

Worksheet for AI Application Planning

This worksheet is aimed to help you, a potential client for an AI application development project, express your intents to an AI expert who would develop the AI application for you. Based on the information you provide, the AI expert would form their understanding of the desired AI application and have discussions with you around project feasibility and potential initial approaches.

Please use this worksheet to help guide you through the necessary decisions you need to consider making to get an AI application development project started. Please utilize the outline shown on the left to freely navigate through the different sections of this worksheet and fill it in in the order you wish to.

I. Overview

First, we will sketch a bird's-eye view of the AI application you wish to build.

1. AI Application Summary

Please fill in the following statement summarize what your AI application would do (feel free to modify the statement as you see fit):

I intend to build an AI application that [application description] so that [target user] [end goal].

(e.g.,

- I intend to build an AI application that [predicts the supply of beef in 6 months] so that [the government employees] [can make informed decisions about policies for stabilizing the price].
- I intend to build an AI application for [generating personalized recipes based on dietary or religious restrictions, preferences, etc. of the consumers] so that [automated kitchen owners where robots do the cooking] [can provide personalized dining experience].
- I intend to build an AI application for [generating detailed sketches of cartoon cuts based on the artist's style based on the artist's rough sketches] so that [cartoon artists] [can easily explore the design space with minimal effort].
- I intend to build an AI application for [providing personalized recommendations of courses] so that [the users of my education platform] [can continue their path of learning on the platform].)

2. Target User

Q1. What is your relationship to the target users of the AI application? If there are multiple classes of target users, please describe your relationship to each class of target users.

A1. (e.g., myself, employees of my company, users of a web service that my company runs, etc.)

Q2. Please provide any demographic characteristics of the target users of the AI application.

A2. (e.g., adults living in the US, children, students, people who regularly use the internet, etc.)

II. Dataset

Dataset is a crucial part of AI application projects as it is used for training the AI model that powers the AI application. Information about the dataset helps AI experts not only determine the feasibility of the AI application development, but also choose the AI model that would work best for the data. Furthermore, AI experts can get a sense of the input and output format of the AI application and know whether the existing data needs additional processing or whether additional data should be collected.

1. Dataset Availability & Collection

Q1. Is the dataset available?

A1. Y/N.

Q1-Y. (If you answered 'Y' to Q1) Are the labels (the expected outputs given the input data) available?

A1-Y. Y/N

Q2. How did you / would you collect the data and the label?

A2.

Dataset Source	(e.g., collected by my company, dataset in the public domain)
Data Collection Method	(e.g., web scraping, user log collection, production line sensor data. If you need to explain further, please do so.)
Data Collection Tools & Environment	(e.g., Scrappy (web scraping tool), Camera D-60 with controlled lighting environment. If you need to explain further, please do so.)

Label Collection / Generation Method (if labels are collected / generated separately)	(e.g., would do it myself, hiring expert labelers, crowdsourcing. If you need to explain further, please do so.)
---	--

Q3. Please provide a link to the dataset (and labels) if they are available. If the full dataset is unavailable at the moment, please provide some sample data points.

A3. <https://example.com> or send later via email.

Q4. How many data points do you have in the dataset / can you collect for the dataset? A rough estimate is fine if you do not know the exact count.

A4.

In Possession	(e.g., 7549, 500-600, 10s of thousands)
Can collect in addition	(e.g., 7549, 500-600, 10s of thousands)

2. Dataset Format

Q1. What is the format of the (input) data?

A1. (e.g., data table, plain text, images, documents with images, voice, mesh, etc.)

Q1-T. If the data is of a tabular format, please list and describe each of the data fields.

A1-T.

Field	Description (if not obvious from the field name)	Value Range
(e.g., # children)		(e.g., 0, 1, 2, ... (integers))
(e.g., temperature)		(e.g., usually -10C – 40C)
(e.g., colors)		(e.g., 3 categories (red, green, blue))
(e.g., userLoggedIn)	(e.g., Whether the user is logged in when using the web)	(e.g., binary (T/F))

	service)	
(Add more rows if needed)		

Q1-I. If the data includes images, please describe the image format.

A1-I.

Color Scheme	(e.g., black-white, RGB, RGBA, 26 channels, etc.)
Image dimensions / resolution	(e.g., 1000x600, 300dpi, etc.)
Other Notes about the Format	

Q1-T. If the data includes plain text, please describe the text format.

A1-T.

Language	(e.g., English, Spanish, Python, etc.)
Other Notes about the Format	(e.g., text length, etc.)

Q1-O. If the data includes other types of data, please describe the format of that data format.

A1-O. (e.g., file format, encoding format, etc.)

Q2. What is the format of the (output) label? If the output is expected to be a different type (e.g., image, sound, mesh, etc.), feel free to modify the table to the format in above questions (Q1-I, Q1-T, Q1-O).

A2.

Label	(e.g., signal color)
Description (if not obvious from the field name)	
Value Range	(e.g., 3 categories (red, yellow, green))

3. Data Insights

Q1. How is each of the fields (if the data is of tabular format) correlated with the (output) label? If the data includes non-tabular data, what aspects of the data may be relevant in predicting or generating the (output) label?

A1.

Field / Aspect	Relationship to (Output) Label
(e.g., Temperature)	(e.g., Optimal production around 36C)
(e.g., Speed)	(e.g., Survival less likely when speed is higher)
(Add more rows if necessary)	

Q2. Based on your knowledge about the domain, what do you expect as the most relevant features or aspects of the data in predicting or generating the (output) label?

A2. (You may give more than one)

Q3. Please provide any insights on the distribution of the data.

A3. (You could include things like: data collection date (or time period), images taken indoors / outdoors, frequency of the data (weekly, monthly, etc.), etc.)

Q4. If the (output) label is categorical or binary, how much data do you have or hope to get for each of the labels?

A4.

Label Category	# of Data Points
(e.g., True)	(e.g., 500)
(e.g., False)	(e.g., 1000)
(Add more rows if necessary)	

Q5. Do you expect any anomalies in the data (e.g., missing data, bad measurements). If yes, please explain.

A5. (No or Yes + explanation)

4. Policies and Regulations on Data Use and Access

Q1. Are there any legal issues with the local government regulations on using the data in your dataset? If yes, please explain and how it may be circumvented.

A1. (No or Yes + explanation; There may be regulations on medical or educational information or other personally identifiable information. For creations, there may be copyright issues. Some regions have specific regulations – e.g., to work with images including faces in Russia, you must get special approval.)

Q2. Are there any policies from your organization that limits AI experts' access to data? If yes, please explain the policy and how it may be circumvented.

A2. (No or Yes + explanation; e.g., Your organization may require signing of NDAs for data sharing, allow access to a subset of the data (would internally improve the model with the remaining private data), or require accessing the intranet within your organization)

III. Defining Success

Next, you should think about the specific goals the AI applications should meet in order to consider it a 'success'. For the AI experts, knowing your priorities in defining success helps them determine what AI model to use (since there are usually tradeoffs between two or more targets). They would also try to figure out whether such an AI model is feasible and may initiate discussions about tempering your expectations. It also helps them define a development strategy and AI metrics to use during development to keep track of their progress.

1. Metrics

Q1. What metrics do you or others in the domain currently use to monitor the performance in the target use case of currently deployed algorithms or humans? How are they measured? How good is the state-of-the-art / your currently deployed algorithm or humans?

A1.

Metric (with explanations if not obvious)	Measurement Method	Current Value
---	--------------------	---------------

(e.g., Pixel defect rate)	(e.g., Have humans check for products with pixel defects visually. Compute as [# of products with defects] / [total number of products])	(e.g., 0.05)
(e.g., Expert-verification)	(e.g., An expert eyeballs the output and either gives it a 'pass' or 'non-pass')	(e.g., unknown)
(Add additional rows if necessary)		

Q2. Which of the above metrics do you care about for evaluating the AI application? If there are more than one, please list them in the order of importance, the most important first.

A2. (A list)

Q3. Are certain types of errors that would be more harmful than others? Please explain.

A3. (e.g., For diagnosis of cancer, false negatives (saying that you don't have cancer when you actually do) may be more harmful than false positives (saying you have a disease when you do not), since false positives can easily be amended through secondary checks, but false negatives can leave cancer undetected.

For used car price prediction (for a company trying to buy used cars), underestimating car costs may be less harmful than overestimating car costs, since underestimating car costs will result in a no-deal without any financial loss, but overestimating car costs will actually incur a loss.)

Q4. Will you need to directly verify the results to check whether the AI model is performing 'well enough'?

A4. (Y/N)

Q5. [OPTIONAL] Do you have any suggestions for AI metrics for the AI experts to use during development

A5. (Optional, if you have thoughts; e.g., AUC, IOU, F1, etc.)

2. Success Threshold

Q1. In your own words, what are your expectations from the model? i.e., when would you consider the deployment of the AI application a success?

A1. (Answer)

Q2. On the aforementioned metrics, do you have specific thresholds you wish to meet with the AI application? If so, what are those thresholds?

A2. (Y/N)

Metric	Target Threshold
(e.g., pixel defect rate)	(e.g., 0.02)
(Add more rows if necessary)	

IV. Model Needs

Now, we will ask you about the needs you have about the model so that the AI experts can decide which AI model to use.

1. Core Model

Q1. Do you have any requests on the specific class of AI model (e.g, GPT/LLM, deep learning, transformers) you would like to use for the AI application?

A1. (N/Y + specific class of AI model)

Q2. It may be the case that there may be a way to build the desired AI application without the use of AI. Depending on the amount and quality of available data and the problem setup, a non-AI-based solution may be less costly and may perform on par with an AI-based solution. Are there specific reasons for deciding to use AI and would non-AI based solutions be of interest to you?

A2. (N/Y + Reasoning on why you need to use AI (e.g., a managerial decision, a requirement by the investors, AI has proven to be giving state-of-the-art results in the domain, competitor has successfully built an application based on AI, etc.) + I would be interested / would not be interested in non-AI-based solutions.)

Q3. How often do you expect the input to enter the system and how fast should the AI application be? What is the expected number of concurrent uses of the system?

A3.

Frequency of Input Entry	(e.g., one-time, monthly, streaming data, every time a user enters our service, etc.)
Runtime of the AI Application	(e.g., real-time, within 5-10 seconds, several hours is fine, etc.)
Concurrent Uses	(e.g., 1 – me, 1 – production monitoring system in the production pipeline, 100-1000 – users of our service, 15 – at each production site)

Q4. Do you need the system to be explainable? i.e., do you need to understand the reasoning process of the AI model (e.g., regions of an image that were used to make a prediction, the step-by-step reasoning process for question-answering, etc.) or the general behavior of the AI model (e.g., contributions of each of the features in determining the (output) label)?

A4. (N/Y + Type of explainability you desire (sample explanation you would like to see would be nice))

Q5. What is the scope of inputs that the AI model should be able to handle, especially if in terms of abnormal inputs or variations in input styles?

A5. (e.g., may assume that the input will always include at least one face but should be able to capture multiple faces, should be able to handle out-of-domain questions, should be able to handle multiple artist 'styles' in input art pieces, etc.)

2. Regional and Corporate Restrictions on AI and Framework

Q1. Are there any legal issues with the local government regulations on using AI for the specific problem? If yes, please explain and how it may be circumvented.

A1. (No or Yes + explanation; There has recently been a movement in various governments around the world in regulating AI (e.g., the EU AI Act for regulating safety standards))

Q2. Are there any policies from your organization that limits AI experts' use of certain frameworks or AI models? If yes, please explain the policy.

A2. (No or Yes + explanation; e.g., Some organizations ban or limit usage of GPT due to the possibility of data leakage. There are also some organizations that limit the use of publicly available code. Furthermore, some organizations ban the use of certain frameworks.)

3. Output Filtering Needs

Q1. Are there types of output that need to be filtered out or modified? If yes, please explain.

A1. (No or Yes + explanation; For image generation, images including nudity, violence, corpses often need to be filtered out or blurred. If you need to filter out or modify something not obvious (as the ones mentioned), you may need to provide examples.)

V. Budgets, Resources & Framework

In this section, you will answer some questions about the availability of resources for development and deployment. This information would also help adjust what AI model the AI expert would use to power the AI application. With unlimited resources, the most powerful AI model could be optimal, but the limitation of resources usually adds limits on the power of the AI model you could use for your AI application.

1. Development Resources

Q1. How much compute power can you provide to the AI expert for development? Specifically, can you provide GPUs (graphics processing unit) (if so, what type of GPUs) or other types of cloud computing resources?

A1. (e.g., we can provide access to GPU (Tesla V100), we can provide \$500 credit for training on the AWS (Amazon Web Services) server, will provide OpenAI API key, etc.)

Q2. What is your budget for development?

A2. (Rough estimate if there is a limit)

2. Deployment Resources & Framework

Q1. Which framework and computing resources will the AI application run on when it is deployed?

A1. (e.g., web service hosted by AWS, mobile (Android), will have access to our GPU (Tesla V100) during deployment, etc.)

Q2. What is your budget for upkeep?

A2. (Rough estimate if there is a limit. Any limits on API usage, if using APIs?)

VI. Understanding the Client Situation

Now, we ask you to provide some information about your situation. While the AI expert is an expert in AI and may have some knowledge about your application domain, they probably do not know well about your specific situation. This information should help them gain background knowledge about the domain and your situation so that they can be on the same page when having further discussions or making decisions.

1. Target User Workflow

Q1. In your own words, please describe the current workflow of the target user without the AI application.

A1. (Description of the target user workflow. You could include the steps that the target goes through, other stakeholders the target interacts with, analysis methods, etc.)

Q2. In your own words, how would the target user's workflow change with the introduction of the AI application?

A2. (Expected changes. Could be as minor as some kind of a productivity improvement, but could also completely automate certain stakeholders, etc.)

Q3. If any, what are the difficulties you (or the target users) face(d) during the workflow that made you seek an AI-based solution?

A3. (List of difficulties and explanation)

2. Related Work

Q1. Have you made any prior attempts to build the AI application? If yes, please provide details of the prior attempt and why you need to go through another attempt.

A1. (N/Y + details + why you are going through another attempt (e.g., failed project – didn't have enough data at the time, but now we have collected sufficient data over time, new technology –

we used a classical AI model (SVM) for building our model, and with the rise of deep learning, we believe that we could do much better))

Q2. [OPTIONAL] If you are aware of related papers or attempts from other organizations, please share them.

A2. [OPTIONAL]

Title	Link (or send us the files via email)	Brief Description
(Add more rows if necessary)		

3. Non-data field factors

Q1. Are there any external factors that are not part of the dataset that could affect the model? Please explain.

A1. N/Y+

Factor	Description (if not obvious)	How it Affects the Model
(e.g., regional education level)		(e.g., could affect how well the target users can parse and understand the output)
(e.g., abrupt changes in new car production)		(e.g., if new cars become harder to come by, prices of used cars would have a tendency to rise)
(e.g., media coverage)		(e.g., media showing ice cream could in general have positive effect on ice cream demand, but media showing potential health risks could have negative effect on ice cream)
(Add additional rows if necessary)		

VII. Client-AI Expert Agreement Terms

You should define the terms that you expect the AI expert to meet for you. The terms would likely be adjusted through further discussion.

1. Timeline

Q1. What is the timeline of the project and what are the key milestones (if any specific milestones exist)? How important are these timelines?

A1.

Milestone	Target Date	Timeline Importance
(e.g., Initial end-to-end prototype)	(e.g., 2024 Mar, 7 weeks, etc.)	(e.g., more of a suggested date, 2-3 month delay is acceptable, etc.)
(e.g., Project completion & deployment)	(e.g., 2025 Dec, 17 months, etc.)	(e.g., must meet the deadline in order to meet the system deployment schedule, must meet the deadline for a paper submission deadline, etc.)
(Add more rows if necessary)		

Q2. Is the project in a rush? If the project is in a rush, it may be possible to expedite the process to some degree, but will incur higher costs as it would require the AI expert to focus and dedicate resources to the AI application development.

A2. (N/Y + I really need the project to complete by [date] and am aware that this will incur higher costs)

2. Deliverables

Q1. What deliverables must the AI expert deliver to you?

A1.

Deliverable	Description (if not obvious from the name)	Deadline (if different from the project completion date)
(e.g., code)		

(e.g., re-trainable model with training instructions)	(e.g., must retrain the model with internal data not shared with the AI expert)	
(e.g., augmented / preprocessed dataset)		(e.g., By the initial end-to-end prototype milestone)
(Add additional rows if necessary)		

3. AI Application Functionality / Features

Q1. Please list out the required functionalities and features of the AI application that the AI expert must meet, other than the underlying AI model itself (input-to-output).

A1.

Functionality / Feature	Description (if not obvious)	Deadline (if different from the project completion date)
(e.g., user input automatic processing for AI model input)	(e.g., user inputs raw image and will require automatic preprocessing)	
(e.g., explainability: weights on each dataset feature in determining the label)		
(Add additional rows if necessary)		

VIII. Other Notes

Finally, provide any further information, including information regarding the nature of the project and your organization.

Q1. What is the nature of the project? (e.g., research prototype building, personal use, internal use, public deployment, etc.) Are there major consequences of AI application malfunction (e.g., death, etc.)? If yes, please explain.

A1. (Nature of the project, N/Y + explanation)

Q2. Are there specific characteristics of you or your organization the AI expert should know?

A2. (This could include organization structure, organization philosophy, etc.)

Q3. Anything else you would like to inform the AI expert about?

A3. (Free-form, you may leave this blank.)