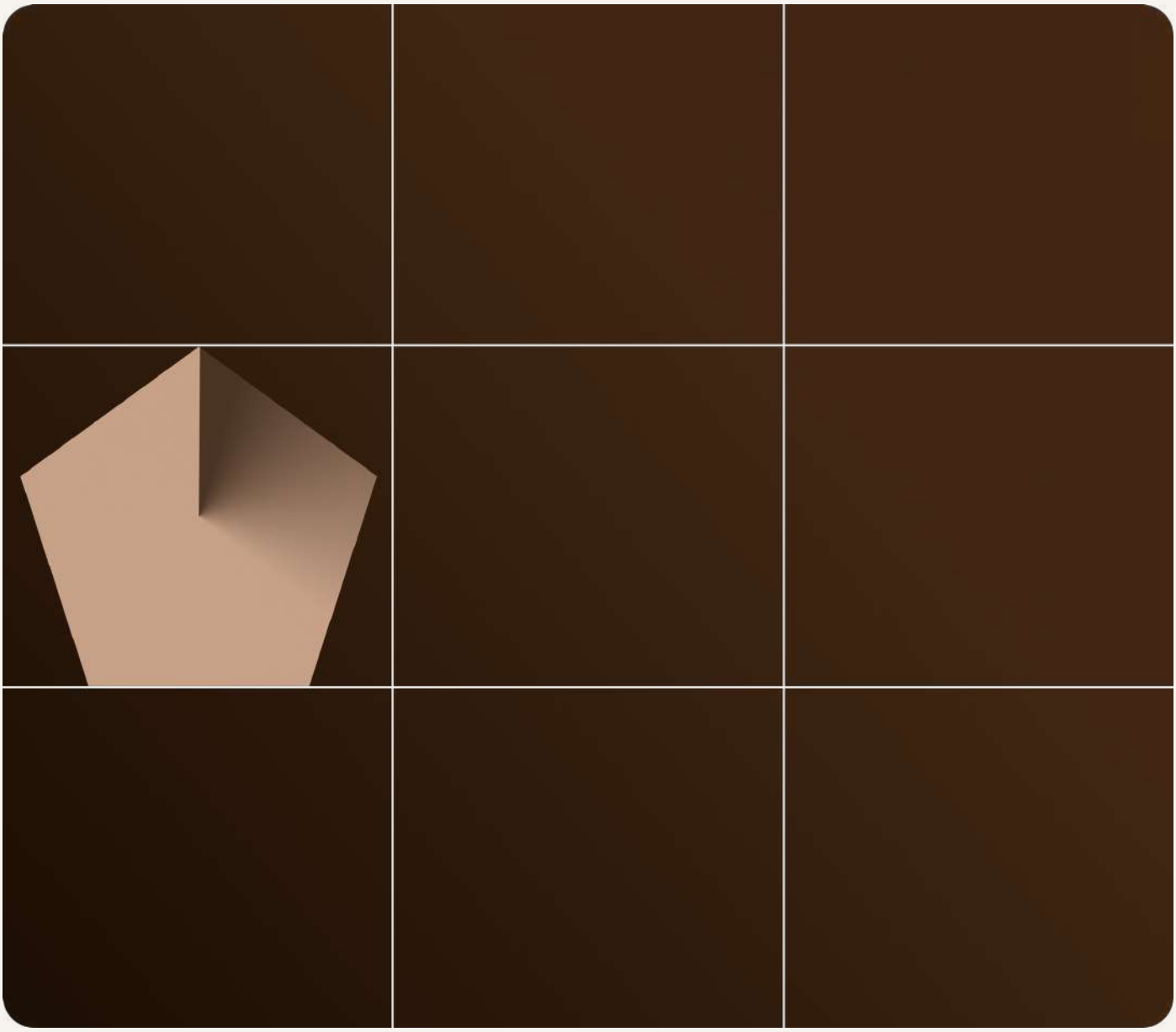
Project Management Institute



PMI-CPMAI™

PMI Training

PMI-CPMAI Workbook

PMI-CPMAI Training & Certification

LAST UPDATED: July. 2025

Check periodically for additional revisions

Table of Contents

About PMI-CPMAI and This Workbook **5**

About this Workbook 5

How to Use this Workbook 5

CPMAI Hierarchy **6**

Mapping Generic Tasks to Specialized Tasks 6

CPMAI Phases **7**

Phase I: Business Understanding 8

Phase II: Data Understanding 8

Phase III: Data Preparation 8

Phase IV: Model Development 8

Phase V: Model Evaluation 8

Phase VI: Model Operationalization 9

Overview of the Phases With Generic Tasks Per Phase 9

CPMAI Phase I: Business Understanding **11**

Task Group: Determine Business Objectives **11**

Subtask: Determine Business Objectives 11

Subtask: Determine Business Success Criteria 13

Subtask: Cost-Benefit Analysis 14

Task Group: Cognitive Project Requirements **16**

Task: Cognitive Requirements 16

Task: AI Pattern Identification 22

Task Group: Assess Situation **24**

Task: Resource Requirements 24

Task: Schedule Requirements 26

Task Group: AI System Performance and Operation **27**

Task: Acceptable Model Performance Values 27

Task: Acceptable KPI Performance Values 30

Task Group: Trustworthy AI Requirements **32**

Task: Use of Trustworthy AI Framework 32

Task: Required Ethical AI Considerations 33

Task: AI Failure Modes 34

Task: Required Compliance With Regulations and Laws 38

Task: Required AI Transparency Considerations 39

Task: Required AI Explainability Considerations 41

Task Group: AI Go/No-Go **42**

Task: Business Feasibility 42

Task: Data Feasibility 44

Task: Execution Feasibility 46

Table of Contents

Generic Task Group: Produce Project Plan **49**

Task: Determine Business Objectives 49

CPMAI Phase II: Data Understanding **51**

Task Group: Collect Initial Data **51**

Task: Collect Initial Data 51

Task: Describe Data 53

Task: Explore Data and Cognitive Data Requirements 54

Task Group: Data Quality **55**

Task: Verify Data Quality 55

Generic Task Group: Machine Learning Model Data Requirements **57**

Task: Training and Test Data Requirements 57

Task: Edge Model Data Needs 59

Task: Pretrained and Third-Party Model Usage 60

CPMAI Phase III: Data Preparation **62**

Generic Task Group: Data Selection **62**

Task: Select Data 62

Generic Task Group: Data Cleansing and Enhancement **63**

Task: Clean Data 63

Task: Enhance and Augment Data 65

Generic Task Group: Data Labeling **67**

Task: Label Data 67

CPMAI Phase IV: Model Development **69**

Generic Task Group: Select Modeling Technique **69**

Task: Select Cognitive-Relevant Algorithm/Modeling Technique 69

Task: Ensemble Methods 71

Task: Usage of AutoML 72

Task: Fine-Tuning/R-training of Pretrained Models 73

Task: Usage of Generative AI 74

Generic Task Group: Model Test and Validation Design **77**

Task: Generate Model Test Design 77

Generic Task Group: Model Training/Model Building **78**

Task: Model Training/Model Building 78

Generic Task Group: Hyperparameter Optimization **79**

CPMAI Phase V: Model Evaluation **81**

Task Group: Evaluate Model Results **81**

Task: Model Performance Results 81

Task: KPI Measurement 83

Task: Model Iteration Approach 84

Task: Review Process 85

Table of Contents

Phase VI: Model Operationalization **87**

Task Group: Model Operationalization Plan **87**

Task: Operationalization Plan 87

Task Group: Model Monitoring and Maintenance **90**

Task: Monitoring and Maintenance Plan 90

Task: Model Governance Framework 91

Task Group: Determine Requirements for the Next Iteration **93**

Task: Determine Next Steps 93

Task Group: Project Report **95**

Task: Produce Final Report 95

Task: Review Project 95

Sources and Endnotes **97**

About PMI-CPMAI and This Workbook

The PMI® Certified Professional in Managing AI (PMI-CPMAI)™ methodology is a vendor-neutral, data-centric, AI-specific, iterative methodology for running and managing artificial intelligence (AI), machine learning (ML), and cognitive technology projects. This approach to project management borrows and extends upon previous approaches and processes for project management, such as agile practices and the Cross-Industry Standard Process for Data Mining (CRISP-DM), which have both pioneered methods for running large, complex, and constantly changing projects that can respond to continuous needs while also focusing on the data-centric aspects of those projects.

This PMI-CPMAI Workbook extends upon those approaches and processes (whose content and concepts are acknowledged and cited herein) with particular focus on the needs for AI projects. These projects specifically need guidance on how to approach machine learning modeling, model iteration, model operationalization, considerations for model evaluation, and other insights that are not specified within other more general methodologies.

About this Workbook

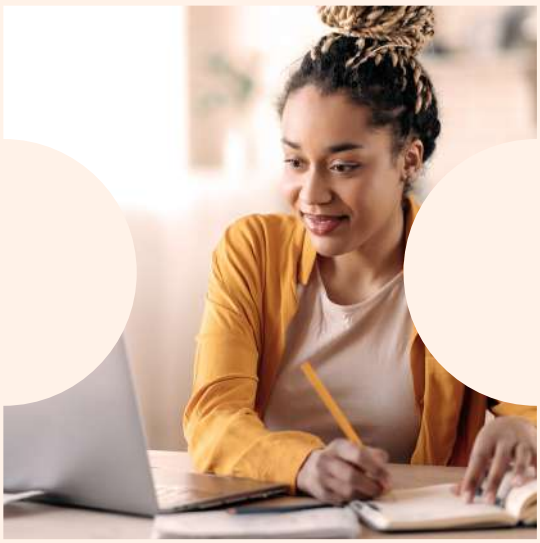
The purpose of this Workbook is to accompany the PMI-CPMAI training and workshops, and provide a tool that implementers of the CPMAI methodology can use to guide their various project iterations.

This workbook itself will be iterated significantly over time, so we encourage users of this Workbook to stay connected with efforts to make sure you are using the latest version available.
Go to learning.pmi.org for updated materials.

How to Use this Workbook

- Please make a copy of this document for each iteration of your project.
- Create a new document for each major iteration so you can keep track of the assets, decisions, and outputs used for each iteration.

To Use this Workbook, MAKE A COPY of this Workbook for EACH project you are running and use the Workbook as your Documentation for Each Iteration



CPMAI Hierarchy

Similar to the CRISP-DM methodology, the CPMAI methodology employs a hierarchy of levels that aims to define the objectives and steps to be performed from the general to the specific. Specifically, the CPMAI hierarchy follows this pattern:

Phases

The highest level of hierarchy that gathers together activities into a grouping, such that all tasks in a phase accomplish a specific objective with regard to the project.

Generic Tasks

A list of tasks to be accomplished for the particular phase to assist in achieving the objective of that phase.

Specialized Tasks

The set of project-specific tasks, adapted from the generic tasks that are specific for a particular project.

Methodology Artifacts

The lowest level of abstraction contains the specific records, outputs, actions, decisions, and results that occurred by executing the specialized tasks defined for the projects that fulfill the goals of the generic tasks that meet the objectives of that particular phase.

An example of an instance of following the CPMAI hierarchy is as follows:

- Within Business Understanding (Phase I), there is a generic task to do an AI Go/No-Go assessment, and specifically within that is the specialized task to do an In/Out (I/O) Flow Reasonability Assessment. The results of the assessment are detailed within a document that details the reasonability and fulfills the task for AI Go/No-Go, which in turn helps to determine the business understanding.

Mapping Generic Tasks to Specialized Tasks

We can map generic tasks to specialized tasks as follows:

- Analyze your specific context.
- Remove any details not applicable to your context.
- Add any details specific to your context.
- Specialize (or instantiate) generic contents according to concrete characteristics of your context.
- Possibly rename generic contents to provide more explicit meanings in your context for the sake of clarity.

Source: CRISP-DM v1.0 Guide

CPMAI Phases

At the highest level of the CPMAI hierarchy are the **phases** (see Figure 1) that separate the activities to be accomplished into a logical group that aim to achieve one primary objective within that part of the project

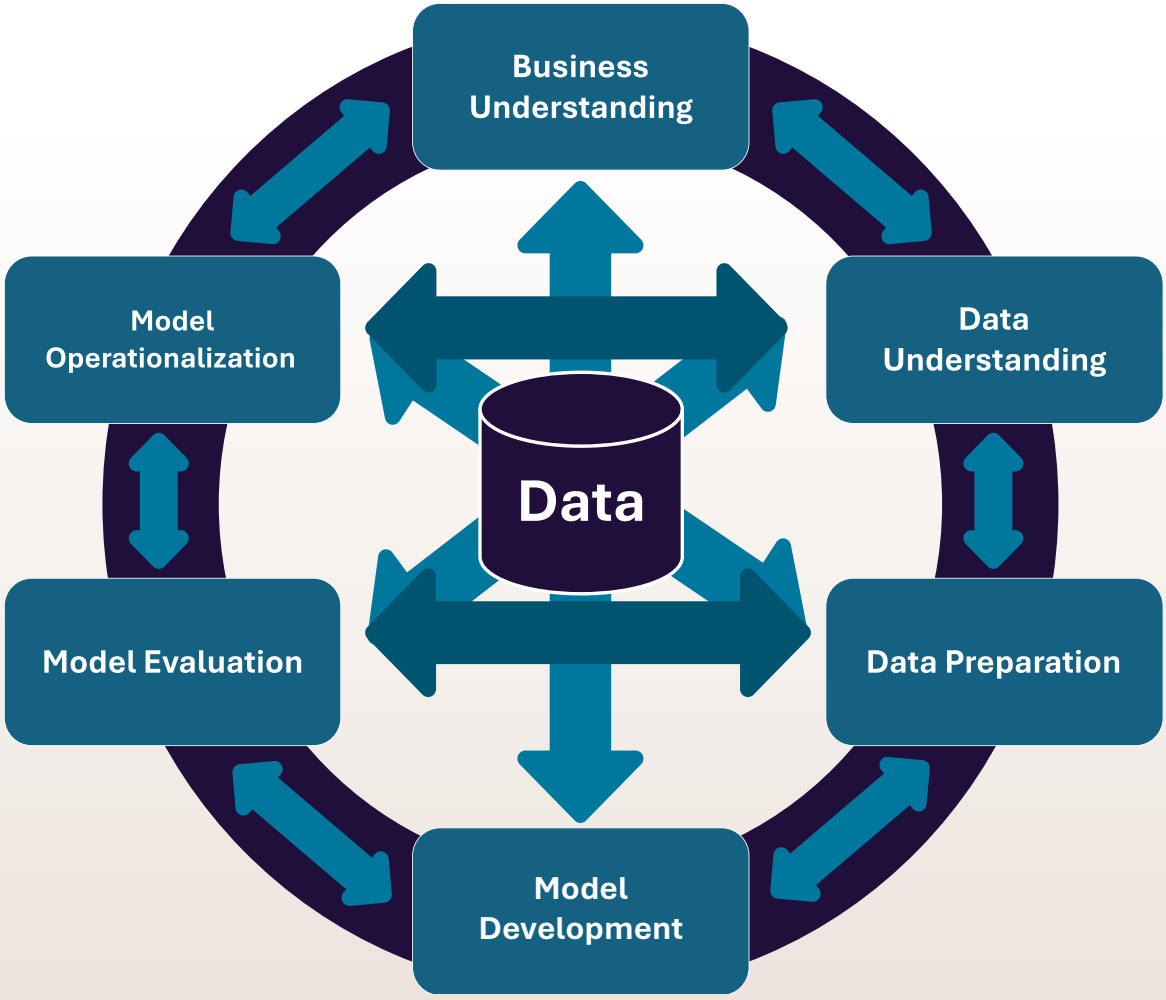


Figure 1. CPMAI phases overview

In a similar fashion to the CRISP-DM methodology, CPMAI utilizes the concept of six primary phases that are to be applied fully for one user story or across multiple user stories as appropriate. Each phase focuses the efforts of the project team on accomplishing objectives for preparing, developing, and operationalizing models that meet specified user objectives and value outcomes.

Unlike CRISP-DM, these phases are all meant to be mutually iterative, which means that if a project team is engaged in activities at the Model Development phase and realizes that there are Data Understanding phase issues, they can shift back to the Data Understanding phase before iterating again on Data Preparation and Model Development phase activities and prior to embarking on Model Evaluation phase activities. Just because you have started a phase does not mean you cannot go back to a previous phase for the same project iteration, as may be required by the business objectives.

Phase I: Business Understanding

The first phase in any CPMAI project is gathering an understanding of the business or organizational requirements. Borrowing from CRISP-DM, but customizing for AI purposes, the Business Understanding phase focuses on understanding the project objectives and requirements from a business perspective, then converting this knowledge into an AI and cognitive project problem definition and a preliminary plan designed to achieve the objectives.

From an agile perspective, the objectives of the Business Understanding phase should tie very closely to the user story that is relevant for that iteration. As mentioned, in a CPMAI project, it is possible for a single sprint or iteration to encompass all of the CPMAI phases, which means that the business understanding will be relevant to that particular iteration's user story.

Phase II: Data Understanding

The second phase in a CPMAI project is the Data Understanding phase, which is focused on data needs identification, initial data collection, data requirements, data quality identification, insights into data, and potential interesting aspects of the data worth further investigation.

Phase III: Data Preparation

The third phase in a CPMAI project is the Data Preparation phase, which focuses on activities needed to construct the data set that will be used for modeling operations. Data preparation includes data cleansing, data aggregation, data augmentation, data labeling, data normalization, data transformation, and any other activities for data of structured, unstructured, and semistructured nature.

Phase IV: Model Development

In the fourth phase of a CPMAI project, we actually embark on the creation and development of machine learning models and other cognitive technology modeling artifacts. This activity includes model technique selection and application, model training, model hyperparameter setting and adjustment, model validation, ensemble model development and testing, algorithm selection, and model optimization.

Phase V: Model Evaluation

Once a model has been satisfactorily created, it should be evaluated to ensure it performs according to the business requirements and other factors set in Phase I of the CPMAI project as well as the other factors as defined in the user story acceptance criteria for that particular sprint or iteration. Model evaluation from an AI perspective includes model metric evaluation, confusion matrix calculations, key performance indicator (KPI) metrics, model performance metrics, and model quality measurements, and a final determination if the model is suitable for meeting the goals of the sprint or iteration, or if earlier phases should be iterated upon to reach those goals.

Phase VI: Model Operationalization

The final phase of the CPMAI methodology is putting the developed model into operation, namely operationalizing it in a manner consistent with delivering the functionality required for the sprint or iteration. Model operationalization may mean deploying the model for use in a cloud environment, edge device, for use within an on-premise or closed environment, or within a closed, controlled group. Model operationalization considerations include model versioning and iteration, model deployment, model monitoring, model staging in development and production environments, and other aspects of positioning the model in a position to provide value to meet the stated purpose. Depending on the requirements, operationalization can be simply the generation of a report from a model to a complex multiendpoint deployment.

Overview of the Phases With Generic Tasks Per Phase

Table 1 outlines each of the phases mapped to generic tasks per phase. The specifics of these generic tasks are outlined in the subsequent sections. Items in bold are the main generic tasks, and items in italics are outputs/artifacts created during the generic task process.

Phase I: Business Understanding	Phase II: Data Understanding	Phase III: Data Preparation	Phase IV: Model Development	Phase V: Model Evaluation	Phase VI: Model Operationalization
Determine business objectives • Business and organizational objectives • Business success criteria • AI success criteria • Necessary KPIs and technical metrics • Desired ROI **Assess situation** • Inventory of resources • Requirements, assumptions, and constraints • Risks and contingencies • Costs and benefits • Tools assessment • AI skills assessment	**Collect initial data** • Big Data environment • Initial data collection report **Describe data** • Data description report • Data source formats • Training data identification • Test data identification • Edge model needs **Explore data** • Data exploration report **Verify data quality** • Data quality report • Training quality • Test quality	**Select data** • Rationale for inclusion/ exclusion **Clean data** • Data-cleaning report • Construct data derived attributes • Generated records **Label data** • Data-labeling needs **Integrate data** • Merged data • Data augmentation **Format data** • Reformatted data • Data anonymization • Data normalization • Data de-noising	**Select modeling techniques** • Modeling technique • Modeling assumptions • Algorithm selection • Ensemble methods **Generate test design** • Test design **Build model** • Model training • AutoML usage • Hyper-parameter optimization • Model descriptions • Pretrained model finetuning or usage	**Evaluate results** • Assessment of data modeling results and business success criteria • Model assessments • Validation results • Confusion matrices • KPI evaluation • Approved models **Review process** • Review of process **Determine next steps** • List of possible actions • Model iteration decision	**Operationalize model** • Deployment plan • Model operational-ization plan • Model scaffolding • Operationalization environment **Monitor and maintain** • Monitoring and maintenance plan • Governance framework **Produce final report** • Final report • Final presentation **Review project** • Experience documentation

Phase I: Business Understanding	Phase II: Data Understanding	Phase III: Data Preparation	Phase IV: Model Development	Phase V: Model Evaluation	Phase VI: Model Operationalization
Outline cognitive project requirements • AI Go/No-Go • Pattern identification Cognitive/ noncognitive parts • Transparency requirements • Acceptable metrics	**Customize pretrained models** • Pretrained and third-party model usage • Transfer learning requirements	**Describe data set** • Data set description	**Assess model** • Model assessment • Model validation • Revised hyper-parameter settings • Scaffolding environment		

Table 1: Generic tasks in each CPMAI phase

CPMAI Phase I: Business Understanding

The first phase of a CPMAI process is gaining a thorough understanding of the business and organizational objectives and other factors that will determine whether the project is worth undertaking, and the conditions under which it will be a success. The Business Understanding phase focuses on understanding the project objectives and requirements from a business perspective, then converting this knowledge into an AI and cognitive project problem definition and a preliminary plan designed to achieve the objectives.

As mentioned earlier, since we are doing CPMAI in the context of agile, the business understanding for the particular CPMAI project is relevant to the specific sprint or iteration you are currently in, and should tie very closely to the user story that is relevant for that iteration. In a CPMAI project, it is possible for a single sprint iteration to encompass all of the CPMAI phases, but it is also possible for a CPMAI project to span multiple sprints or iterations, which means that the business understanding will be relevant to that particular iteration's user story.

Task Group: Determine Business Objectives

These tasks help the project team determine overall business objectives as they are relevant to the AI and cognitive aspects of the project.

Subtask: Determine Business Objectives

Description

The first objective of the project team is to thoroughly understand, from a business perspective, what the customer really wants to accomplish in a manner that is consistent with cognitive technology goals. Often the customer has many competing objectives and constraints that must be properly balanced. The project team's goal is to uncover important factors, at the beginning, that can influence the outcome of the project. A possible consequence of neglecting this step is to expend a great deal of effort producing the right answers to the wrong questions.

Task Artifacts

Background: Record the information known about the organization's business situation at the beginning of the project.

Business objectives: Describe the customer's primary objective from a business perspective. In addition to the primary business objective, there are typically other related business questions that the customer would like to address. For example, the primary business goal might be to keep current customers by predicting when they are prone to move to a competitor. Examples of related business questions are "How does the primary channel used (e.g., ATM, branch visit, internet) affect whether customers stay or go?" or "Will lower ATM fees significantly reduce the number of high-value customers who leave?"

What problem are you solving with AI in this iteration?

Subtask: Determine Business Success Criteria

Business success criteria: Describe the criteria for a successful or useful outcome of the project from the business point of view. This might be quite specific and able to be measured objectively, for example, reduction of customer attrition to a certain level, or it might be general and subjective, such as "give useful insights into the relationships." In the latter case, it should be indicated who makes the subjective judgment.

What are the objective measures of success for this project iteration?

Subtask: Cost-Benefit Analysis

Costs and benefits: Construct a cost-benefit analysis for the project, which compares the costs of the project with the potential benefits to the business if it is successful. The comparison should be as specific as possible. For example, use monetary measures in a commercial situation.

What is the cost and time budget for this project?

What is the expected ROI for this project?

Task Group: Cognitive Project Requirements

In this set of tasks, we address business understanding objectives that are specifically relevant to cognitive projects. Not all business requirements need to be met with cognitive projects when simpler or more direct or more programmatic/deterministic/heuristic methods apply. The objective in this task group is to uncover those cognitive relevant requirements, assess the AI Go/No-Go test and additional needs to help sharpen the cognitive requirements and uncover additional considerations for AI projects.

Task: Cognitive Requirements

Description

The goal of this section is to make it clear what and how the cognitive solution will be relevant to the business objectives. How is the business objective being addressed currently? In what ways would noncognitive approaches solve this problem? Would those noncognitive approaches solve this problem equally well? If not, in what ways would machine-based cognitive approaches solve it better? What would the cognitive solution need to do in order to solve the problem better than noncognitive approaches?

Task Artifacts

Noncognitive/heuristic approach: A heuristic approach is a simple, straightforward way of accomplishing something that represents the baseline, noncognitive approach. Some people refer to heuristic as "quick and dirty" while others simply consider it to be the simple, obvious method, or others think of it as the brute-force approach, depending on the situation. For example, if the goal is to develop a system to provide movie recommendations, the noncognitive heuristic approach might be to simply present the top 10 list of most popular movies. In another example, if the solution is to replace human effort to do a similar task, then the heuristic approach might be simply that same human effort.

The goal of a cognitive solution is to do the task better, faster, cheaper, more reliably, or more scalable than the heuristic approach. In this task, identify the heuristic approach if it exists. If not, determine what the heuristic approach could be. In some instances, simply defining a heuristic approach may make the need for a cognitive approach unnecessary as it may not be effective to replace a "good enough" noncognitive solution.

For this task

List the current, noncognitive approaches used to address the current task.
List any additional noncognitive approaches that could be "good enough" to use to meet the requirements.
List the specific disadvantages of these noncognitive approaches in terms of cost, effort, complexity, scalability, time, or other considerations that any cognitive approach would need to improve upon.
If the noncognitive approaches are indeed good enough, provide a firm justification for why a cognitive approach is required.

Why does this project need a cognitive (AI) solution?

What noncognitive (non-AI) alternatives are there to solving the current business problem?
For those alternatives, why are they not feasible for this project?
If noncognitive alternatives are feasible, then why are they not being used for this project?

What are the noncognitive (non-AI) portions of this project that will be used in conjunction with the cognitive components?

Are non-cognitive automation alternatives possible for this iteration? If so, why are they not being used for this project iteration?

Cognitive objectives: Enumerate the specific objectives that the cognitive solution must accomplish. For example, “Predict how many clicks a viewer will make for a specific video, enabling the presentation of videos sorted by predicted total clicks.” The form of these goals should be in a tangible format relevant to cognitive specific activities.

What are the cognitive objectives for this project?

Cognitive outcomes: Detail the specific outcomes or goals using business terminology that the cognitive approach should address. For example, the business goal might be “decrease time and effort required to categorize documents,” and so the cognitive goal might be “classify documents with accuracy of at least 90% at a rate of at least 60 per hour.”

Define the criteria for a successful outcome of the project in technical terms—for example, a certain level of predictive accuracy or a propensity-to-purchase profile with a given degree of “lift.” As with business success criteria, it may be necessary to describe these criteria in subjective terms, in which case the person or persons making the subjective judgment should be identified.

What are the cognitive outcomes and goals for this project?

AI success criteria: Describe the criteria for a successful or useful outcome to the project for an AI project. What would the AI project need to successfully do that a non-AI project would not be able to do? In what ways would the AI system need to be better than a non-AI system? If some people on your team are skeptical about the use of AI, what would you need to prove in your AI system to satisfy their needs?

What would the AI project need to successfully do that a non-AI project would not be able to do? In what ways would the AI system need to be better than a non-AI system?

Task: AI Pattern Identification

Description

CPMAI uses the seven patterns of AI shown in Figure 2 as a way to shortcut and speed up cognitive projects. Each of the seven patterns represents projects that share similar objectives, technology basis, and other aspects that once acknowledged will help fill in the missing blanks as to how any other project in that pattern should run. Identifying the pattern and even other projects that have been developevd in the pattern will prove to be convenient for project teams.

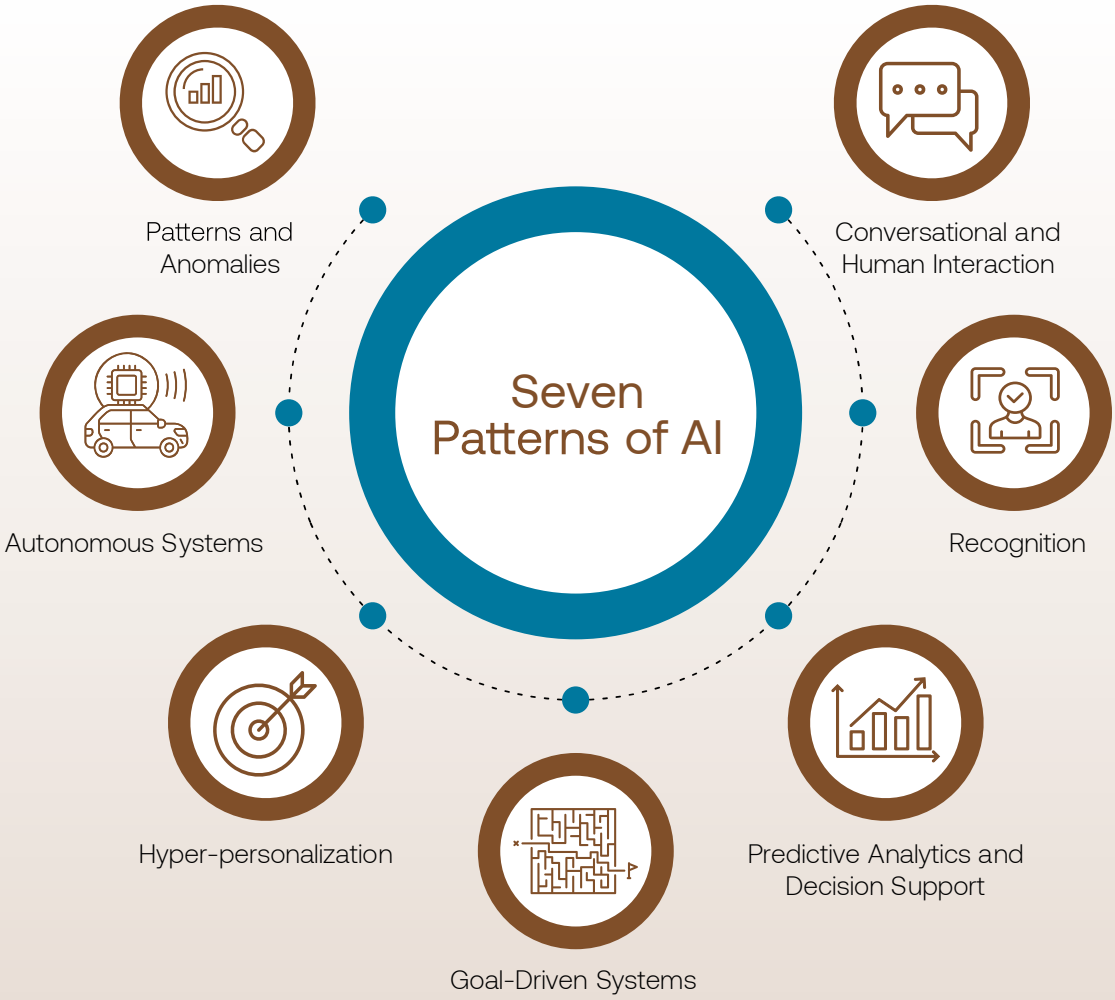


Figure 2. The seven patterns of AI

Task Artifacts

Pattern(s) identification: Which patterns are you implementing in this project? How do you know this is the pattern(s) you are implementing? What other projects have you or others implemented using this pattern that you can learn from or glean any existing artifacts from?

Which pattern(s) of AI are you implementing for this project iteration?

Task Group: Assess Situation

This set of tasks focuses on figuring out where the business is with regard to its current situation and considerations. You can consider these tasks to be the "prerequisites" for the business objectives or constraints under which the business objective should be met.

Task: Resource Requirements

Description

To accomplish this task, the project team should do some detailed fact finding about the resources, constraints, assumptions, and other factors that should be considered in determining the goal and project plan as relevant to the cognitive task at hand. In this task, we expand upon the business objectives to understand what will be available to address those objectives and any constraints that stand in the way of successful project execution.

Task Artifacts

Inventory of Resources

List the resources available to the project, including personnel (business experts, data experts, technical support, data-mining experts), data (fixed extracts; access to live, warehoused, or operational data), computing resources (hardware platforms), and software (data-mining tools, other relevant software).

What are the project iteration schedule requirements or constraints?

List the cognitive-specific technology resources you have available. What infrastructure do you have that is available for use? What cognitive tools do you have that is available for use?

What technology resources do you need for this project?

List the cognitive skills you have available. What expertise and skills are available that you can use for this project? What are the skill gaps you need to address for this project?

What skills do you need for this project iteration?

Task: Schedule Requirements

Description

What are the critical project schedule, timing, and dependency requirements that will impact this iteration?

Task Artifacts

List all requirements of the project, including schedule of completion, comprehensibility and quality of results, and security, as well as legal issues. As part of this output, make sure you are allowed to use the data.

What talent/team resources do you need for this project?

List the constraints on the project. These may be constraints on the availability of resources but may also include technological constraints such as the size of the data set that it is practical to use for modeling.

What are the other project constraints that might impact the ability to deliver this iteration?

Task Group: AI System Performance and Operation

Task: Acceptable Model Performance Values

Description

Cognitive projects differ from traditional application development projects in that they are probabilistic. They provide results in ranges of probability and with some element of error. Project teams need to identify the minimum acceptable performance values they will accept for a particular project. These acceptable limits might start low during initial iterations, while the project proves its value, and may be increased in subsequent iterations. In a machine learning project, using the confusion matrix shown in Figure 3 illustrates how to calculate various measures.

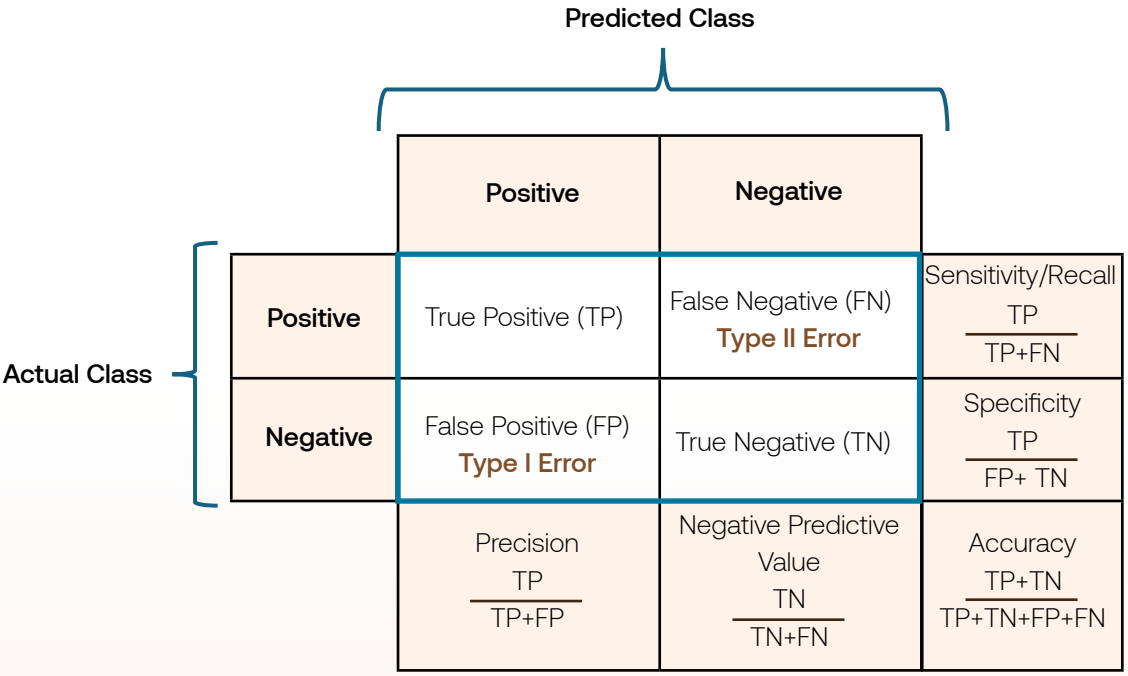


Figure 3. Confusion matrix (for binary classifier)

Task Artifacts

Minimum acceptable accuracy: Accuracy is calculated as the sum of the true positives and true negatives divided by all of the results. In essence, you are calculating how much the model "got right" over all of its guesses. For example, imagine that the model is predicting whether a given email is spam. How many times did the system correctly classify the email as "spam," and how many times did it correctly classify the email as "not spam" across all of its guesses? Obviously, 100% correct would be the best outcome. What is the minimum value for accuracy you will accept? This depends on the cost, complexity, and time involved in handling inaccurate results.

Minimum acceptable precision/positive predictive value: Precision is measured as the number of times the model correctly predicted a positive result over all of the times it guessed a positive result. In the spam classifier example above, this is how many times the system correctly classified the email as "spam" over all of the times it classified a message as "spam."

Minimum acceptable recall/sensitivity: Recall is measured as the number of times the model correctly predicted a positive result over all of the times it should have predicted a positive result. In the spam classifier example, this is how many times the system correctly classified the email as "spam" over all of the times that the messages were actually spam.

Minimum acceptable F1: The F1 is calculated as a combination of precision and recall. The exact formula is F1 = (2 x Precision x Recall)/(Precision + Recall). Think of it as a weighted average of precision and recall. What minimum value of F1 would you accept?

Maximum acceptable false positive rate: Some projects have a high penalty for incorrectly predicted positive results. As such, in situations where being incorrect about a positive result has significant cost or other detriments, you need to indicate what your maximum accepted false positive rate should be. This is calculated as the number of times the model predicts an incorrect positive result divided by the total number of actual negative results. In the spam classifier example, the false positive rate is the percentage of time the model predicted an email was "spam" when it was not spam. The goal is to get this number as close to 0% as possible. What is the maximum value you will accept?

Maximum acceptable false negative rate: As noted above, some projects may have high penalties for incorrectly predicted negative results. The false negative rate is calculated as the number of times the model predicts an incorrect negative result divided by the total number of actual positive results. In the spam classifier example, the false negative rate is the percentage of time the model predicted an email was "not spam" when it was spam. The goal is for this number to be as close to 0% as possible. What is the maximum value you will accept?

What are the desired or required performance metrics for the model?
What sensitivities are there to false positives or negatives in the case of a binary classifier or inaccurate responses in the case of generative AI solutions?

Task: Acceptable KPI Performance Values

Description

Cognitive projects differ from traditional application development projects in that they are probabilistic. They provide results in ranges of probability and with some element of error. Project teams should to identify the minimum acceptable performance values they will accept for a particular project. These acceptable limits may start low during initial iterations, while the project proves its value, and may be increased in subsequent iterations. In a machine learning project, using the confusion matrix shown in Figure 3 illustrates how to calculate various measures.

Task Artifacts

Required Business KPIs.

What are the required business KPIs—such as revenue goals, customer acquisition goals, reduction in errors, desired efficiency, or other metrics—that this project will aim to achieve? These KPIs should be assessable metrics that the system and project can be measured against.

What are the desired or required business KPI performance metrics for this AI project iteration?

Required Technology KPIs

What are the required technology KPIs that would be relevant to the IT team or technical operation—such as response time, server or cloud usage, data system requirements, error rates, or other metrics—that this project will aim to achieve? These KPIs need to be assessable metrics that the system and project can be measured against.

What are the desired or required technology KPI performance metrics for this AI project iteration?

Task Group: Trustworthy AI Requirements

Task: Use of Trustworthy AI Framework

Description

One of the best practices for trustworthy AI projects is the use of a comprehensive, trustworthy AI framework that guides all of the steps of AI projects and builds a cohesive, consistent application of trustworthy AI concepts across all AI projects. If you are using a trustworthy AI framework, you should make that known. Otherwise, if you are not, you should make it clear that the trustworthy AI implementations might be unique and specific to this AI project.

Task Artifacts

Identification of trustworthy AI framework: Identify which, if any, trustworthy AI frameworks or guidance will be used for this AI project.

What, if any, trustworthy AI framework will you be using for this project?
If none, how will you ensure consistent application of trustworthy AI across this project and others?

Task: Required Ethical AI Considerations

Description

Cognitive projects can face ethical challenges including ensuring no harm is done, elimination or avoidance of bias in data and results, and fairness in the model results and usage.

Task Artifacts

Do No Harm
What potential physical, financial, emotional, environmental, or other harms could be caused by this AI project? What approaches will you use to mitigate those potential harms?

What potential physical, financial, emotional, environmental, or other harms could be caused by this project?
What approaches will you use to mitigate those potential harms?

Task: AI Failure Modes

Description

Since cognitive projects are inherently probabilistic and include failure, project teams need to identify how they will handle failure of the model to produce expected results.

Task Artifacts

Failure Contingency
How will you handle situations where the model fails to provide a positive result?
How will you handle situations where the model provides a result with less-than-required accuracy/confidence?
What are acceptable probability ranges to automatically accept model results?
What are probability ranges for results that may require intervention by a backup system or a human?
In what situations can you resort to a heuristic approach if the model fails to provide results with high confidence?
How will you handle negative results in situations of sensitivity to false negatives?
How will you handle positive results in situations of sensitivity to false positives?

How will you know when the AI project is failing to provide adequate results?
How will you handle AI system failures for this iteration?

Risks and Contingencies

List the risks or events that may delay the project or cause it to fail.
List the corresponding contingency plans—what action will be taken if these risks or events take place.

What do you see as the most significant risks for this project that could lead to project failure?

Human-in-the-Loop

In what ways will a human be in the loop of machine learning model usage?
When will the system be put on hold or otherwise halted if the cognitive project is failing to provide valuable results?
What monitoring will you put into place to ensure the system does not provide undesirable results?

How will you maintain a human in the loop or otherwise involved in the AI project operation?

Bias Identification

How will you minimize exposure to informational bias?
What approaches will you take to verify that there is not any human-induced bias in your data?
Can you produce any metrics or measures to gauge if your results are skewed by human bias in data?

How will you identify and minimize exposure to informational bias?

Task: Required Compliance With Regulations and Laws

Description

Your AI project and the data it uses may be subject to compliance with laws and regulations based on your location, industry, and other factors. You need to identify what regulations and laws you may be required to comply with, so your AI project can be completed without legal or liability issues.

Task Artifacts

Identify Laws and Regulations for Compliance

For this AI project iteration, what laws, regulations, or other compliance may be required? If you do not know, how will you find out?

Task: Required AI Transparency Considerations

Description

Cognitive projects can face challenges in providing transparent results. It is important to note what the transparency requirements are for this project in terms of explaining the algorithm and any other transparency measures.

Task Artifacts

Data Transparency
Is it important and possible to provide visibility into the sources of data and even access to the data for anyone who inquires?
What steps and processes will you put into place to provide transparency into data sources?

What transparency are you going to provide to others about the source(s) of the data used in this AI project?

Methods of Data Selection Transparency

Is it important to provide visibility into how you decide which data to include in the model and the methods by which you selected that data or excluded other data?
What steps and processes will you put into place to provide transparency into methods of data selection, data exclusion, data filtering, or data transformation?

What transparency are you going to provide to others about the methods you use to select and filter the data you are using for your AI project?

Task: Required AI Explainability Considerations

Description

Certain legal, compliance, and risk considerations may require that the AI system used for decision-making or other uses provides some level of explainability for audit, root cause analysis, or other purposes. If you have such requirements, you should know about them as they will impact algorithm selection choice and other factors.

Task Artifacts

Model explainability: Is it important for the results of the model to be explainable? If so, to what degree? Will this necessitate not using black box algorithms, that are opaque and not easily understandable, even to the developers who created them?

What are the requirements for explainable algorithms for this AI project?

Task Group: AI Go/No-Go

The AI Go/No-Go test, seen in Table 2, is a set of nine factors upon which a determination of whether or not an AI project is possible depends. Three of these factors focus on business feasibility, three on data feasibility, and three on execution feasibility.

		Risk	Cautious	Go
Business Feasibility	Is there a clear problem definition?			
	Is the organization willing to invest and change?			
	Is there sufficient ROI or impact?			
Data feasibility	Do you have the required data that measures what you care about?			
	Is there sufficient quantity of data needed to train systems, and do you have access to that data?			
	Is the data of sufficient quality?			
Technology/ Execution feasibility	Do you have the required technology and skills?			
	Can you execute the model as required in a timely manner?			
	Does it make sense to use the model where you plan to use it?			

Table 2: AI Go/No-Go factors

Task: Business Feasibility

Description

The business feasibility factors aim to determine whether or not the key business conditions exist for successful implementation and deployment of an AI project.

Task Artifacts

Business feasibility: Answer the following questions, with specificity such that the answers must be "yes" to continue the cognitive project:

Do you have a clear description of the problem with regard to the business objectives? If so, mark this a "GO."
If not, what additional definition is required? Mark this as a "No-Go."

GO	MAYBE	NO-GO
☐	☐	☐

Is the customer/business owner/product owner willing to implement or put in production the cognitive solution that your team will be producing? Can you get a verbal or written commitment to put your solution into production without any obstacles? If there are obstacles for using the project team's output, what are those obstacles, and can they be achieved within the iteration?

Is the customer/business owner/product owner willing to implement/put in production the cognitive solution that your team will be producing? If so, mark this as "Go."
If not, what obstacles are in the way of implementing your AI project? Mark this as a "No-Go."

GO	MAYBE	NO-GO
☐	☐	☐

Does the cognitive solution provide enough ROI or impact to the business, assuming it can be produced? And will the solution be better or more impactful than the current, noncognitive, heuristic approach?

Does the cognitive solution provide enough ROI or impact? If so, mark this as "Go."
If not, what do you need to modify in your project to provide a positive return? Mark this as a "No-Go."

☐ GO	☐ MAYBE	☐ NO-GO

Task: Data Feasibility

Description

The data feasibility factors aim to determine whether or not the key data conditions exist for successful implementation and deployment of an AI project.

Task Artifacts

Data feasibility: Answer the following questions, with specificity such that the answers must be "yes" to continue the cognitive project:
Is the data that is required to create the cognitive model available, and does it actually measure what you need to make the cognitive project a reality?

Is the data required to create the cognitive model available, and does it actually measure what you need? If so, mark this as "Go."
If not, what data do you need? Mark this as a "No-Go."

GO	MAYBE	NO-GO
☐	☐	☐

Do you have enough data, and do you have the means to access it?

Do you have access to the data you need? If so, mark this as "Go.
If not, what do you need for access to the data? Mark this as a "No-Go."

GO	MAYBE	NO-GO
☐	☐	☐

Does the data have a sufficient level of quality to be useful? We will revisit this Go/No-Go during CPMAI Phase II, so make your best guess based on what you know now.

Does the data have a sufficient level of quality to be useful? If so, mark this as "Go."
If not, mark as a "No-Go."
We will revisit this Go/No-Go during CPMAI Phase II, so make your best guess based on what you know now.

GO	MAYBE	NO-GO
☐	☐	☐

Task: Execution Feasibility

Description

The execution feasibility factors aim to determine whether or not the AI project can be executed as required for successful implementation and deployment of an AI project.

Task Artifacts

Execution Feasibility

Do you have access to the technology you need to build the cognitive solution? Do you have the required expertise and skills to use it? If there is a cost to use that technology, do you have the funds and approvals to use that technology? If approvals are needed, do you have a clear path to obtain such approvals?

Do you have access to the technology and expertise you need for this iteration? If so, mark this as "Go." If not, what technology or expertise do you need? Mark as a "No-Go."

GO	MAYBE	NO-GO

Is it feasible to build a model that executes in sufficient time to address the specific problem? Will the solution require a substantial investment in technology or infrastructure to meet such performance goals? Will the resultant model fit within the constraints in which the system will be used? If you do not know the answers to these questions, can you obtain the answers quickly before moving to the next phase?

Is it feasible to implement the model where and how you want to? If so, mark this as "Go." If not, how can you obtain the answers needed to make it feasible? Mark as a "No-Go."

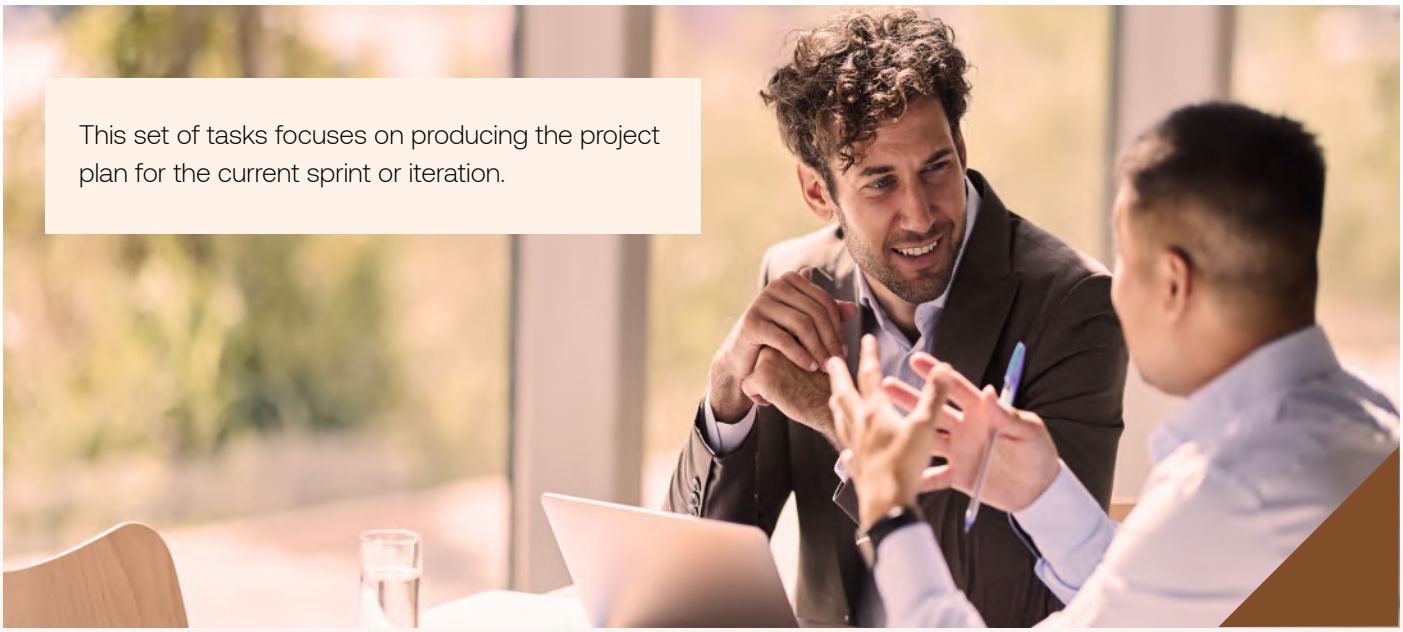
GO	MAYBE	NO-GO

Does it make sense to put the model where it should be operationalized in order to function? Will the model actually work where you intend it to be operationalized? Have you gathered the required understanding to know the cost and complexity to operationalize the model where it will be intended to go? If not, do you have a means to obtain the answers to these questions quickly before moving to the next phase?

Does it make technical, operational, business, and financial sense to implement the model in the way and in the location you want to? If so, mark this as "Go."
If not, how can you get the answers needed to make it feasible? Mark as a "No-Go."

☐ GO	☐ MAYBE	☐ NO-GO

Generic Task Group: Produce Project Plan



This set of tasks focuses on producing the project plan for the current sprint or iteration.

Task: Determine Business Objectives

Description

Describe the intended plan for achieving the cognitive project goals and thereby achieving the business goals. The plan should specify the steps to be performed during the rest of the project, including the initial selection of tools and techniques.

Task Artifacts

Project Plan

List the stages to be executed in the project, together with their duration, resources required, inputs, outputs, and dependencies. Indicate if iterations will be required between phases of the project—for example, iterations of the modeling and evaluation phases depending on what is learned as part of other phases.
List any dependencies between time schedule and risks. Mark results of these analyses explicitly in the project plan, ideally with actions and recommendations.
Explain the process by which each step in the project plan will be evaluated for progress and achievements and how this will map to the sprint or iteration and communication plan as developed as part of the agile team charter. Make sure the agile sprint plan includes any assessments that will be developed by the project team as part of the project plan.

Initial Tools and Technique Selection

From the survey of available tools made in the assessment task defined previously, which tools and techniques will you be using?
Will purchases be required to obtain required tools?
Will any pre-project evaluation be required to determine which tools to use?
How long is the procurement cycle required to acquire required tools?
As an example, you select a technology that supports various methods that would be used at different stages of the project. It is important to assess tools and techniques early in the process since the selection of tools and techniques may influence the entire project.

CPMAI Phase II: Data Understanding

Cognitive projects are dependent on having access to the right sources of clean data in the right formats. The second phase in a CPMAI project is the Data Understanding phase, which is focused on data needs identification, initial data collection, data requirements, data quality identification, insights into data, and potential interesting aspects of the data worth further investigation.

Task Group: Collect Initial Data

This set of tasks focuses on collecting data, to understand what data is available, its various formats, and levels of quality.

Task: Collect Initial Data

Description

This task requires the project team to acquire the necessary data that was identified earlier in the project resources. If necessary, this means data loading and some initial data preparation to understand the data. If you acquire multiple data sources, integration is an additional issue, either here or in the later Data Preparation phase.

Task Artifacts

Initial Data Collection Report

- List the data set(s) required, together with their locations, the methods used to acquire them, and any problems encountered.
- Record problems encountered and any resolutions achieved. This will aid with future replication of this project or with the execution of similar future projects.

Detail the list of data and locations of that data you will need for this iteration of the AI project. If you are encountering any issues with locating or accessing data, document resolution to these issues.

Task: Describe Data

Description

The goal is to examine the data—and the properties of that data as required for the project—and report on the results.

Task Artifacts

Data Description Report

- Describe the data that has been acquired, including the format of the data, the quantity of data (for example, the number of records and fields in each table), the identities of the fields, and any other surface features that have been discovered.
- Evaluate whether the data acquired satisfies the relevant requirements.
- Is the data properly structured if you need structured data as per pattern requirements?
- Is unstructured data available as per pattern requirements?

Document the nature of the data you need. What structure is it? Does it have the elements that you need for your AI project iteration?
If not, how will you resolve issues of mismatch with the data you need and what you have?

Task: Explore Data and Cognitive Data Requirements

Description

This task focuses on inspecting the data to assess its characteristics, methods of querying and visualization, distribution of key attributes (for example, the target attribute of a prediction task), relationships between pairs or small numbers of attributes, results of simple aggregations, properties of significant subpopulations, and simple statistical analyses. These analyses may directly address the project goals; they may also contribute to or refine the data description and quality reports, and feed into the transformation and other data preparation steps needed for further analysis.

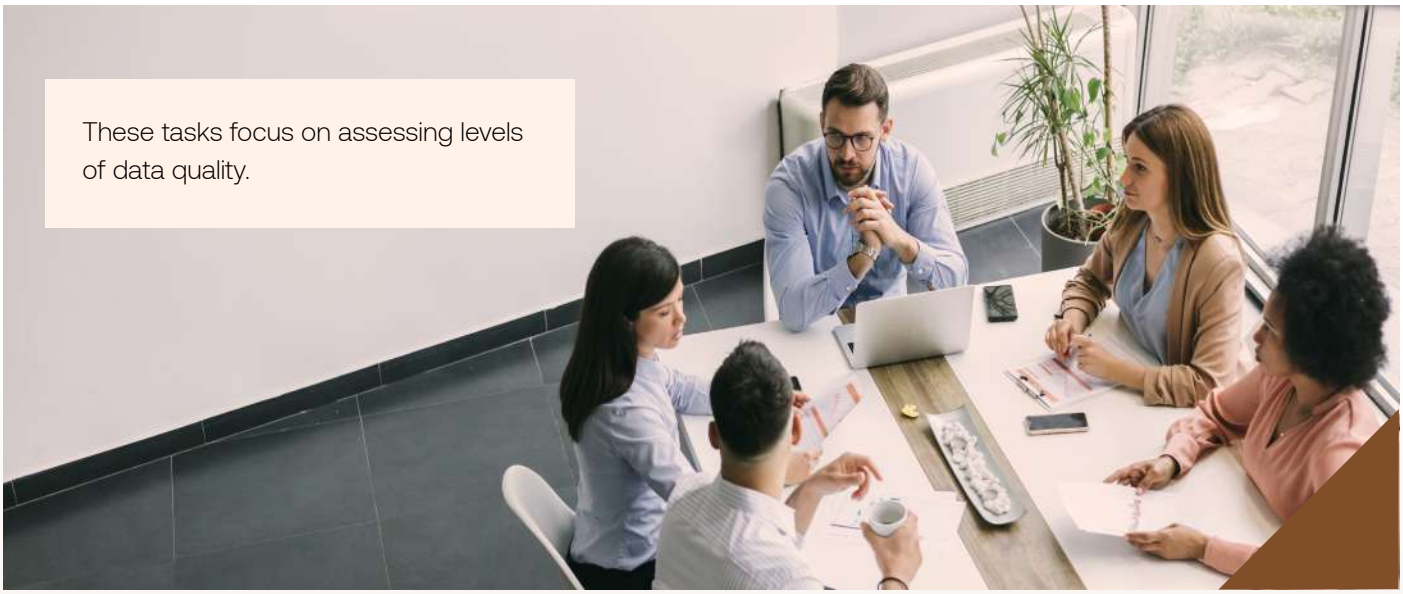
Task Artifacts

Data Exploration Report

- First findings or initial hypothesis and their impact on the remainder of the project.
- If appropriate, include graphs and plots to indicate data characteristics that suggest further examination of interesting data subsets.
- If the pattern requires unstructured data, do you have sufficient quantities of that data?
- Do you have access to sufficient data to be used for training purposes?
- Do you need to augment data from third-party sources for model training purposes?

Have you inspected and selected some of the data to ensure it meets your needs? Detail what you discovered. Is the data a sufficient quantity for your AI project iteration needs? If not, how will you resolve the lack of data?

Task Group: Data Quality



These tasks focus on assessing levels of data quality.

Task: Verify Data Quality

Description

This task focuses on examining the quality of the data and addressing questions such as: Is the data complete (does it cover all the cases required)? Is it correct, or does it contain errors? If there are errors, how common are they? Are there missing values in the data? If so, how are they represented, where do they occur, and how common are they?

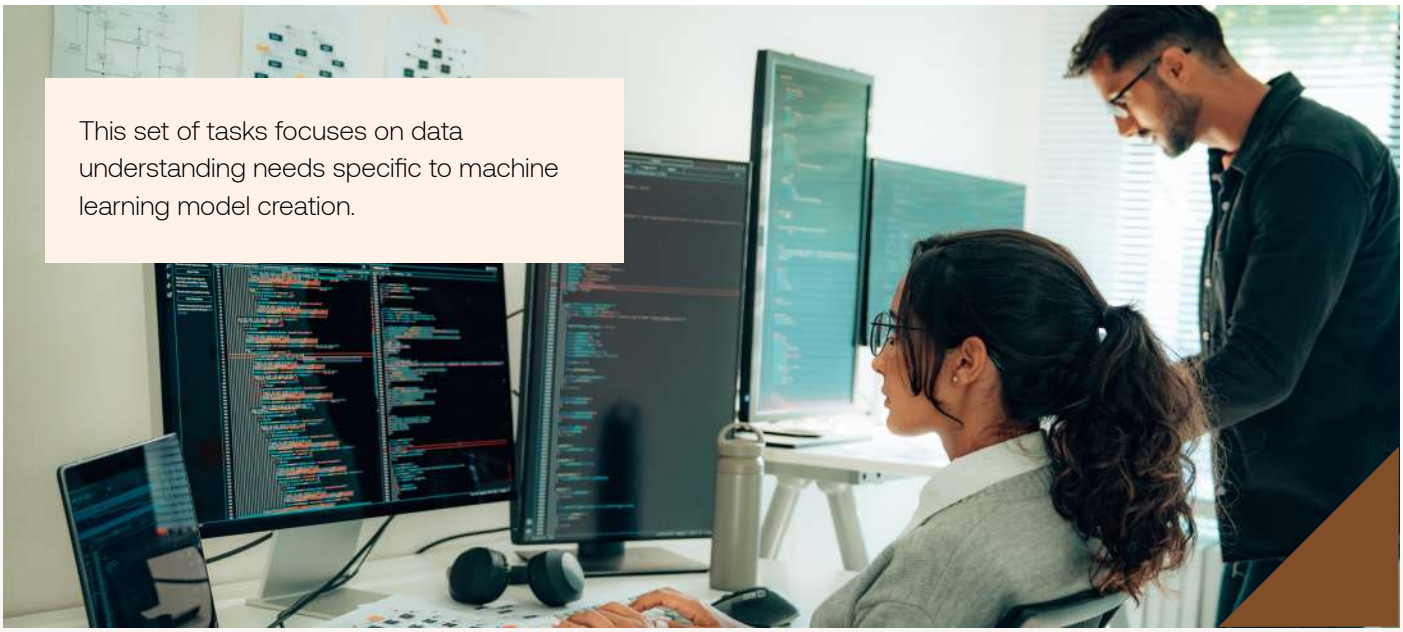
Task Artifacts

Data Quality Report

- List the results of the data quality verification.
- What errors are present in the data?
- What missing values are in the data?
- Is the data complete? If not, how is it not complete?
- Is the data in the required format necessary?
- What additional features or attributes are missing in the data?
- What transformations are required to change the data into the right formats?
- What additional data augmentation is necessary?
- If quality problems exist, list possible solutions. Solutions to data-quality problems generally depend heavily on both data and business knowledge.

What is the current quality of the data you located for your AI project?
What needs do you have for data preparation, augmentation, enhancement, and transformation?

Generic Task Group: Machine Learning Model Data Requirements



This set of tasks focuses on data understanding needs specific to machine learning model creation.

Task: Training and Test Data Requirements

Description

Supervised learning and other forms of machine learning often require well-labeled training data sets upon which they must learn, and these training data sets are then further split into test and validation data sets that can be used to validate the hyperparameter settings and test the model for overall performance. In this task, the project team will specify details of the training, validation, and test data sets.

Task Artifacts

Training Data Set Requirements

- Determine the minimum amount of data necessary for training and its format.
- Determine the source and quality of training data.
- What transformation or multiplication activities can be done to increase training data volume while maintaining quality?
- What data-labeling activities should occur for training data?
- What third-party sources of relevant training data are available?
- What additional augmentation, transformation, and modification activities should be done to change internally sourced or third-party sourced training data into a form that can be used for training?
- Generate a report on the above determinations.

Validation Data Set Requirements

- Determine how training data will be split with validation data.
- Determine if and how cross-validation will be used and the value for “k” if using k-fold cross- validation.
- Determine how to interpret validation results.
- Determine how hyperparameters will be optimized based on validation results.
- Determine frequency and iterations of retraining and to what extent you will be optimizing based on validation results.

- Generate a report on the above determinations.

Test Data Set

- Determine how the test data will be split from existing training data.
- Determine the method of test data set selection.
- Determine how test results will impact the decision to use the model.
- Determine what results will require retraining of the model.
- Generate a report on the above determinations.

What additional, specific needs do you have for training data for your AI project? How will you make sure you have sufficient training, test, and validation data?

Task: Edge Model Data Needs

Description

In some instances, the model produced will be used “at the edge” on devices or other systems that will have limited computing or data capability and/or limited connectivity to online resources. In this task, project teams will define unique requirements for edge models that might constrain the way in which models are developed and used.

Task Artifacts

Edge Constraints

- List requirements to use the model on edge systems.
- List constraints on the model or model usage including data constraints, performance constraints, requirements for model result response time, constraints on external data access, and other model operationalization constraints.
- List requirements for data capture on the edge device.
- Determine methods for model versioning and update, methods for monitoring edge model performance, and other operationalization requirements as determined relevant by the project team.

Do you have special needs for data to or from edge devices? If so, detail those needs here.

Task: Pretrained and Third-Party Model Usage

Description

In many cases, project teams will decide to use an existing model already developed by another team or organization, or developed by the same team in a previous project iteration or in another project. In these situations, the team should evaluate the appropriateness of the third-party or pre-trained model to ensure it is still applicable to the current iteration and what revisions, fine-tuning, or changes need to be made so that the model provides required results.

Task Artifacts

Selection of Pretrained/Third-Party Models

- Determine if the team will use a model that has already been trained in a previous iteration by another team or a third-party organization.
- If the model was produced by the current project team in a previous iteration, are the assumptions for that model still the same or have any changed?
- If the model was produced internally by another team, gather information to understand the source of data for that model, how it was trained, and performance parameters
- If the model was produced by a third-party organization, gather information about costs to use the model, origin of data and methods of training, performance measures, and other factors that will determine appropriateness of the model.
- Ideally, run the pretrained model through tests using current iteration data to verify acceptable performance.
- Determine the method for monitoring third-party model usage cost and performance measures over time.
- Document decisions and information gathered.

Needs for Fine-Tuning/Transfer Learning:

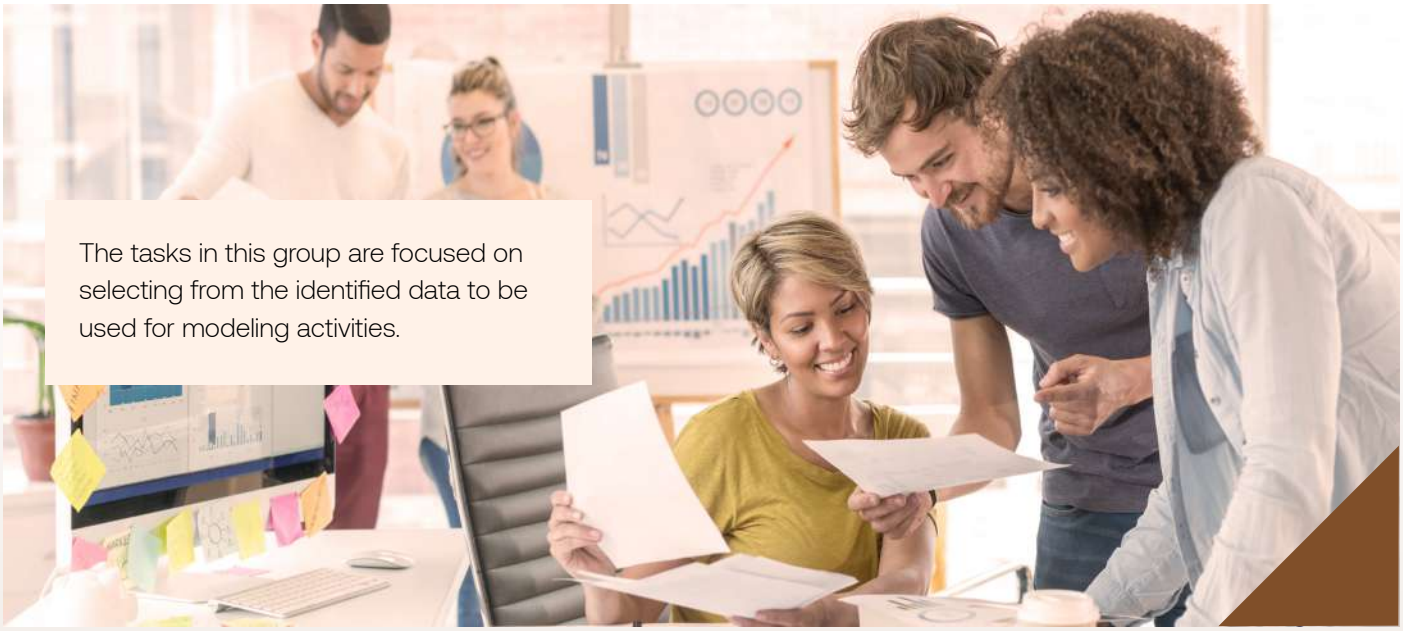
- If pretrained/third-party models cannot be used as-is, determine how the models should be modified to make them useful.
- If transfer learning will be applied to fine-tune models, determine how the transfer learning will be used and decisions made along with training requirements.
- If third-party models are used that cannot be retrained, determine if the model will be used in combination with others in an ensemble or other methods to guarantee the appropriateness of third-party model usage.
- Document fine-tuning and other augmentation decisions.

Can you make use of any pretrained models, models from third parties, or foundation models? If so, detail those models and where and how you will access them.
If you need to extend the models through transfer learning or fine-tuning, detail those needs here.

CPMAI Phase III: Data Preparation

The third phase of CPMAI is focused on directing the required data into a form and shape that can be used for modeling purposes. The output of this phase is the data set that will be used for modeling activities.

Generic Task Group: Data Selection



The tasks in this group are focused on selecting from the identified data to be used for modeling activities.

Task: Select Data

Description

Decide on the data to be used for analysis. Criteria include relevance to the data-mining goals, quality, and technical constraints such as limits on data volume or data types. Note that data selection covers selection of attributes (columns) and selection of records (rows) in a table.

Task Artifacts

Description of Data Selected

- List the sources of data to be selected and queries or selections being made for data.

Rationale for Inclusion/Exclusion

- List the data to be included or excluded and the reasons for these decisions.

Select the data that is needed for the project and provide some documentation of the data selection method. If data sources or data within a data source was excluded, detail that exclusion for future reference.

Generic Task Group: Data Cleansing and Enhancement

The tasks in this group are focused on directing data into a usable format for modeling activities.

Task: Clean Data

Description

The goal of this task is to bring the quality of the data selected to the level required for modeling. This may involve selection of clean subsets of the data, insertion of suitable defaults, estimation of missing data by modeling, and formatting or cleansing operations.

Data pipelines:

- Create a reusable data pipeline to collect, ingest, and prepare data for training purpose ("training data pipeline").
- Create a reusable data pipeline that would be similarly used for real-world inference data.

Rationale for Inclusion/Exclusion: Describe what decisions and actions were taken to address the data quality problems reported during the Verify Data Quality task of the Data Understanding phase.

- Transformations of the data for cleaning purposes and the possible impact on the analysis results should be considered.
- Operations are taken to resolve data issues including de-duplication, fixing erroneous data, identification and replacement of missing data, numeric transformations, textual transformations, or other formulaic transformations applied.

Reordering of attributes: Some tools have requirements on the order of the attributes, such as the first field being a unique identifier for each record or the last field being the outcome field the model is to predict.

Reordering of records: It may be important to change the order of the records in the data set. Perhaps the modeling tool requires that the records be sorted according to the value of the outcome attribute. Commonly, records of the data set are initially ordered in some way, but the modeling algorithm requires them to be in a fairly random order. For example, when using neural networks, it is best for the records to be presented in a random order, although some tools handle this automatically without explicit user intervention.

De-duplication: Some modeling operations may require the removal of duplicate rows or records. List steps taken to identify and eliminate duplicates if necessary.

Erroneous fields: Data may be placed into incorrect fields. List steps taken to identify incorrect fields for data and methods used to fix erroneous placement.

Syntax changes: Additionally, there are purely syntactic changes made to satisfy the requirements of the specific modeling tool. Examples: removing commas from within text fields in comma-delimited data files, trimming all values to a maximum of 32 characters.

Normalization: It might be required to change the value of numeric fields to be within a predefined range for modeling operations. For example, instead of having house pricing ranges of arbitrary values, these can be placed on a scale from 0 to 1 with intermediate values calculated from 0 and the maximum value of the data field.

De-noising: Some data may have erroneous or spurious values due to the way that the data is collected. Formatting may involve some form of preprocessing to remove spurious data or noise from results that will then be fed into modeling operations.

Anonymization: In certain instances, data from various sources may require personally identifiable information (PII) removed prior to being used in models. This may be due to regulatory, privacy, security, or other reasons. Any needs for anonymization should be identified as well as methods used for anonymization and verification that such anonymization will not negatively impact performance of resulting models.

Perform data-cleansing and preparation operations. Detail the methods used for preparation. Document the data pipeline used, from data collection and ingestion through data preparation.

Task: Enhance and Augment Data

Description

This task includes constructive data preparation operations such as the production of derived attributes or entire new records, or transformed values for existing attributes. These methods combine information from multiple tables or records to create new records or values.

Task Artifacts

Derived Attributes

- Derived attributes are new ones that are constructed from one or more existing attributes in the same record. Example: area = length × width.

Generated Records

- Describe the creation of completely new records. Example: Create records for customers who made no purchase during the past year. There was no reason to have such records in the raw data, but for modeling purposes it may make sense to explicitly represent the fact that certain customers made zero purchases.

Enhanced Data

- **Enhancement from additional sources:** Merging data refers to joining together two or more data sources that have different information about the same information resources.

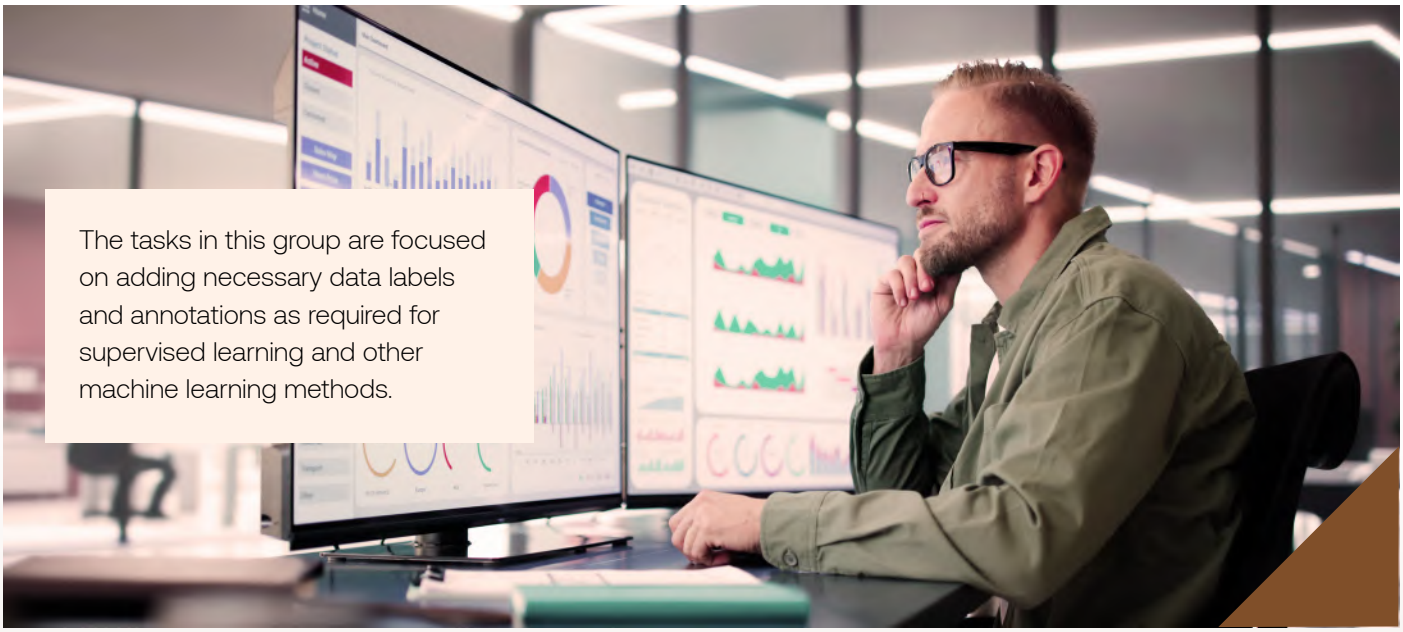
Example: A retail chain has one table with information about each store's general characteristics (e.g., floor space, type of mall), another table with summarized sales data (e.g., profit, percent change in sales from previous year), and another with information about the demographics of the surrounding area. Each of these tables contains one record for each store. These tables can be merged together into a new table with one record for each store, combining fields from the source tables.

- **Aggregations of sources:** Merged data also covers aggregations. Aggregation refers to operations in which new values are computed by summarizing information from multiple records and/or tables. For example, converting a table of customer purchases where there is one record for each purchase into a new table where there is one record for each customer, with fields such as number of purchases, average purchase amount, percent of orders charged to credit card, percent of items under promotion, etc.

- **Enhancement from third-party sources:** In some instances, data may need to be enhanced with third-party data. If this is deemed necessary, identify such third-party sources, any costs or complexity involved in sourcing data, methods for data augmentation, and means for verifying accuracy of third-party data.

Perform data augmentation and enhancement operations. Detail the methods used for augmentation. Document additions or modifications to the data pipeline for augmentation.

Generic Task Group: Data Labeling



The tasks in this group are focused on adding necessary data labels and annotations as required for supervised learning and other machine learning methods.

Task: Label Data

Description

This task involves identifying methods for data labeling and engaging in data-labeling efforts.

Task Artifacts

Identification of Labeling Needs

- List specific requirements for data labeling and annotation.

Identification of Labeling Method

- Identify methods to be used for data labeling, including the use of internal labor, third-party labor, preexisting labeled data, or mixed labor modes.
- Identify costs for data labeling.
- Identify method for verifying quality of data labeling.

Data Labeling

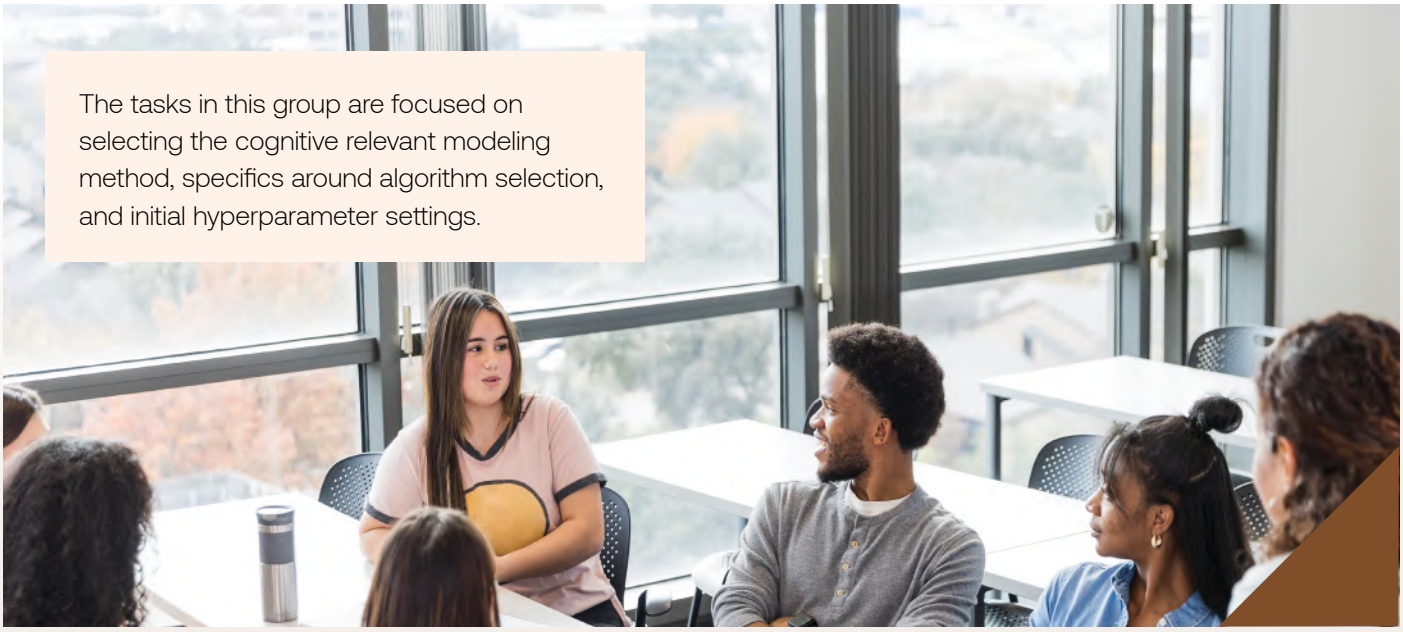
- Conduct necessary data-labeling activity and report on data-labeling outcomes.

Perform data labeling as necessary for data. Detail method and approach used for data labeling and how the costs will scale as data-labeling needs increase.
Document additions or modifications to the data pipeline for data labeling.

CPMAI Phase IV: Model Development

In the fourth phase of CPMAI, the project team creates and iterates on the model based on the identified and prepared data to provide the results as required in the business understanding. The output of this phase is a model that can then be evaluated.

Generic Task Group: Select Modeling Technique



The tasks in this group are focused on selecting the cognitive relevant modeling method, specifics around algorithm selection, and initial hyperparameter settings.

Task: Select Cognitive-Relevant Algorithm/Modeling Technique

Description

Using the data gathered and prepared in previous phases, the project team then selects the machine learning algorithm and/or cognitive-relevant modeling technique. This selection should be related to the tools identified in the Business Understanding phase but independent of that selection. For example, if it is determined that an unsupervised clustering algorithm such as K-means will be used, the tool should support such needs; if not, then the tool choice should be changed, and the project should iterate on that portion of the project. This task refers to the specific modeling technique. If multiple techniques are applied, perform this task separately for each technique. In many instances, the choice of AI pattern in Phase I will guide the decision on algorithm selection and modeling technique.

Task Artifacts

Algorithm Selection

- Document the actual algorithm or modeling technique that is to be used.
- Identify supervised, unsupervised, and/or reinforcement learning approaches to be used.
- Verify that selected tools support algorithm decision.

Initial Parameter/Hyperparameter Settings

- Document decisions made for initial hyperparameter settings and the setting values.

Modeling Assumptions

- Many modeling techniques make specific assumptions about the data—for example, that all attributes have uniform distributions, no missing values are allowed, class attribute must be symbolic, etc. Record any such assumptions made.

Select the appropriate algorithm and approach to be used for model development.
If a generative AI, foundation model, or pretrained model will be used, see a later workbook task.

Task: Ensemble Methods

Description

This task focuses on the use of multiple models combined in the form of an ensemble to improve overall performance in certain instances. Project teams should determine if an ensemble is useful or feasible and how it should be set up if it is determined to be useful. Some machine learning algorithms, such as random forests, are inherently ensemble methods.

Task Artifacts

Ensemble Determination

- Determine and document if an ensemble method or approach is necessary or helpful to improve overall performance.

Ensemble Method and Configuration

- Document ensemble method (i.e., bagging, boosting, gradient, averaging) and any configuration necessary for proper operation.

If an ensemble of models will be developed for this iteration, detail the configuration of that ensemble.

Task: Usage of AutoML

Description

In many instances, model-building steps can be significantly accelerated and facilitated through the use of automatic machine learning (AutoML) tools. If a decision has been made to use such tools, the specifics around AutoML tool usage should be defined.

Task Artifacts

AutoML Usage

- Document how AutoML tools will be used for model creation and any requirements for the tool to operate, as well as how the output of such tools will be used.
- Verify that models created by AutoML tools will be usable as part of the I/O flow determined in the AI Go/No-Go task as part of Phase I.

If you will use AutoML tools to accelerate model development, detail the tool(s) used and how they will be applied for this AI project iteration.

Task: Fine-Tuning/Retraining of Pretrained Models

Description

If a pretrained model or foundation model was determined to be useful for the project, and if it requires additional fine-tuning or retraining, those steps should be identified and performed for the model-building task.

Task Artifacts

Model Fine-Tuning

- Determine and document what approach will be used to fine-tune or retrain pretrained models.
- Verify that selected tools can fine-tune existing models.

If a pretrained model or foundation model will be used, detail which model(s) will be used and the method, if any, used to fine-tune the model for your specific AI project iteration.

Task: Usage of Generative AI

Description

Using generative AI can accelerate your AI project. The primary method for using generative AI is through prompt engineering approaches or the use of additional tools to chain together generative AI/Large Language Model (LLM) interactions. If this is the approach you are going to use for this iteration, detail the method used for prompt generation and changing of LLM or generative AI interactions.

Task Artifacts

Detail generative AI used:

- If the generative AI solution is a third-party API, detail how the API will be accessed and the incremental cost(s) and limitations of that API.

If needed, detail which generative AI approach will be used; if hosted through an API, detail the costs and limitations of the API.

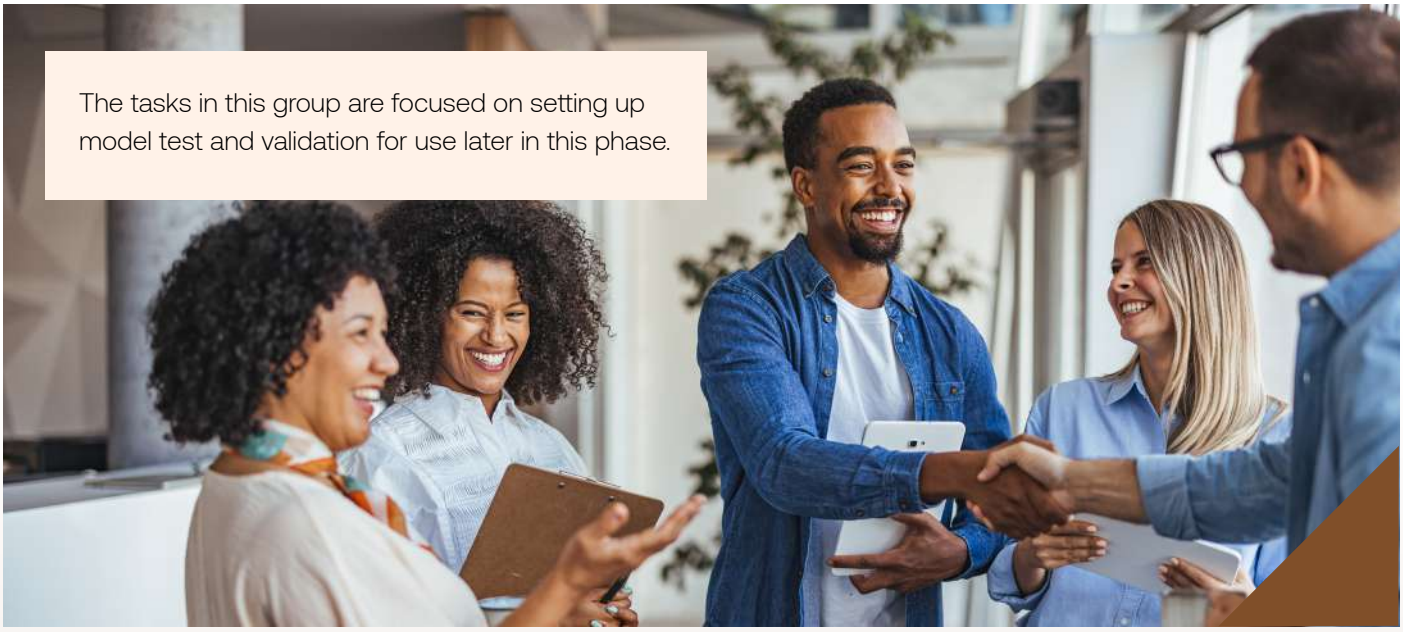
Detail prompt engineering approach: Generative AI solutions will require a prompt engineering approach, so detail that here.

Determine the approach for prompt engineering to be used with the generative AI solution.

Chaining method/approach: If you are going to use a generative AI or LLM approach in your solution that requires chaining to access additional resources or pre- or post-processing, detail the chaining method and approach you will be using.

Determine the approach used, if any, to chain LLM or generative AI results to be used with additional resources or data sources of pre- or post-processing of generative AI/LLM results.

Generic Task Group: Model Test and Validation Design



The tasks in this group are focused on setting up model test and validation for use later in this phase.

Task: Generate Model Test Design

Description

Before we build a model, we need to generate a procedure or mechanism to test the model's quality and validity. For example, in supervised data-mining tasks such as classification, it is common to use error rates as quality measures for data-mining models. Therefore, we typically separate the data set into train and test sets, build the model on the train set, and estimate its quality on the separate test set.

Task Artifacts

Test and Validation Design:

- Describe the intended plan for training, testing, and validating the models. A primary component of the plan is determining how to divide the available data set into training, test, and validation data sets.
- Documentation of k-fold cross-validation settings if that method is used.
- Verify that tools selected support methods of testing and validation.

Determine the approach used to validate the model and ensure that it does not overfit or underfit data and achieves desired machine learning objectives.

Generic Task Group: Model Training/Model Building

The tasks in this group focus on building the model itself, usually through training if a supervised model, or through other means for unsupervised, reinforcement, or other cognitive-relevant approaches.

Task: Model Training/Model Building

Description

Conduct model training or model building as determined in prior tasks, using the prepared data set to create one or more models and/or model ensembles.

Task Artifacts

Models

- These are the actual models produced by the modeling tool, not a report.

Model Descriptions

- Describe the resulting models. Report on the interpretation of the models and document any difficulties encountered with their meanings.

Perform model training/model development, fine-tuning, or prompt engineering activities. Document or detail any results or outputs that can be shared for future phases or iterations.

Generic Task Group: Hyperparameter Optimization

Tasks in this group are focused on maximizing fit and other model-related measures.

Optimized Hyperparameter Settings

- With any modeling tool, there are often a large number of user-configurable parameters that can be adjusted. In a previous task, initial hyperparameter settings were chosen. In this task, we list the final, optimized settings. List the hyperparameters and their chosen values, along with the rationale or method for the choice of the settings.

Revised Hyperparameter Settings

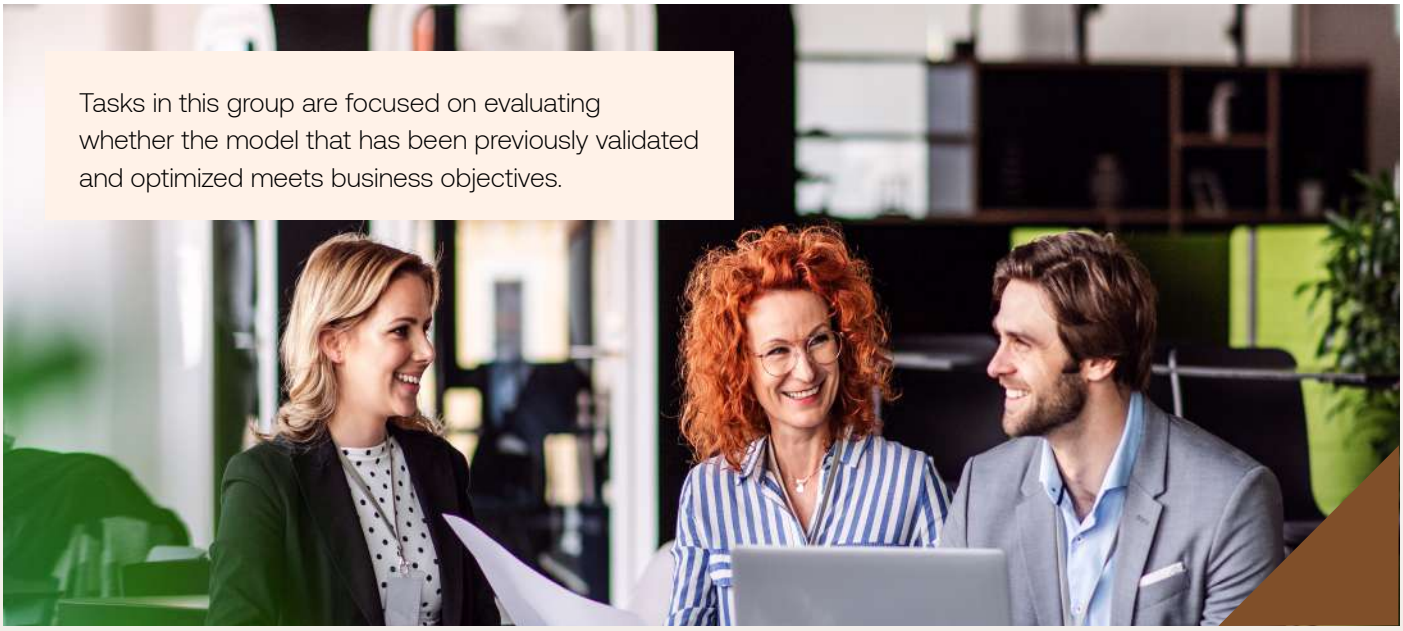
- According to the model assessment, revise hyperparameter settings and tune them for the next run in the Build Model task. Iterate model building and assessment until you strongly believe that you have found the best model(s).
- Document all such revisions and assessments.

Document measurements of model fit against training, test, and validation data.
Determine methods for hyperparameter optimization and perform optimization, documenting results.

CPMAI Phase V: Model Evaluation

In the previous phase, the Model Assessment task dealt with factors such as the accuracy and generality of the model. This phase focuses on assessing the degree to which the model meets the Phase I objectives and seeks to determine if there is some project-relevant reason this model is deficient. This may entail testing the model in a production environment or staging environment against a real application to see how it performs.

Task Group: Evaluate Model Results



Tasks in this group are focused on evaluating whether the model that has been previously validated and optimized meets business objectives.

Task: Model Performance Results

Description

Project managers should validate and interpret the resulting models based on a variety of factors including fit determination (overfit versus underfit), accuracy and performance measures, other factors that would impact changes to hyperparameter settings, changes to algorithm selection, and other factors that may require retraining or other iterations to model building. Please note that this task only considers models, whereas the evaluation phase also considers all other results that were produced in the course of the project.

Task Artifacts

Model Assessment/Validation

- Provide a model validation report from tooling based on selection and method of model validation.
- Rank resulting models and assess according to the evaluation criteria.
- Summarize results of this task, list qualities of generated models (e.g., in terms of accuracy, precision, performance, and other facts), and rank their quality in relation to each other.

Model Measurements

- Provide any measurement output from modeling tools such as learning curves, receiver operating characteristic (ROC) curves, precision versus recall, lift/gain charts, and other training and validation measures to assist with assessing model performance.

Evaluate the model and produce evaluation measures including, as relevant, confusion matrices, ROC curves, and other assessments of model performance.

Task: Key Performance Indicator Measurement

Description

This task focuses on measuring and evaluating the model against Phase I objectives.

Task Artifacts

Assessment of Model With Respect to Business Success Criteria

- Summarize assessment results in terms of business success criteria, including a final statement regarding whether the model and project overall meet the initial business objectives.

Measure model performance against the business KPIs detailed in Phase I.
If the model does not perform adequately, detail the steps needed to improve the KPI performance of the model in this iteration.

Technology KPI Evaluation

Evaluate the performance of the model against project metrics identified in Phase I. This should follow business/project objectives that provide the ROI or value promised for the engagement.

Measure model performance against the technology KPIs detailed in Phase I.
If the model does not perform adequately, detail the steps needed to improve the KPI performance of the model in this iteration.

Task: Model Iteration Approach

Description

Models will require continuous iteration, especially if they are only marginally providing the desired results, or if you want to improve the results in future iterations.

Task Artifacts

Required Model, Data, or Process Changes:

- Document and explain changes required to any of the previous tasks within the phases to achieve project objectives.
- Note that this may require changes within the tasks in Phases II, III, and IV.

Detail approach that will be used to iterate this model to improve on any of the results in this phase. If any iteration is required to previous phases to improve the results, detail what previous phases need iteration.

Task: Review Process

Description

At this point, the resulting models appear to be satisfactory and to satisfy the objective needs. It is now appropriate to do a more thorough review of the resulting models to determine if there is any important factor or task that has somehow been overlooked. This review also covers quality assurance issues—for example: Did we correctly build the model? Did we use only the attributes that we are allowed to use and that are available for future analyses?

Task Artifacts

Review of Process

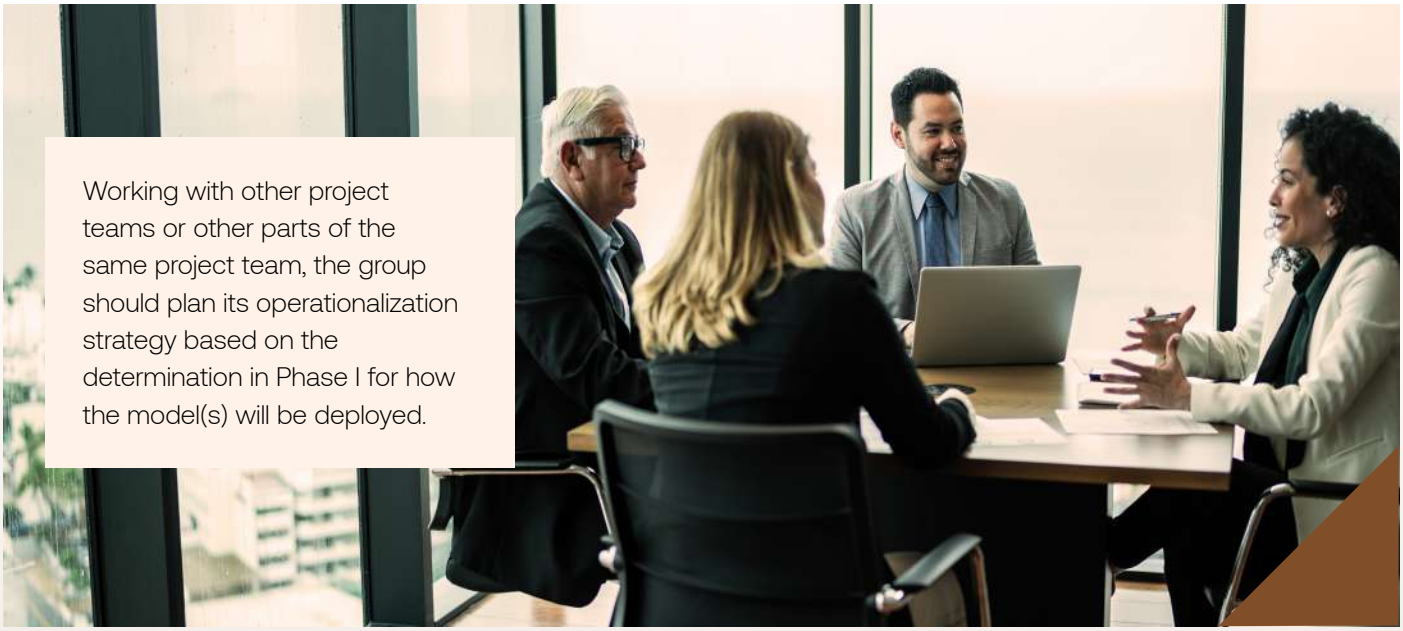
- Summarize the process review and highlight activities that have been missed and those that should be repeated.

Detail any required approvals or reviews to be conducted before the model can be operationalized in production.

Phase VI: Model Operationalization

In this final phase, the project team works to get the model operationalized in all of the locations identified earlier in the project.

Task Group: Model Operationalization Plan



Working with other project teams or other parts of the same project team, the group should plan its operationalization strategy based on the determination in Phase I for how the model(s) will be deployed.

Task: Operationalization Plan

Description

This task takes the evaluation results and determines a strategy for operationalization. Usually, the project team will work with noncognitive portions of the project to discern the steps necessary to put the model into a place where it can be used.

Task Artifacts

Operationalization Plan

- Summarize the operationalization strategy (i.e., batch mode, microservice, real-time, streaming, on-premise, in the cloud, etc.), including the necessary steps and how to perform them.

How will this model be operationalized, and in what mode and location(s)?

What IT processes must be followed to operationalize the model as planned?

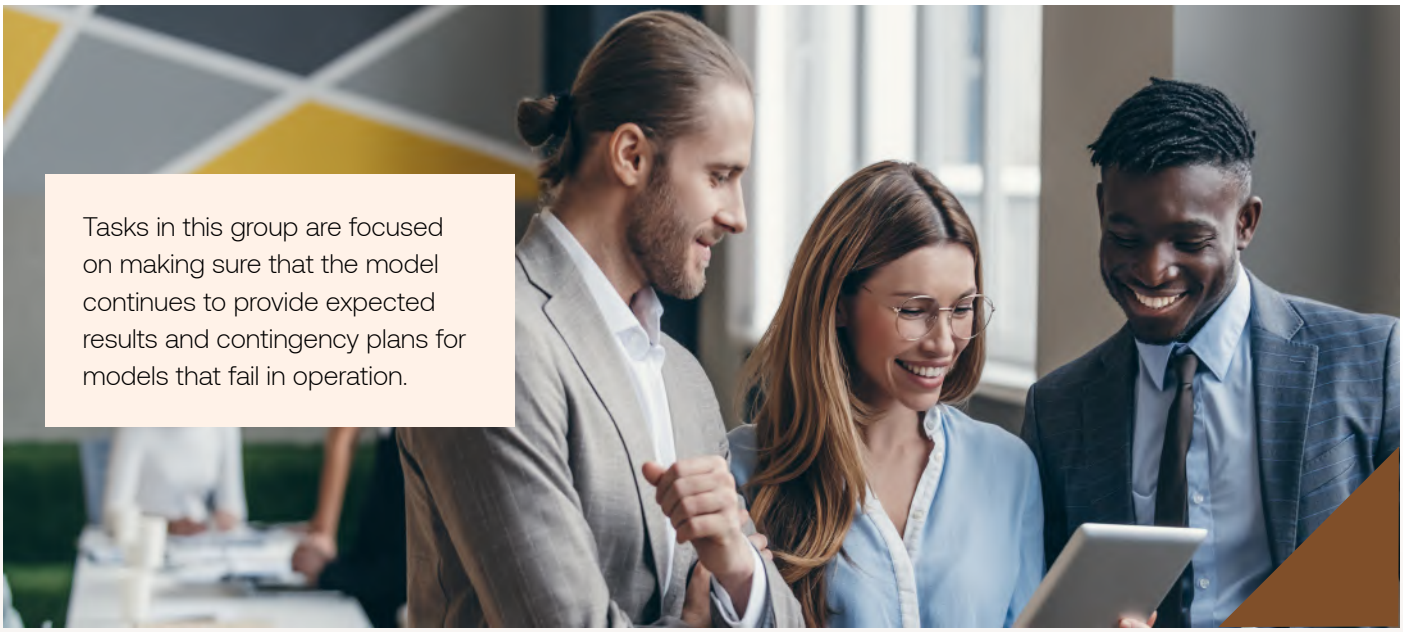
Model Scaffolding

- Detail the additional application development, coding, or other work that should be done to position the model where it can be used.
- This may require the creation of additional user stories or sprints or iterations to put the model into production, thus necessitating placing the completion of this phase on hold until such work is done.

What additional non-AI/noncognitive activities, such as application development should be done to operationalize this model as intended for this iteration?

Perform the task of operationalizing the model as planned, given the processes and constraints above.

Task Group: Model Monitoring and Maintenance



Tasks in this group are focused on making sure that the model continues to provide expected results and contingency plans for models that fail in operation.

Task: Monitoring and Maintenance Plan

Description

Monitoring and maintenance are necessary when models are used as part of the day-to-day business and its environment. The careful preparation of a maintenance strategy helps to avoid unnecessarily long periods of incorrect usage of model results. To monitor the deployment of the model(s), the project requires a detailed monitoring process plan. This plan takes into account the specific type of deployment.

Task Artifacts

Monitoring and Maintenance Plan

- Summarize the monitoring and maintenance strategy, including the necessary steps and how to perform them.

What continuous monitoring and management approach and tools will be used for the model in this iteration?

Implement the monitoring and management as detailed above.

Task: Model Governance Framework

Description

Cognitive models, probabilistic in nature, face a number of real-world challenges when applied in decision-making and other roles. Project teams planning long-term usage of models should define a structure by which decisions will be made about model modification and ongoing usage, based on input and feedback from non-project team users.

Task Artifacts

Determination of Governance Team

- Documentation and identification of members who will serve as the owners of the produced model and responsible for its usage, soliciting feedback from users of the model(s), and fielding questions or concerns about how the model or models are used.

Structure of Governance Team Operation

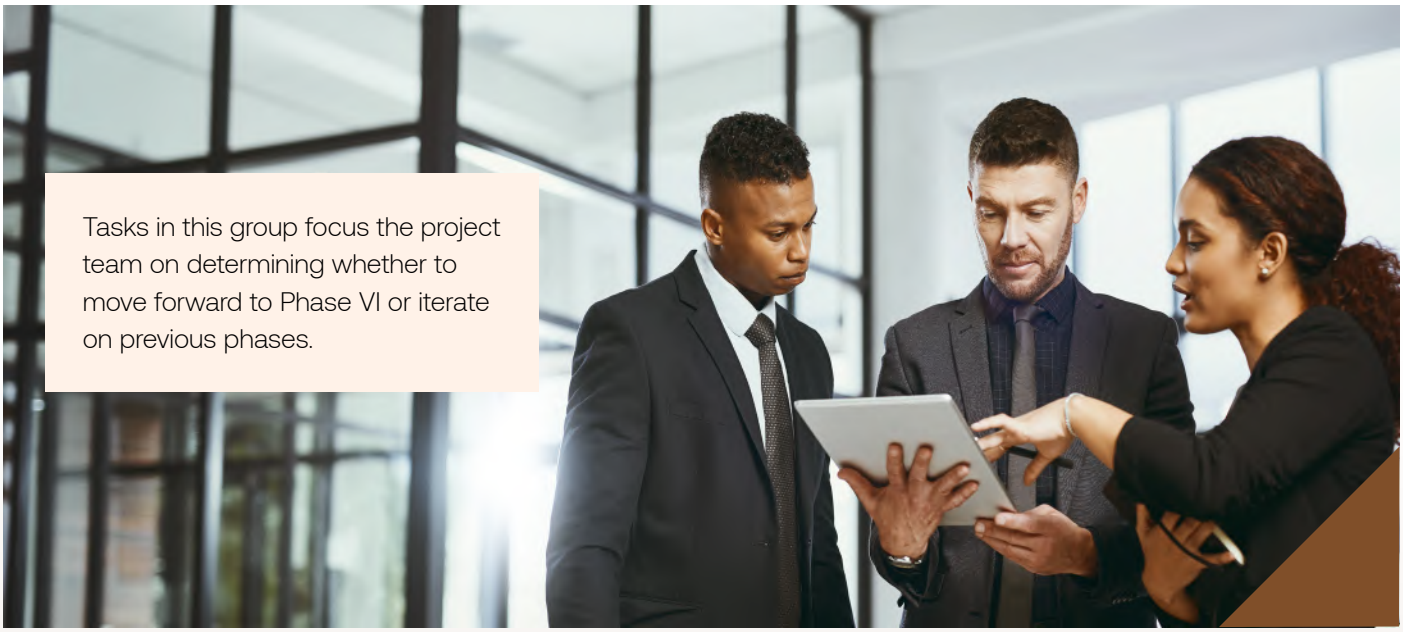
- Summarize a plan for how the governance team will operate, including a means for soliciting feedback from users, an approach for evaluating feedback and concerns, and the steps the team will take to address any issues or concerns.

Required Documentation

- List any required documentation or compliance measures the team will have to take when the model or models have been put into full operationalization.

Determine which group(s) will be responsible for governance and ownership of the model for this and future iterations, as well as the means by which the group will respond to various needs after this iteration.

Task Group: Determine Requirements for the Next Iteration



Tasks in this group focus the project team on determining whether to move forward to Phase VI or iterate on previous phases.

Task: Determine Next Steps

Description

Depending on the results of the assessment and the process review, the project team decides how to proceed. The team decides what the next iteration should require for this project or other AI projects. This task includes analyses of remaining resources and budget, which may influence the decisions.

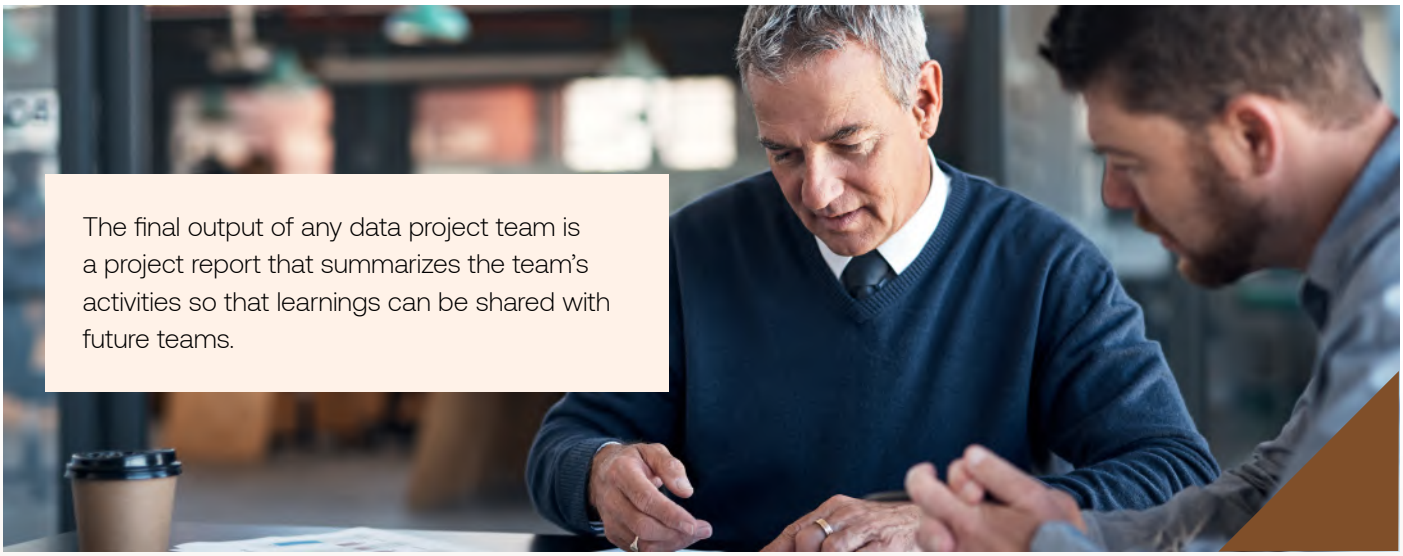
Task Artifacts

Future Iterations

- List the potential further iterations, along with the reasons for and against each option.
- Decide how to proceed, along with the rationale.

What should be done in the next iteration for this AI project?
What resources exist to pursue the next iteration?

Task Group: Project Report



The final output of any data project team is a project report that summarizes the team's activities so that learnings can be shared with future teams.

Task: Produce Final Report

Description

At the end of the project, the project team writes up a final report. Depending on the operationalization plan, this report may be only a summary of the project and its experiences (if they have not already been documented as an ongoing activity), or it may be a final and comprehensive presentation of the project result(s).

Task Artifacts

Final Report and Presentation

- This is the final written report of the project engagement. It includes all of the previous deliverables, summarizing and organizing the results.
- There will often be a meeting at the project's conclusion at which the results are presented to the customer.

Task: Review Project

Description

Assess what went right and what went wrong, what was done well, and what should be improved.

Task Artifacts

Experience Documentation:

- Summarize important experience gained during the project. For example, pitfalls, misleading approaches, or hints for selecting the best-suited data-mining techniques in similar situations could be part of this documentation. In ideal projects, experience documentation also covers any reports that have been written by individual project members during previous phases of the project.

Perform an iteration post-mortem. What went well during this iteration? What did not go well? What can be improved for future AI project iterations?

Sources and Endnotes

- Some material was sourced from CRISP-DM v1.0 documentation (crisp-dm.org), the ownership and copyright of which remains with CRISP-DM.