisis computed by Prediction, considering the impact of vertical risks on company performan­ce. |

### I/O interfaces of AI Modules

The I/O data of Data Compression and Understanding AIMs are given in *Table 2 – I/O data of Use Case AIMs*.

*Table 2 – I/O data of Use Case AIMs*

|  |  |  |
| --- | --- | --- |
| **AI Module** | **Input** | **Output** |
| **Data Conversion** | Financial statement data  Governance data  Risk assessment data | Financial statement data (converted)  Governance data (converted) |
| **Financial assessment** | Financial statement data | Financial features |
| **Governance assessment** | Governance data | Governance features |
| **Risk matrix** | Technical data from internal assessment on risks (i.e., cyber security) | Severity |
| **Decision** | Financial features, Governance features | Ranking of features importance |
| **Prediction** | Financial features, Governance features | Probability of company crisis  Adequacy of organizational model |
| **Perturbation** | Probability of company crisis (index); severity from Risk Matrix | Index of business continuity |

### Technologies and Functional Requirements

#### Governance data (raw)

By Governance data we mean attributes that indicate the governance structure of a company and of the roles of key personnel.

The most basic roles are shareholder, manager, sole administrator, president/member of the board of directors, auditor, president/member of the statutory board of directors. They can be considered as “universal”, as commonly recognized across all countries. ISO 37000 (still under develop­ment) [6] aims at proposing a consistent set of recommendations, including definitions, for organizations in terms of governance. However, a governance data ontology is missing.

**To Respondents**

Respondents are invited to propose a governance data ontology that captures today’s practice at the global level.

#### Financial statement data (raw)

The Financial statement (raw data) are produced based on a set of accounting principles driving maintenance and reporting of company accounts, so that financial statements are consistent, trans­parent, and comparable across companies.

A set of principles [3], identified by the International Accounting Standard/International Financial Reporting Standard (IAS/IFRS), can be considered as “universal”, as they are commonly recognized across all countries. Indeed, although different countries can consider different accounting principles based on their jurisdictions, they are endorsed and standardised by the IAS/IFRS to guarantee their convergence [5].

An example of a tool that provides digital representation of Financial Statement data is the eXtensible Business Reporting Language (XBRL). The requirement of any such languages is that it should reflect the balance sheet structure in terms of assets, liabilities, and shareholders' or owners' equity.

The Financial statement (raw data) are converted to a standard format by the Data conversion AIM.

**To Respondents**

Respondent are invited to propose a digital representation of Financial statements data that is applicable to a minimum set of Financial statements having a universally valid semantics. JSON and XBRLS are primary examples of such digital representations. However, other representations are possible.

Respondents are invited to either select one of the two choices above or suggest alternative formats. In all cases justification of a proposal is requested.

Preference will be given to formats that have been standardised or are in wide use.

#### Risk assessment technical data (raw)

By Risk assessment technical data, we mean attributes that indicate the internal assessment that the company performs to identify and measure potential or existing vertical risks, and their impact on business continuity.

This data contains values of likelihood, impact, gravity, residual risk and treatments. All are and are described in ISO 31000 – “Risk management -- Principles and guidelines” [7].

**To Respondents**

Respondents are invited to propose a digital representation of Risk assessment technical data.

#### Governance

This is the Governance data (raw) after conversion. JSON appears to be a convenient format.

**To Respondents**

Respondents are requested to comment on this choice.

#### Financial statement

This is the Financial statement data (raw) after conversion. JSON appears to be a convenient format.

**To Respondents**

Respondents are requested to comment on this choice.

#### Risk assessment technical data

This is the Risk assessment technical data (raw) after conversion. JSON appears to be a convenient format.

**To Respondents**

Respondents are requested to comment on this choice.

#### Financial features

Financial features are a set of indexes and ratios computed using financial statement data. Examples of Financial features are given by

*Table 3*.

*Table 3 – Financial features*

|  |  |  |
| --- | --- | --- |
| **Feature** | **Feature value** | **Feature type** |
| **1** | Absolute value | Revenue/Profit |
| **2** | Index/Percentage (%) | Revenue/Profit |
| **3** | Absolute value | Revenue/Profit |
| **4** | Absolute value | Revenue/Profit |
| **5** | Index/Percentage (%) | Revenue/Profit |
| **6** | Index/Percentage (%) | Cost/Debt |
| **7** | Absolute value | Cost/Debt |
| **8** | Index/Percentage (%) | Cost/Debt |
| **9** | Absolute value | Cost/Debt |
| **10** | Index/Percentage (%) | Cost/Debt |
| **11** | Absolute value | Production |
| **12** | Absolute value | Production |
| **13** | Index/Percentage (%) | Revenue/Profit |
| **14** | Absolute value | Production |
| **15** | Index/Percentage (%) | Cost/Debt |

**To Respondents**

Respondents are requested to propose Financial features suitable for financial assessment. The Financial features should include those given by

*Table 3* and may include other features as well that satisfy the requirement of being extracted or computed from Financial statement data.

#### Governance features

Governance features are a set of indexes/parameters that are used to assess the adequacy of the organizational model. *Table 4* gives examples of Governance feature.

*Table 4 – Governance features*

|  |  |  |
| --- | --- | --- |
| **Feature** | **Feature value** | **Feature type** |
| **1** | Absolute value | Decision maker data |
| **2** | Index/Percentage (%) | Shareholder data |
| **3** | Absolute value | Shareholder data |
| **4** | Absolute value | Decision maker data |
| **5** | Absolute value | Decision maker data |

**To Respondents**

Respondents are requested to propose Governance features suitable for assessing the suitability of governance, e.g., those reported in *Table 4*. Proposed Governance features shall satisfy the requir­ements of:

1. Being extracted or computed from the Governance data.
2. Being expressed by numerical values.
3. Adding insight to the data of *Table 4*.

#### Severity

A set of values, each reflecting the level of risk for a specific vertical risk, cyber and seismic in the phase, as evaluated by the company. This severity is computed according to ISO 27005 [8], considering the levels of probability of occurrence, business impact and gravity of a specific risk.

**To Respondents**

Respondents are invited to comment on this choice or to propose and motivate alternative solutions.

#### Decision Tree

It is a tree-like decision model, built starting from the financial and governance features. An example is provided by [9] where the Random Forest supervised learning method has been used to predict the probability of company crisis and bankruptcy.

**To Respondents**

Respondents are invited to propose a decision support tool.

#### Default probability

It is a score in the 0 to 1 range that represents the likelihood of company default in a specified number of future months dependent on financial data. It is computed by Prediction using the financial features and decision tree.

**To respondents**

Respondents are requested to comment on the description above and to propose extensions.

#### Adequacy of organisational model

It is a score in the 0 to 1 range that represents the adequacy of the organisational model. It allows to identify potential critical points or conflicts of interest that can lead to an increase in the risk of default. It is computed by Prediction using the governance and financial features.

**To respondents**

Respondents are requested to comment on the description above. Suggestions about multidimensional measures of adequacy are welcome.

#### Business continuity index

It is a score in the 0 to 1 range that represents the likelihood of company default in a specified number of future months dependent on financial non-financial data. It is computed by Perturbation using default probability and severity.

**To Respondents**

Respondents are requested to comment on the description above and to propose extensions.

# Terminology

*Table 5* identifies and defines the terms used in the MPAI-CUI context.

*Table 5 – MPAI-CUI terms*

|  |  |
| --- | --- |
| **Term** | **Definition** |
| Access | Static or slowly changing data that are required by an application such as domain knowledge data, data models, etc. |
| AI Framework (AIF) | The environment where AIM-based workflows are executed |
| AI Module (AIM) | The basic processing elements receiving processing specific inputs and producing processing specific outputs |
| Communication | The infrastructure that connects the Components of an AIF |
| Data Processing (DP) | A legacy technology that may be used to implement AIMs |
| Decision Tree | A decision support tool that uses a tree-like model of decision, given the financial and governance features |
| Delivery | An AIM that wraps data for transport |
| Execution | The environment in which AIM workflows are executed. It receives external inputs and produces the requested outputs both of which are application specific |
| Financial features | A set of indexes and ratios computed using financial statement data. |
| Financial statement | Data produced based on a set of accounting principles driving maintenance and reporting of company accounts, so that financial statements can be consistent, transparent, and comparable across companies. |
| Governance features | A set of indexes/parameters that are used to assess the adequacy of the organizational model. |
| Knowledge Base | Structured and unstructured information made accessible to AIM (especially DP-based) |
| Management and Control | Manages and controls the AIMs in the AIF, so that they execute in the correct order and at the time when they are needed |
| Risk assessment | Attributes that indicate the internal assessment that the company performs to identify and measure potential or existing vertical risks, and their impact on business continuity |
| Severity | A set of values, each reflecting the level of risk for that specific vertical risk as evaluated by the company |
| Storage | Storage used to e.g., store the inputs and outputs of the individual AIMs, data from the AIM’s state and intermediary results, shared data among AIMs |

# References

1. MPAI-AIF Call for Technologies, N100; <https://mpai.community/standards/mpai-aif/#Technologies>
2. MPAI-AIF Use Cases and Functional Requirements, N74; <https://mpai.community/standards/mpai-aif/#Requirements>
3. MPAI-AIF Framework Licence, MPAI N171; <https://mpai.community/standards/mpai-aif/#Licence>
4. International Financial Reporting Standard. List of IFRS Standards. Available online: <https://www.ifrs.org/issued-standards/list-of-standards/>
5. European Commission. Financial reporting. Available online: <https://ec.europa.eu/info/business-economy-euro/company-reporting-and-auditing/company-reporting/financial-reporting_en>
6. International Organization for Standardization. ISO 37000 Guidance for the Governance of Organizations. Available online: <https://committee.iso.org/sites/tc309/home/projects/ongoing/ongoing-1.html>
7. International Organization for Standardization. ISO 31000 Risk Management. Available online: <https://www.iso.org/files/live/sites/isoorg/files/store/en/PUB100426.pdf>
8. International Organization for Standardization. ISO 27005 Information technology -- Security techniques -- Information security risk management
9. Perboli G., Arabnezhad E., A Machine Learning-based DSS for Mid and Long-Term Company Crisis Prediction. To be published in *Expert Systems with Applications*. 2021.