AI BASED TRANSFORMATION PROJECTS-ML AND NLP MANAGED AUDIT, CONTROL, AND GOVERNANCE (ML_ACG)-THE BASICS

Antoine Trad

IBISTM, Transformation Projects, Paris La-Defense France, France

ABSTRACT

This article focuses on Artificial Intelligence (AI) based Transformation Projects, in which Machine Learning (ML) and Natural Language (NL) Processing (NLP) are used to interface, integrate, and manage enterprise's Audit, Control, and Governance (ML_ACG) Framework (ML_ACGF). It also promotes an In-House Implementation (IHI) approach for such a holistic framework. Knowing that today the fields related to audit-analytics, and the use of AI and NLP, have become a strategic factor for transformation projects. Such projects even revolutionize the manner in which auditors extract and use insights that are sourced from huge data-volumes and then used by specialized frameworks, like the ML_ACGF. The ML_ACGF, uses various AI Subdomains or fields (AIS), in order to: 1) Enable accurate analysis, interpretation, and generation of human-language based reports; and also to improve risks' assessment processes; 2) Discover hidden insights in documents, data-sets, specialized audit-medias, design-documents, source-code, emails, and other contents' formats; 3) Automate audit procedures by interfacing NLP, Machine Learning (ML), and Natural Language Generation (NLG) modules, where ML learns from historic data-sets to find hidden patterns and to enable the prediction of possible important projects' risks, knowing that predicting such risks is a very critical subject, because transformation projects have an eXtremely High Failure Rates (XHFR), which is about 95%; 4) Project's architecture and related models, are used to instantiate solutions by applying standard Enterprise Architecture (EA) and audit methodologies; 5) Use decision-making processes to mitigate critical risks; and 6) Present interfaces' and integrations' technics for existing market' auditing frameworks like the Control Objectives for Information and Related Technologies (COBIT), Committee of Sponsoring Organization's (COSO) Enterprise Risk Management (ERM), and other. For this article the author: 1) Uses his research framework that is based on a mixed-method and a Polymathical approach, which is specialized in supporting transformation projects; 2) Implements and presents a Conceptual Proof of Concepts (CPoC), which illustrates a concrete case and solution; and 3) Presents a list of conclusions and recommendations on ML_ACG's feasibility and advantages, where the main goal is to provide models and concepts that support and implement ML_ACGF's different features, and not just to use quantitative analysis to show commercial opportunities.

KEYWORDS

Machine Learning, Artificial Intelligence, Natural Language Processing, Auditing and control procedures, Transformation Projects, Enterprise Architecture, Automation, and Decision-making.

1. INTRODUCTION

This article (and all author's research works) focus(es) on AISs' usage in audit activities, which can include a wide range of heterogenous AI domains and other associated fields. Where AISs are used and integrated, to execute automated tasks in various APplication Domains (APD), and tries to mimic the Human Factors (HF) Brain (HBB). And that needs a Polymathic approach to AIS (PAIS), which can be considered as avant-garde and cross-functional foundation of EA based Composite Models (CModels), which include functional, mathematical, and scientific fields. A Business, Organizational, Governance, or Common Transformation Project (simply Project) uses CModels to integrate automated Audit, Control, and Governance (ACG) Procedures (ACGP). ACGPs are called using NLP Environments (NLPE) or other environments, to support various types of implementations like routine audit processes, which can use also ML modules. Using subdomains like ML, needs fine-tuning of common pre-trained AI "foundation models". These models are trained on huge amounts of In-memory DataSets (IDS) (or data columns) to be used by AI-based interactions. As shown in Figure 1, PAISs are a new form of AI-based interactions that are needed to combine various AISs like ML, NLP, IDSs, and many other topics [1]. The PAIS adopts a "generalized intelligence" approach, which contrasts with the common commercial products' approach, which promotes a specialized AIS-approach that is used in specific and well-defined contexts. PAISs contain a large set of capabilities, like NLPEs' interfacing, Image recognition, Intelligence based decision-making, and Logical reasoning, which are effective in various APDs. PAISs are closely related to the evolution of Artificial Generic Intelligence (AGI), which focuses on developing Information and Communication Systems (ICS) that can: 1) Apply Learning Processes (LP) for complex audits; and 2) Understand and apply gained experiences (and knowledge) as broadly as an HF Intelligence's (HFI) capacity. But, PAISs raise major critical considerations for AISs based Project's integration and inter-disciplinary learning processes, like the fact of challenging how ICSs can mimic complex HFI's intellectual-traits. PAISs have the ultimate goal of offering versatile, and general-purpose decision-making; especially for more generalized initiatives like ML_ACG's activities.

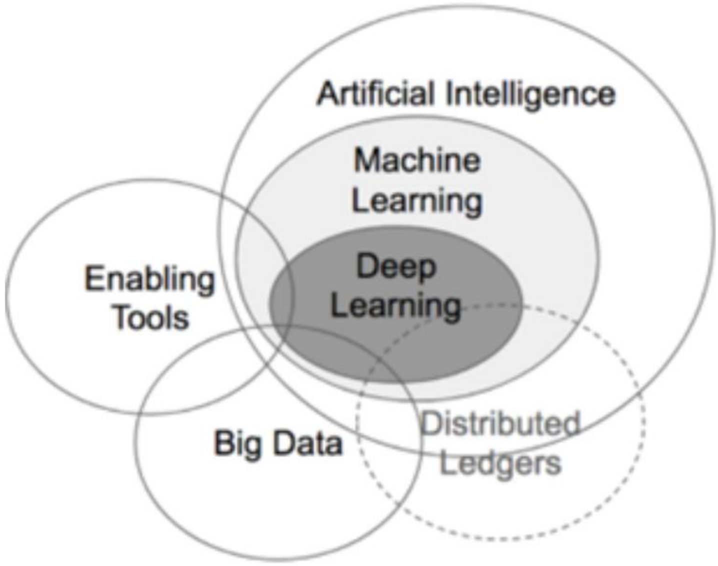


Figure 1. The PAIS' approach and AIS' interactions.

Knowing that actually, environments or more precisely commercial AIS products, like OpenAI and DeepMind, are contributing heavily in the development of Polymathic capabilities in AISs [2]. The ML_ACG is analysed in the context of PAISs based Projects, where:

- It is used to manage ML_ACGF's interfaces, integration processes, and enterprise's ACGPs related operations.
- Adopts an IHI concept for the ML_ACGF and PAISs integrations.
- AISs based audit-analytics process is strategic for maintaining an ICS, where it focuses on automating ACGPs (or internal audit controls and procedures) by using NL, NLP, ML, DL, and other.

- The ML_ACGF contains an LP Container (LPC) that persists experiences, solutions, and historic-data-volumes; which is used to predict Projects' risks and avoid XHFRs.
- Proposes an optimal Project's architecture and CModels, to support possible solutions by using EA.
- The use of decision-making process to evaluate possible risks and to launch corrective ACGPs' Actions (ACGA).
- Proposes interfaces for interacting with existing major audit frameworks like COBIT, COSO's ERM…
- Uses a specific mixed-method and a Polymathical approach for complex (re)engineering undertakings, like the Project.

ACGAs are the needed actions that are defined by enterprise's ACG policies and procedures, which support the implementation of Project's management visions and directives to mitigate possible risks until the achievement of Project's objectives. ACGAs are applied at all Project's levels, and in various Business Processes (BP) or WorkFlows (WF) parts, and in other ICS' artefacts. These ACGAs are preventive and include a set of (semi)manual and automated tasks like: Authorizations (and approvals); Verifications; Reconciliations, and Performances' reviews. Segregation of duties is typically built into the selection and development of control activities [3]. This research article has been adapted for transformational activities and can appear as complex, therefore, it is recommended that the valuable reader refers to the author's In-House Implemented (IHI) Polymathic Transformation Framework (IHIPTF) related articles, works, guides, and resources, like: The IHIPTF Guide [4]; The IHIPTF Glossary [5]; A related syllabus [6]; The AHMM4PROJECT and business transformations' implementation [7,8]; The AI based Projects [9]; and this article's specific used abbreviations and terms, which are found at the end of this article.

2. RDP FOR ML_ACG

The following section describes the research and development environment which was implemented by the author.

2.1. Basic Constructs

The RDP is made up of the following artefacts, facts, modules, and constraints [3,13,14]:

- The Research Question (RQ) is: "How can the ML_ACG support Projects audit, control, and governance activities?". And this article's auxiliary question is: "How can AISs support ML_ACG's automation?".
- Applies a Polymathic research mixed-method approach, where it is mainly a qualitative heuristics approach.
- Uses a collection or set(s) of initially selected, initialized, and related Critical Succes Factors (CSF), Critical Success Areas (CSA), Key Performance-Indicators (KPI), and ICS' concrete NLP (or ICS' programming-language) VARiables (VAR).
- A CSA corresponds and links to important ML_ACG's (or major) topics (that are organized in sections); and this article 1st CSA (or section) is the RDP.
- CSFs and CSAs are used to evaluate ML_ACG's integration statuses and Project's feasibility; which are supported by the Applied Holistic Mathematical Model (AHMM) for AI (AHMM4AI) and IHIPTF's modules.
- The evaluation processes are supported by the Decision-Making System (DMS) and Knowledge-Management System (KMS) (simply Intelligence).

- CSFs, CSAs, KPIs, and VARs are labelled as Factors; and these Factors are processed by DMS' IHI Heuristics-Decision-Tree (IHDT).
- Uses sets of Factors and transformed/generated services, to solve Project's and ML_ACG's problem-types.
- It is assumed that the Project's generated services are organized in Blocks to be used by common, ICS', and AIS' calls.
- It is assumed that the Entity's Unbundling (and transformation) Processes (EUP) of legacy common-functional, ICS's modules, templates, and resources, was successful and the needed services were successfully generated.
- Uses different modelling disciplines like standard Enterprise Architecture (EA) Modelling or Models (EAM), Composite AI Models (CAIM), ICS Models (ICSM), and Intelligence interfaces.
- Applies CAIMs which in turn contain EAMs, and other types of models (simply CModels).

2.2. Advanced Constructs

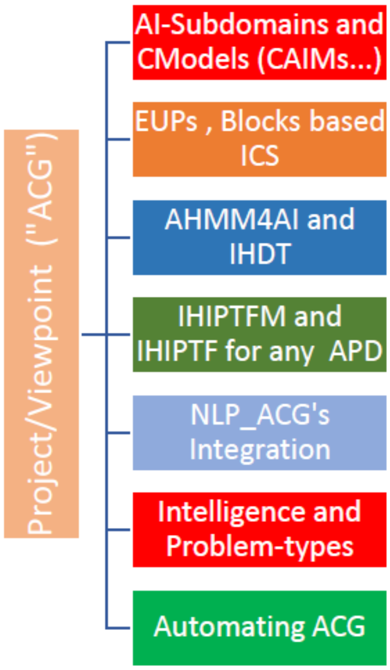


Figure 2. RDP's and Project's Viewpoints.

As shown in Figure 2, the RDP is made up of the following advanced phases and actions [10,11]:

- Executes the Polymathic Literature Review Process (PLRP) which is an in-depth analysis. It proved that there isn't a similar approach or framework that uses the already mentioned keywords or parts like IHIPTF, EUP, AHMM4AI, CModels... And there isn't any relevant scholar-references or resources that analyzes the mentioned RQ and topics.
- Projects face XHFRs and that was proven in previous PLRPs.

- RDP's RQ main gap and weakness is that there is an extreme lack of Polymathic concepts and skills, like in critical fields such as AISs, where it is recommended to use a PAIS approach.
- The PLRP used various resources like: 1) Articles, books, professional experiences, and resources related to the RQ, AISs, CAIMS, EAMs, ICSMs, and EUPs; 2) Author's published articles, and the IHIPTF; 3) ML_ACG's conceptual models and feasibility-checks; 4) Various ACG topics; and 5) The usage of the Polymathic Empirical Engineering Research Model (PEERM).
- As an important research-gap was acknowledged, therefore there is the need to continue with ML_ACG's research and analysis.
- The RDP has two phases, which are: 1) Phase "1", which uses CSA Decision-Tables (CSA_DT), and is RDP's empirical and heuristics evaluation phase; and 2) Problem-types' solving phase.

2.3. The Framework-IHIPTF

The Project's main goal is to offer ML_ACG recommendations, patterns-concepts, and a real-world ML_ACG Roadmap (ML_ACGR). The ML_ACGR presents a timeline for ML_ACG's main components integration, like CModels that can be presented with generated cartographies. These generated cartographies are produced and updated by Generative AI (GenAI) technics and are synchronized by the Architecture Development Method (ADM). The ML_ACG uses existing ADM that is included in The Open Group's Architecture Framework (TOGAF); which is specialized in transformations' (re)architecture reengineering. As shown in Figure 3, the ADM based IHIPTF has the following Sections (IHIPTFS) (2025a): 1) Is applied to support Projects and ML_ACG; 2) It offers the IHIPTF Platform (IHIPTFP); 3) It offers cases on Projects in the IHIPTF Cases (IHIPTFC); 4) It includes the refined Blocks (and patterns) (IHIPTFB); 5) It includes an extensive IHIPTF Dictionary (and educational Syllabus) (IHIPTFD); and 6) It includes an implementation IHIPTF Environment (IHIPTFE).

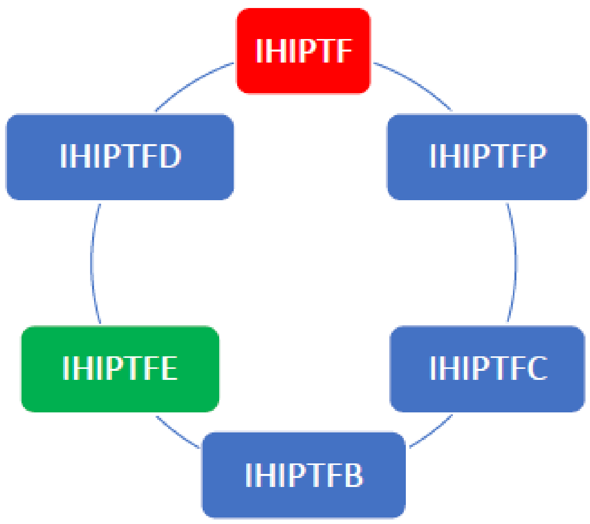


Figure 3. The IHIPTFSs.

The Project and ML_ACG, need all IHIPTFSs, that in-turn offers the following capabilities [34]:

- The IHIPTFB delivers sets of Blocks (and various categories of patterns), and associations that can be used by existing standard EA methodologies, resources, artefacts, and models; like TOGAF and its ADM.
- IHIPTFCs describe Project’s and ML_ACG’s optimal Applied Case-Studies (ACS) that are used in the CPoC.
- The IHIPTFD is a common Project-dictionary; and in this article (it is a part of this article, and located at its end), contains the specific list of abbreviations and terms.
- The IHIPTFE offers a concrete sample on how to implement a transformation environment (and the IHIPTFE was implemented using Microsoft .NET Visual Studio).
- The IHIPTFM manages the Project and ML_ACG’s abstractions to various ICS-technologies and external methodologies.
- The IHIPTFP presents the Project’s and ML_ACG’s heterogeneous ICS-platform(s) and Entity’s infrastructure.

The IHIPTFFM supports CModels which are used by the Project and standard methodologies, like TOGAF, Unified Modelling Language (UML), Domain-Driven Design (DDD), Entity Relational Modelling (ERM)… For the ML_ACG all the mentioned sections need to be modelled and implemented; where the IHIPTFM and IHIPTFB are the most critical sections.

2.4. The PEERM and RDP’s Phases

This RDP includes the following phases: 1) Phase one or “1” that uses CSA_DTs, and is RDP’s empirical phase; in which the IHIPTF evaluates the selected CSAs and the results are synthesized in Figure 20. The selected CSAs are: a) RDP; b) ICS’, EA, and technologies integration; c) The AISs and NLP’s integrations; d) Existing ML_ACGFs; e) ML_ACGF’s implementation; and f) ACSs’ usage; and 2) Phase two or “2”, presents how to solve a concrete ACG problem-type(s), by using Intelligence and its embedded IHDT. This RDP applies the PEERM, which is optimal for Projects, AISs, and NLPEs. The CPoC’s final outcome, is synthesized in Figure 20; and the IHIPTF supports ML_ACG’s implementation, an ML_ACGR, and offers a set of solutions, and recommendations [34].

This RDP applies a specific mixed-method that is different from conventional (mainly quantitative-statistics based) approaches, and is mainly a qualitative research-model. The mentioned mixed-method includes [15,16]: 1) IHDT’s adapted heuristics-basic-reasoning module; 2) Quantitative-methods for Projects and risks’ evaluations (Qn); 3) Qualitative-methods for Projects (Ql) and risks’ evaluations, to support empirical concepts for complex research initiatives; and 4) An IHDT based LPs (ILP), that is basically built on the Action Research (AR) approach. The AHMM4AI supports and structures the ML_ACGF and also supports its activities like interfacing the IHDT and Intelligence. Such complex undertakings need a well-built ICS, which uses established EA methodologies, and performant technologies.

3. ICS’, EA, AND TECHNOLOGIES INTEGRATION

This section describes the used ICS, EA approach, avant-garde technologies, and common resources that are related and are important for this article and RQ. The description of their basics is intended to help the valuable reader to follow this complex article.

3.1. The Role of Automation

This section describes the needed automation that is based on ICS, EAMs, CModels, and, other methodologies and technology components, which are designed and integrated to automate various types of ACGPs (or controls), and mitigate risks. That can include various steps in the ADM based lifecycle, like: Gathering evidence; Validating evidence; Analysing evidence; Assessing; and Concluding effectiveness and documentation or certification [17]. Such a lifecycle happens after that a control is designed as part of an Agentic AI (AIA) WorkFlows (AIAWF). ML_ACG's automation helps in: 1) Efficiencies' improvements, by shortening labour-time, producing and analysing evidence(s), and promoting effectiveness; 2) Risks' reduction by increasing the coverage of ICS' and technologies' assets; and 3) Effectiveness by implementing continuous tracing and monitoring, in which ACGPs (or controls) are automatically validated and tested for compliance. To ensure successful ACGPs' automation in risks' management process(esses), by using AIAWFs, there is the need for [17]:

- Understanding AIAWFs and their related controls' concepts, and ICS' availability.
- Access management, where users' access operations management are complex and need manual interventions; although ICS' have centralized granting schemes.
- Change management that includes multiple AIAWFs running on a centralized ICS, which is optimal for automation processes.
- Cybersecurity schemes' operations include: Alerts' generation; Data-losses' prevention monitoring and applied Cybersecurity WFs have immense volumes; Centralized ICS; and Automation of WFs is a strategic priority.
- AIAWFs are used to synchronize ACGPs (or controls), and Data-Analytics (DA) tools to provide possibilities to reduce execution cycles. Automation processes cannot solve maturity or deficiencies in ML_ACG, but it can increase efficiencies in the ACGPs' execution.

3.2. Integrating ACGPs

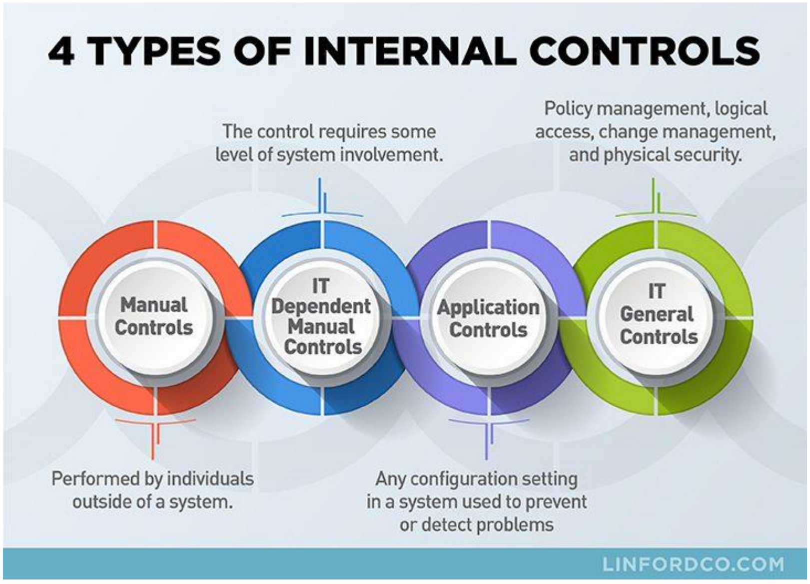


Figure 4. Types of IACGPs (or controls) [18]

The ML_ACG creates touchpoints in ICS' or AISs' processes that are located, reviewed; and that enables Project's accountability and to avoid risks like: Fraud; Waste; Abuse; and Errors. According to Committee of Sponsoring Organizations (COSO), an Internal ACGPs (IACGP or internal control) is a sub-process, which is managed by a specialist (like Auditor, EA specialist, Security specialist, Director, Manager, or Project's team-member), and is designed and implemented to provide a sufficient level of assurance of achieving objectives related to ML_ACGs' operations, reporting, and compliance. The ultimate objective is to track and evaluate: Progresses; Sustainability; and Performances. IACGPs include: Manual interventions; ICS related guides (or manuals); and ICS general, and application specific procedures. IACGPs are used to evaluate and recommend if the Project (or its strategic objectives) are feasible and realistic in the defined enterprise's contexts, expectations, operations, law/policies…[18]. As shown in Figure 4, IACGP's (or internal controls) types are [18]:

- Manual, which are performed by persons (like team-members) who are ICS' delimiters, like Auditors' review processes, Sign-off of reports, Financial reconciliation… And if IACGP(s) fails, it is documented in the audit-report.
- ICS's Dependent Manual Controls (DMC) are similar to manual IACGPs, because they include manual processes, but in which the ICS does some automation, like generating the list of users that have accessed the ICS.
- Application of IACGPs that prevent problems-ypes.
- ICS' general IACGPs, includes System and Organization Controls (SOC), that are used in audits processes and are incorporated in the Entity's: Policies management; Logical accesses; Change management; and Physical security. Like in the case of enabling users' access-rights.

There are also preventative and detective IACGPs that are efficient and can reduce costs, when problems happen. The ML_ACG enables the usage of metrics that are linked to various selected Factors, like Efficiency; Statuses' checking; and Effectiveness of a processes. IACGPs' weaknesses are mainly due to: Building CModels; Operating effectiveness… And that all can be enhanced by applying regular audit-review sessions, that can be executed by 3rd parties in the form of SOC 1 or SOC 2 reports. Concerning critical topics like auditing robustness, automation, and availability, IACGPs can be used by AISs and NLP based ML_ACGF.

4. AISs AND NLP'S INTEGRATION

4.1. The Needed AISs Basics

This section describes the used AISs, ICS, NLP, common concepts and resources that are related and are important for this article and RQ. The descriptions are basic in order to help the valuable reader to follow this complex article. The PAIS and AI based transformations, include many sub-domains like Human2Computer linguistics (and communication), Vision processing, (Re)Planning, Robotics (and automation), NLP and Models (NLPM), Decision-making sciences, and many others; and as illustrated in Figure 5, PAISs' topics and actions like [19,20,21,22,23]:

- NLP processes HFI-languages, NLs, and interprets IDSs by searching, and analysing information from inputted-text.
- Large Language Models (LLM) are setup and trained for any type of inputted-text to mimic HFI's analysis.
- ML is applied and algorithms are used (and is sufficiently trained) to enable ICSs to apply Supervised, Unsupervised, or Reinforcement Learning.
- Deep Learning (DL) uses DA for (self)learning, and delivers acceptable results; and it focuses on: 1) Human emotions; 2) Identifying objects on pictures; 3) Recognizing and memorising voices of different people; 4) Processes videos and audios…
- Neural Network (NN) manages cognitive-science in ICSs, to be able to launch knowledge actions and it mimics the HFI. It also includes a large number of neurons, to simulate HFI-neurons.

- Cognitive processing supports Human-Machine Interaction (HMI) to accomplish HFI like tasks (and solving problems).
- Vision processing identifies, and analyses visual-inputs and it uses DL.
- Expert-systems are specialized in unique tasks and act like HFI-experts.
- Robotics-automation focuses on programmed-nodes that automate sets of actions, and are managed by an HFI.
- Fuzzy-logic subsystems solve problems by delivering a Boolean value: “true or false”.
- There also associated subdomains like DA, Data Science (DS), Big-Data (BD), Business Intelligence (BI)…
- DA uses proven IDS to interface structured data-sources and uses also historical-data and predictive-modelling.
- DS is applied before DA processing, to prepare, extract insights, and filter massive IDSs.
- DA generally processes question(s) and results, and DS has a macro-view that proposes precise-questions.
- DA and DS use BD to manage multiple data-sources.
- BD manages huge volumes of structured and unstructured data, and uses the three V’s: Volume, Velocity, and Variety.

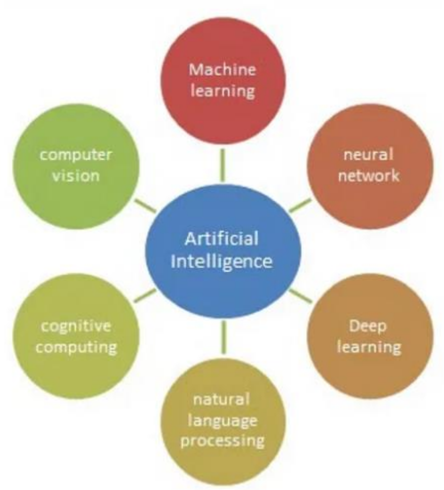


Figure 5. AISs’ main topics [19].

The mentioned PAIS enables ML_ACG’s application and IACGPs’ integration.

4.2. NLP Basics

In the context of ML_ACG, NLPEs support [24,33]:

- The NLPE combines ICS, AISs, algorithms, and linguistics technics (or language studies), to support the understanding, processing, and creation (or generation) of human NL, in a manner that results have (make) sense, logic, and are useful.
- The actual hyper-evolution of volumes of text, and data from social medias, websites, and other sources, makes NLPEs strategic and critical environments. Because they deliver insights and scripts for the automation of complex NLP-ACG’s tasks, such as text analysis or text translations.

- As shown in Figures 6 and 7, NLPEs are used in many APDs that need to use languages, like text-translation, voice-recognition, text-summarization and chatbots. They can also support businesses' improvements of their efficiency, and performance, by applying lingual simplifications.
- NLPEs encompass a long list of technics and capabilities that enable ICS' nodes (ICSN or Servers) process and understand NLs. And these technics are: 1) Text processing and preprocessing; 2) Syntax and parsing; 3) Semantic analysis; 4) Information extraction; 5) Text classification; 6) Language generation; 7) Speech processing; 8) Question answering; 9) Dialogue systems; and 10) Sentiment and emotion analysis.
- Text processing and preprocessing include: 1) Tokenization, which divides text (or string) into smaller units, like words or sentences; 2) Stemming and lemmatization, reduce words to their base (or root-forms); 3) Stop-word removal, removes common words (like and, the, or, is…) which do not contain significant meanings; and 4) Text normalization which standardizes text, including case normalization, removing punctuation and correcting spelling-errors, which is optimal for NLP-ACG's reports.
- Syntax and parsing include: 1) Part-Of-Speech (POS) tagging to assign parts of speech to each word in a paragraph; 2) Dependency parsing analyses the grammatical-structure of a sentence to localize relationships between words; and 3) Constituency parsing breaks-down sentences into its constituent parts or phrases (like verb phrases).
- Semantic analysis includes: 1) Named Entity Recognition (NER) which identifies and classifies objects-entities in text, like names of people, locations…; 2) Word Sense Disambiguation (WSD) determines the meaning of a word in a given context; and 3) Coreference resolution, identifies different words which refer to the identical object.
- Information extraction includes: Entity's extraction identifies specific objects and their relationships within the paragraph-text; and 2) Relations' extraction, identification, and categorizations of relations that exist between objects in the inputted text.

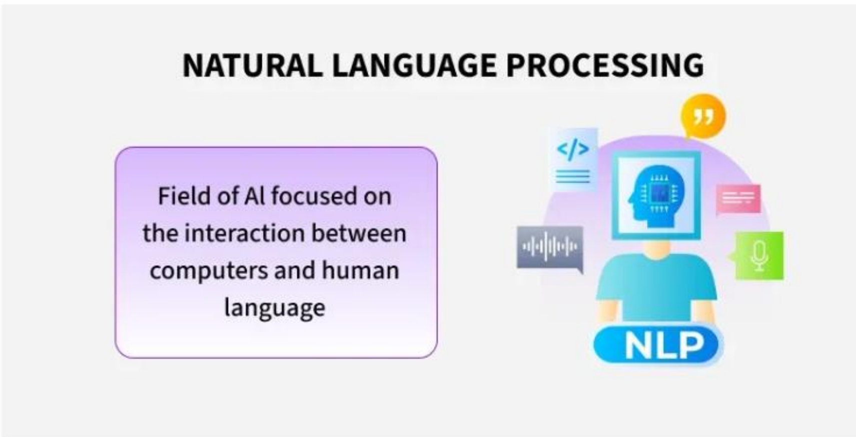
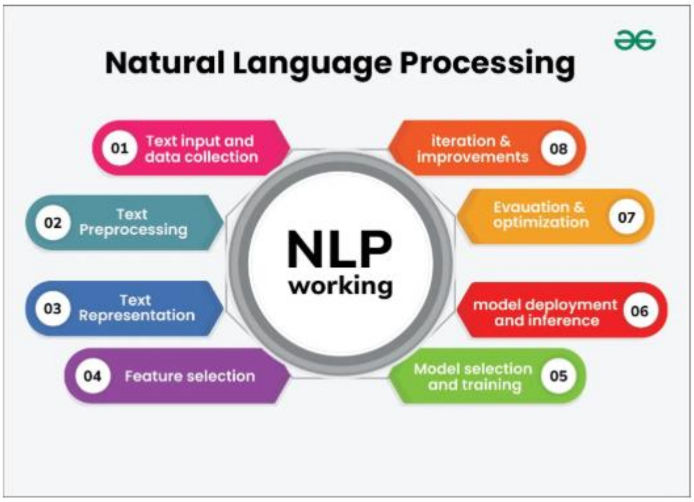


Figure 6. NLP's interactions [24]

- NLP can be adapted to ML to offer a hybrid NLP-ML optimal approach, which makes AISs usage less expensive and more efficient, offering direct answers to all types of queries.
- Pretrained ML based systems are actually available to streamline various applications of NLPIE.
- Text classification includes: 1) Sentiment analysis determines the sentiment (or emotional tone-status) expressed in a text (like positive, negative, or neutral); 2) Topic

modelling identifies topics (or themes) in a huge volume of documents; and 3) Spam detection classifies text as “spam” or “not spam”.

- Language generation includes: 1) Machine translation of text; 2) Text summarization generates precise summaries; and 2) Text generation automates the generation of coherent and contextually relevant text.
- Speech processing includes: 1) Speech recognition and conversion of spoken language into relevant text; and 2) Text-to-Speech (TTS) synthesizes by converting written-text into spoken-language.
- Question Answering (QA) includes: 1) Retrieval-based QA to find and return the most relevant response (text-passage) as the response to the query; and 2) Generative QA (GQA) generates answers based on information available in a text volume.
- Dialogue systems include Chatbots and Virtual Assistants which enable subsystems to participate in conversations with users, delivering responses and execute tasks activated by user’s input.
- Sentiment and Emotion Analysis (SEQ) includes: 1) Emotions’ detection by identifying and categorizing emotions found in the text; and 2) Opinion mining analyses opinions or reviews to comprehend public sentiments toward various fields like products, services.



NLP Working

Figure 7. NLP’s domains [24]

4.3. NLP Models

As already mentioned, NLP processes include ICS’ computational power and specialized NLPMs needed to understand HF’s language(s); and such processes include language’s understanding, generation, and interaction capabilities. NPLM’s include the selection and training of ML Models (MLM), and/or DLMs to support ML_ACG processes by offering [24]:

- Supervised learning that uses labelled data or IDSs, to train needed models like Support Vector Machines (SVM), Random Forests, or DLMs like Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs).
- Unsupervised Learning that applies technics like clustering or topic modelling.
- Pre-trained models that use pre-trained language models such as BERT, GPT or transformer-based models that have well trained.

4.4. NLPs and IACGPs

The NLP and LLM support IACGPs to offer the following advanatges and characteristics [25]:

- NLPs and AISs are transforming the financial-services industry by offering automated analysis of huge-volumes of unstructured data/IDSs, as shown in Figure 8.
- Enhance ML_ACG's efficiency and accuracy, in the domain of financial-reporting.
- Processes unstructured data/IDSs from financial (and audit) reports, contracts, and market-analysis, to transform them into well-structured, and usable insights.
- Supports automation to reduce the needed time required for the original manual data-entry and analysis, and enables faster and accurate financial analysis operations.
- Generates personalized financial-reports, indicating errors and standardizing IDSs or data, which reduces costs.

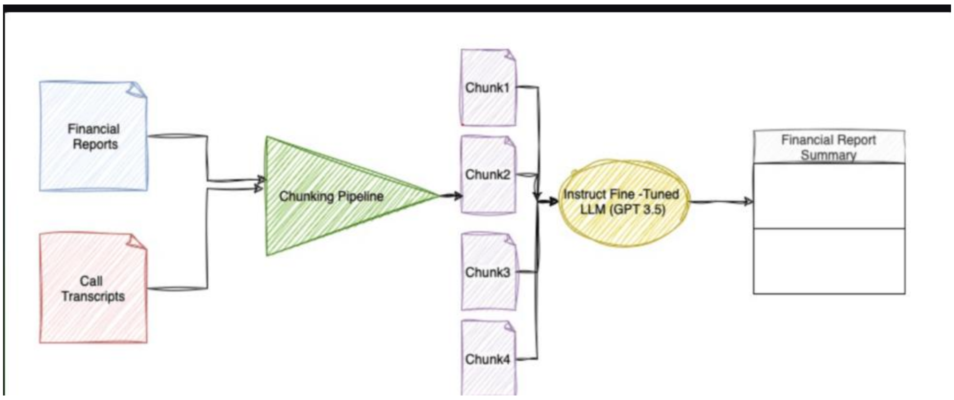


Figure 8. NLP's and LLM's interactions to support IACGPs' integration [25]

- Supports the Fraud Detection Systems (FDS), to identify and prevent fraud-based actions in various APDs, especially in the financial world. This is achieved by using NER to analyse transactions' descriptions, clients' communications, and claims narratives, to discover hidden-patterns and issues which can be cases of fraud.
- Supports audit documents' processing pipeline (or AIAWF) in accounting activities by using automated management and analysis of financial-documents, like invoices, purchase orders, receipts, and financial statements.
- Applies AIAWFs in finance to enhance efficiency, accuracy, and scalability, and as shown in Figure 9, AISs process huge IDSs of financial documents, like invoices, by using Optical Character Recognition (OCR) and that enhances documents WFs.
- Summarizes lengthy financial-reports and disclosures, like annual-reports and regulatory-filings, which are usually complex and have legal specific text. This permits auditors to accurately assess financial health reports.
- Supports financial news and market sentiment analysis, by deriving actionable insights from financial-news and predict market trends. Market insights are extracted from financial-news (or articles), this process include parsing texts to find mentions of specific keywords like economic events…
- Supports Real-Time (RT) financial risk assessment which uses sentiment analysis with selected financial indicators, to enable the assessment of risks' levels of investments.
- Supports AIS-based financial forecasting which includes: 1) Auto-regressive actions that capture the relationships between actual observation and their previous values; 2) Integrated actions for time-series IDSs to exhibit trends or seasonality; 3) Moving average, models relationships between observations and residual errors from previous

processes forecasts; and 4) SARIMA which exhibits regular, repeating-patterns, like monthly (or quarterly) financial gains.
- Supports the automation of compliances' checks that includes labour's-intensive processes that analyses huge IDSs of legal and financial-documents for specific regulatory requirements.

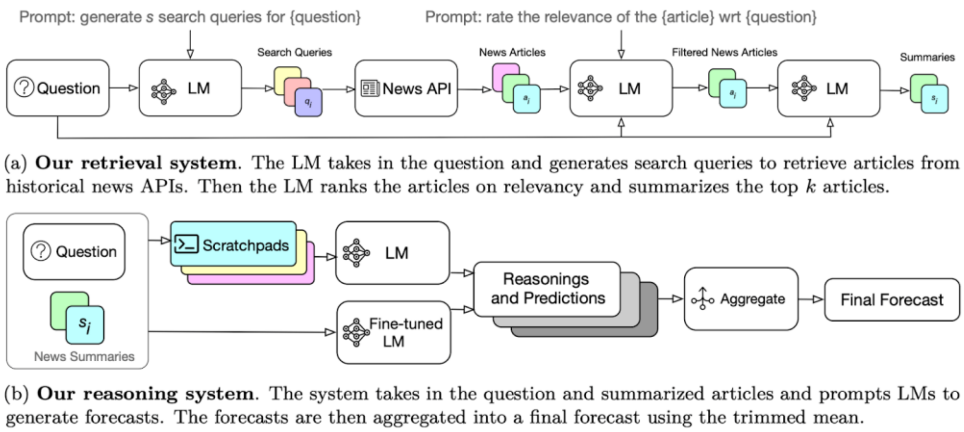


Figure 9. NLP's and AIAWFs interactions enhance various IACGP activities [25]

5. SETTING THE IMPLEMENTATION AND OPERATIONS ENVIRONMENTS

5.1. Agentic AI

The ML_ACGF uses the following AIA's features, capabilities and components [26,27]:
- AIAs in AISs refer to autonomous sub-systems (or component-program) which perform tasks on behalf of a human (or user) or another sub-system; which needs designed WFs and tools.
- Use LLMs that are important for processing (and generating) NLs; and use specific parameters to tune the outputs and final results.
- The required tools for LLMs' management are used for acquiring and processing IDSs, and they dependent on well-trained models. Such tools include IDSs' filters, Web-searches (or crawlers), and Advanced Application Programming Interfaces (API) sub-systems.
- Feedbacks' management scripts and mechanisms like the Human-In-The-Loop (HITL) (or other)…
- AIAs support Intelligence to tune the set of possible solutions.
- Prompt and concise engineering processes support AIAWFs' performances that dependent on the modelling and the quality of provided prompts. Where such an engineering concept supports GenAI Models (GenAIM) to better understand, comprehend, and respond to a wide range of questions/queries. It can also use Chain of Thought (CoT), one-shot, zero-shot, and self-reflection concepts.
- Multi-AIA (MAIA) collaboration and communication interfaces for coordinated processing and solving problem-types in the context of MAIS Systems (MAIAS).
- MAIASs are optimal for complex problem-types and advanced ACSs. Where every agent in MAIAS is linked to a set of environments/tools, AIS-algorithms, and APDs.
- AIAs share gained practical experiences and empower related (I)LPs which are also shared with existing MAIASs.

- Project's integration processes, synchronize used NLP-ACG processes, AIAFWs, which need to be well-integrated in the ICS. The ICS depends on the defined goals and requirements.
- Data-storages' integration and consolidation is the 1st step; and there are commercial products for such integration processes, like LangChain, LangGraph, crewAI and IBM's BeeAI; and they are AIA's Orchestration Frameworks (AIAOF).

5.2. AIAWFs' Integration

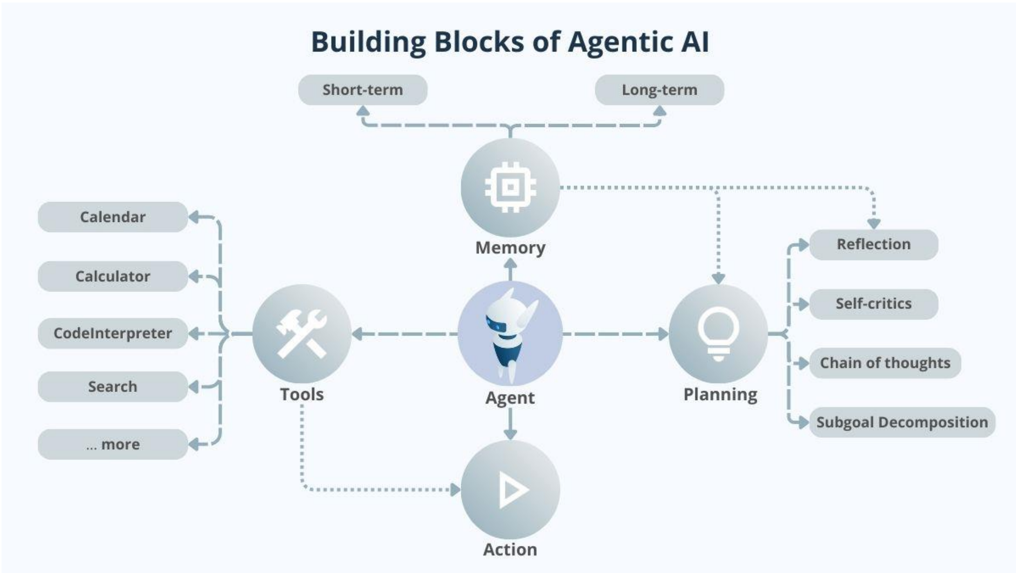


Figure 10. AIA's building blocks [26]

The ML_ACGF uses AIAWFs that are implement and offer services as follows [26,27]:

- They are AIS-driven processes, in which autonomous AIA's blocks, as shown in Figure 1à, deliver decisions, actions, and manage/coordinate tasks with non (or minimal) HF's interventions.
- WFs leverage core modules and components of AIA like reasoning, planning, and an environment to launch complex tasks.
- Are dynamic and flexible so they can adapt to RT data-streams, and complex/unexpected constraints and conditions.
- Has a concept for complex problems that need an iterative approach, to enable AIAs, to unbundle BPs, and to adapt dynamically to different situations, and to refine needed IACGPs.
- Uses GenAI to handle intricate WFs, which enhances operational efficiency, scalability, and Intelligence's capacities.
- Use an NLPE, and related AISs, to automate and optimize various types of processes and to reduce HFs' interventions.
- Empower AI Models (AIM) and CModels, in various APDs.
- Main steps are: 1) Understanding the problem-type by gathering detailed information; 2) Executing diagnostic steps; 3) Using adaptive tools for various APDs; 3) Applies an iterative approach; and 4) Finalizing and learning.

- The AIAWF's core components are: 1) AIAs; 2) LLMs; 3) Tools for managing LLMs; 4) Feedbacks' mechanisms for analysis; 5) Prompt engineering; 5) MAIAs' collaboration and communication; and 6) Integration processes;
- Makes the ICS adaptable to frequent changes and minimizes HF's interventions.
- Includes various building blocks, as shown in Figure 10, like: 1) Planning WFs to achieve the defined goals, and includes reflection, self-critics…; 2) Tools, like Calculator, Code-interpreter…; and 2) Actions needed to complete IACGPs assigned by the audit plan.
- Offers blocks to build an optimal ACG compliant AIA Architecture(s) (AIAA).

5.3. An ACG Compliant AIAA

An AIAA is related to AIAWFs, but differ in how they serve various AISs' functionalities. In the context of ML_ACG and AIA is used by auditing processes for interactions. The AIAA shapes Entity's virtual-space and AIAWFs' structure and support the automation of CAIMs and AISs. AIAA supports and regulates the behaviors of AIAs collaborating with GenAI components; and to provides a structured context for LLMs to automate AIAs to finalize complex tasks. AIAWFs offer the steps that an AIA follows and executes to assist the auditing-process; where the AIA is autonomous and decides which set of optimal actions should be executed. But such processes need a well-designed architecture (or AIAA), which is capable of apprehending ML_ACG's inputs, learn from previous interactions (and experiences), and improve solutions. An AIAA [28,29]:

- Enables the AIA(s) to process data, deliver decisions, and learn from experiences.
- Its primary aim is to enable AIAs to think, learn, based on IDSs, and previous experiences.
- Includes components like Intelligence's algorithms, data-processing modules, and LPs (and models).
- Defines the role of an AIA, and its basic infrastructure and capabilities.
- Enables flexibility to adapt and deliver decisions in various complex contexts, ensuring that AIAs work across different processes or tasks; and enhancing interoperability.
- Adapts autonomously to new contexts and situations based on LPs, and delivers decisions.
- Applies continuous optimization processes to tune AIA's Intelligence's (especially the DMS) by improving LPs and DA's processing.
- Uses various types of interfaces like APIs.
- Includes modules that address Planning; Monitoring ICS' performance; Forethought (and Self-reactiveness); and Self-reflectiveness.

5.4. ACG Framework's Integration

The ML_ACG has to interface and integrate existing major market and standard ACG standards and frameworks which are relevant for AI and classical ICS implementations and operations. AISs driven APDs need accountability mechanisms and that implies that the Entity must implement: Comprehensive internal audit procedures and controls (similar to IACGP); Risk biases mitigations; Eventual misuse of monitoring; and Regulatory roadblocks. The LPE_AGF ensures that risks are minimized and mitigated, and that means that regulators have to improve IACGPs to block issues and biases that may exist in AIS' algorithms. IACGPs integrate AI-audibility into AIAWFs to manage the mentioned risks; by including: Ensuring governance mechanisms; Reducing security risks; Supporting Project's lifecycle(s); Applying architecture/design; and Offering Development, Deployment, and Monitoring phases. The

LPE_AGF is not just a documentation process, because it includes also tracking inputs-outputs, and the application of MLs. It privileges a proactive approach which enables auditors to discover issues before they can cause major damages; like in the case of a breached algorithm… Existing regulations like the AI-Acts and GDPR, focus on the importance of protecting data-privacy and on the implementation of risks' management. Which implies that auditors must take into account HFs' aspects, in order to ensure that AISs do not interfere with ethical Intelligence (or DMS) [30]. As shown in Figure 11, IACGPs are affected by Entity's EA and audits' executive management, and in the same time they offer assurance for achieving Project's objectives that are mainly related to: Operations; Reporting; Compliance; and Re-structuring actions [31,32].

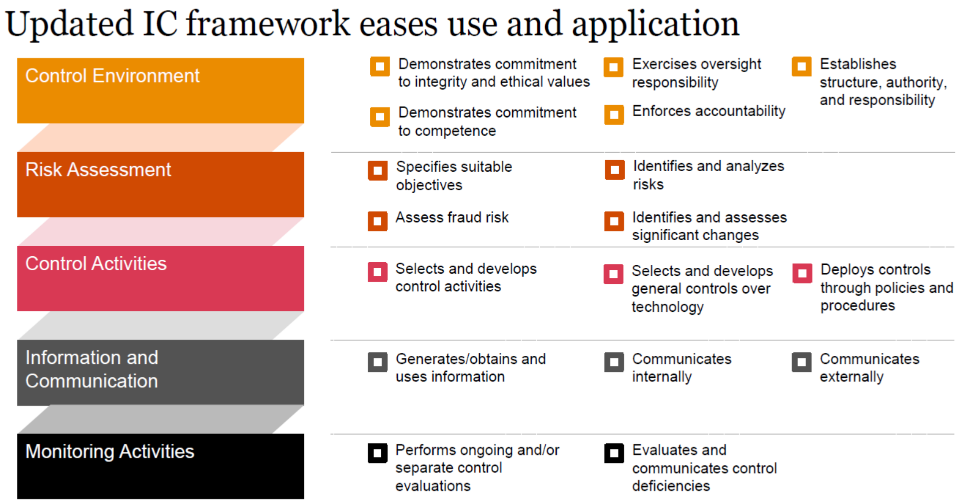


Figure 11. IACGP's applications [31]

6. Conclusions

This article proposes the following set of ML_ACG's recommendations:

- The RDP is related to ML_ACG, AIAs, AIAWFs, ADM, and CModels; and uses an adapted mixed research method.
- The proposed AISs and avant-garde technologies replace legacy ICS components.
- Entity's efficiency, and sustainability depend on ML_ACGF's and IHIPTF's integration.
- A Project adopts a Polymathic (and holistic) approach and uses CModels.
- Intelligence mitigates risks and solves problem-types.
- NLP can be adapted to ML to offer a hybrid NLP-ML which is an optimal approach.
- The ML_ACG and IACGPs replace traditional auditing processes.
- The ML_ACG interfaces the IHIPTF and other frameworks.
- The IHIPTF integrates various methodologies like TOGAF and interfaces other ACG frameworks.
- The ADM supports ML_ACG's activities, and CModels' implementations.
- CModels enable the integration of AIAWFs, and related AISs.
- The ML_ACG supports RT monitoring, and tracing.
- The ML_ACG supports IACGPs' integration for Entity's auditing processes.
- The ML_ACGR fits in the Entity's global transformation strategy.
- The ML_ACG modules' design, development, and operations are managed by MLOps.
- The implementation environment and CPoC used Factors to link AISs' components.
- The CPoC's outcome proves that the ML_ACG is feasible [34].

ACKNOWLEDGEMENTS

The author has reviewed and edited the output and welcomes all types of comments and feedbacks.

REFERENCES

[1] Cranmer, M. Polymathic AI: Foundation Models for Science. University of Cambridge, 2023.
[2] Envisioning. Polymathic AI. Envisioning. 2025. https://www.envisioning.io/vocab/polymathic-ai
[3] Deloitte. COSO – Control Activities. Deloitte. 2020. https://www.deloitte.com/ng/en/services/audit/perspectives/coso-control-activities.html
[4] Trad, A. The In-House Polymathic Transformation Framework (IHIPTF)-The Guide. IBISTM. France. EU. 2024. https://www.ibistm.org/docs/PRWC.pdf
[5] Trad, A. The In-House Polymathic Transformation Framework (IHIPTF)-The Glossary. IBISTM. France. EU. 2024. https://www.ibistm.org/docs/Glossary.pdf
[6] Trad, A. The In-House Polymathic Transformation Framework (IHIPTF)-The Syllabus. IBISTM. France. EU. 2024. https://www.ibistm.org/docs/Syllabus.pdf
[7] Trad, A. The Applied Polymathical/Holistic Mathematical Model for Enterprise Transformation Projects (AHMM4PROJECT). IBISTM. France. EU. 2024.
[8] BIE Executive. Guide to Business Transformation. 2025. https://www.bie-executive.com/guides/guide-to-business-transformation/
[9] Trad, A. Polymathic Approach for Enterprise Transformation Projects: Implementing a Business Meta-model that Respects Environmental Sustainability (PETP-IBMMES). Book: Financial, Economic, Educational, and Technological Determinants of Environmental Sustainability. Chapter 8. Cambridge Scholars Publishing. 2024.
[10] Trad, A. Enterprise Transformation Projects-The Applied Polymathical/Holistic Mathematical Model for Enterprise's Business Process Modelling (AHMM4EBPM). Journal: International Journal of Computers. Volume 9. International Association of Research and Science. iaras.org. 2024.
[11] Burke, P. The Polymath a Cultural History From Leonardo Da Vinci To Susan Sontag. Yale University Press. New Haven and London. UK. 2020.
{12] O'Riordan; B. INNOVATION-Why Transformations Fail And How They Can Succeed With People Power. Forbes. 2021.
[13] Trad, A. Organizational and Digital Transformation Projects-A Mathematical Model for Building Blocks based Organizational Unbundling Process. IGI. USA. 2023.
[14] De Miguel, M., Exertier, D., & Salicki, S. Specification of model transformations based on meta templates. Thales. 2002. https://www.researchgate.net/publication/228953908_Specification_of_model_transformations_based_on_meta_templates.
[15] Easterbrook, S., Singer, J., Storey, M., & Damian, D. Guide to Advanced Empirical Software Engineering-Selecting Empirical Methods for Software Engineering Research. F. Shull et al. (eds.). Springer. 2008.
[16] Quinlan Ch. Business Research Methods. Dublin City University. Cengage Learning. Ireland. 2015.
[17] Michael, P. Technology Control Automation: Improving Efficiency, Reducing Risk and Strengthening Effectiveness. ISACA. 2022. https://www.isaca.org/resources/news-and-trends/isaca-now-blog/2022/technology-control-automation-improving-efficiency-reducing-risk-and-strengthening-effectiveness
[18] Finney, J. What Are Internal Controls? The 4 Main Types of Controls in Audits (with Examples). CISA. 2022. https://linfordco.com/blog/types-of-controls/
[19] Rancho Labs. 6 Major Sub-Fields of Artificial Intelligence. Rancho Labs. Medium. 2021.
[20] Garanhel, M. What are the top 7 branches of artificial intelligence? AI Accelerator Institute. 2023. https://www.aiacceleratorinstitute.com/tag/artificial-intelligence/
[21] OMNI-SCI. Data Science - A Complete Introduction. OMNI-SCI. 2021. https://www.omnisci.com/learn/data-science

[22] Riedl, M. A Very Gentle Introduction to Large Language Models without the Hype. Medium. 2023. https://mark-riedl.medium.com/a-very-gentle-introduction-to-large-language-models-without-the-hype-5f67941fa59e

[23] PMI. Leading AI-driven Business Transformation. 2022. Project Management Institute. https://www.pmi.org.

[24] Geeks. Natural Language Processing (NLP). Geeks. 2025. https://www.geeksforgeeks.org/natural-language-processing-overview/

[25] Mathavan. The NLP in Accounting-NLP and LLM Applications in Accounting. Mercity. 2023. https://www.mercity.ai/blog-post/nlp-and-llm-in-accounting

[26] Stryker, C., & Gutowska, A. What are agentic workflows? IBM. USA. 2025.

[27] FabriXAI. Introduction to Agentic AI and Agentic Workflow. FabriXAI. 2025.

[28] Codewave. Understanding Agentic Workflows: Patterns and Use Cases. Codewave. 2025. https://codewave.com/insights/agentic-workflows-patterns-use-cases/

[29] Winland, V., Bozorg, J., & Stryker, C. What is agentic architecture? IBM. 2025. https://www.ibm.com/think/topics/agentic-architecture

[30] Auditboard. 5 AI Auditing Frameworks to Encourage Accountability. Auditboard. 2024. https://auditboard.com/blog/ai-auditing-frameworks

[31] PwC. Fine tuning your internal controls with COSO. PwC. 2019.

[32] Leland, A. Fundamentals of the COSO Framework: Building Blocks for Integrated Internal Controls. Audit Board. 2024. https://auditboard.com/blog/coso-framework-fundamentals

[33] Kili Technologies. Natural Language Processing (NLP) for Machine Learning. KILI. 2025. https://kili-technology.com/data-labeling/nlp/natural-language-processing-machine-learning

[34] Trad, A. AI Based Transformation Projects-Language Processing Environments for Audit, and Governance (LPE_AG)-The Proof of Concept. MLAIJ. 2025.

AUTHORS

I am a holder of a Dr degree in computer sciences degree and a Dr degree in business administration (DBA), with more than 500 articles and projects related to transformation initiatives. Actually, I am a researcher/professor at IBISTM.

My research fields are related to Enterprise Transformation Projects, AI Subdomains based Mathematical Models, Enterprise Architecture… And the needed educational curricula.

This work is intended for future Managers, Professors, Architects, and Engineers in Complex Business Innovation and Transformation Projects. My Research and Development Projects (RDP) have the lead in the mentioned fields. The educational curricula include: complex projects, AI, and enterprise architecture in the context of digital transformations.

After many years working in and researching the mentioned fields, I realize that there is a need for a major change in education; especially in cross-functional domains. And there is the need for a Polymathic approach as presented in my research chapter.

ABBREVIATIONS

The following abbreviations are used in this manuscript:

ACG	Audit, Control, and Governance
ACGA	ACG (or control) Activities or Actions
ADM	Architecture Development Method
AIA	Agentic AI
AIADP	AIA Design Pattern
AIAOF	AIA’s Orchestration Frameworks
AIAWF	AIA WorkFlows
AIS	AI Subdomains

AGI	Artificial Generic Intelligence
AHMM	Applied Holistic MM (AHMM)
AHMM4AI	AHMM for Integrating an AI
APD	APplication Domains
API	Application Programming Interfaces
AR	Action Research
BP	Business Processes
BPM	BP Model
CAIM	Composite AI Model
CModel	Composite Model
CSA	Critical Success Area
CSA_DT	CSA Decision Table
CSF	Critical Succes Factor
COBIT	Control Objectives for Information and Related Technologies
COSO	Committee of Sponsoring Organizations
DA	Data-Analytics
DMS	Decision-Making System
DL	Deep Learning
DLM	DL Model
DT	Digital Transformation
EA	Enterprise Architecture
EAM	EA Model
ERM	The COSO Enterprise Risk Management
EUP	Entity’s Unbundling (and transformation) Processes
Factors	Is CSFs, CSAs, KPIs, and VARs
GAPA	GAP Analysis
GAO	US Government Accountability Office
GenAI	Generative AI
GenAIM	GenAI Models
HF	Human Factor
HFB	HF Brain
HFI	HF Intelligence
IACGP	Internal ACG Procedures
ICS	Information and Communication Systems
ICSM	ICS Models
IDS	In-Memory DataSets
IHI	In-House Implemented
IHIDT	IHI Heuristics-Decision-Tree
IHIPTF	IHI Polymathics Transformation Framework
IHIPTFM	IHIPTF Method or Methodology
ILP	IHDT based LP
Intelligence	KMS and DMS
ISACA	Systems Audit and Control Association
KMS	Knowledge-Management System
KPI	Key Performance-Indicators
LP	Learning Process
LPC	LP Container
MAIA	Multi-AIA
MAIAS	MAIS Systems
MCP	Model Context Protocol
ML	Machine Learning
MLM	ML Model
MLOps	ML Operations
MM	Mathematical Model
NER	Named Entity Recognition
NL	Natural Language
NLG	NL Generation

NLP	NL Processing NL
ML_ACG	NLP for ACG
ML_ACGF	ML_ACG Framework
ML_ACGR	ML_ACG Roadmap
NLPE	NLP Environnements
PAIS	Polymathics AIS
PDPC	Singapore's Personal Data Protection Commission
PEERM	Polymathic-Empirical Eng Research Model
POS	Part-Of-Speech
PLRP	Polymathic Literature Review Process
XHFR	eXtremely High Failure Rates
VAR	VARiables
WF	WorkFlows