

AI Project Worksheet

Elements that stakeholders, business owners, and client-side investors should look into for successful AI project implementation at scale.

Introduction

Over the years, the software development communities became much more aware of having an interactive AI lifecycle for product development and refinement. The AI project cycle is the key process of initiating, planning, executing, monitoring, controlling, and developing feedback loops for continuous improvement of any software product. With time, planning out every stage of this lifecycle from earlier stages of AI project development became an indeed practice. If you are a business planning to adopt AI for scaling a solution, product, or service having an interactive AI lifecycle is key.

Eventhough it may be tempting to skip the purpose of answering and understanding every step in the process, it is crucial if you want to achieve the desired results by the expected date and time. Communicating the needs and objectives more effectively with data scientists from the get-go can eliminate potential project roadblocks and helps you in preparing for further steps in the AI lifecycle to achieve business growth.

To develop AI software solutions that meet your objectives it is important to plan every detail along the interactive AI lifecycle with care. Whether you're working with an in-house team or an external partner to build your AI solution, it is always important to understand and answer every question around the project requirements to ensure timely delivery. The goal of this AI project worksheet is to help you address five imp, optimized costs, and valuable outcomes. Important conditions that are essential and will lead you to successful project completion.

Business Need

Defining the business objectives, needs, and target domain help to bridge a knowledge and language gap between domain experts and data scientists. The more understanding or specialized domain expertise support you provide, the better the data science team will perform and assure model accuracy or in general, which may cost them more effort to prepare for the project.

The questions that should be asked/defined in your initial conversation with a data science partner.

Describe a use case that your solution will address?

Describe the relevant scientific and/or technical area.

Define what you as a business expect from the project?

Specify the business goal by explaining what results you as a business expect to see from the project after deployment?

Explain how this project results will affect end-users workflows.

Explain how results will affect decision-makers' workflows.

Explain how project results will be evaluated?

If there is an existing project that offers the same solution, explain how it works and why it needs to be replaced or improved.

Data

Data is the backbone of any AI project, and it is important to help data scientists by providing the contents of your data like:

- Specify the length of the historical data available (in months):
- Specify the quantity of data (size or number of records, e.g. X tables totaling Y TB of data):

- **Mentioning the data format**

- ☐ Image/Video data
- ☐ Database or Data warehouse
- ☐ Flat file (Excel, CSV, JSON, XML, etc.)
- ☐ Other

- **Assess and specify data quality:** Even if you don't have an expert who is qualified to provide a definitive assessment, a rough estimate can help the data science team evaluate the data preparation requirements.

- ☐ Perfect (no wrong/missing data)
- ☐ Reliable (very little amount of wrong/missing data)
- ☐ Sufficient (some wrong/missing data)
- ☐ Insufficient (a large amount of wrong/missing data)

Environment and Tools

Specify how the data science team can access the data. Whether the data will be provided on physical media or shared digitally. If the data is stored on a remote desktop, define the operating system, whether it's Linux, Windows, macOS. If you have activated cloud service specify the details for the service provider.

After defining the data environment, specify your preferred data science tool. Eventhough python is the most widely used data science programming language, it is important to mention your data science partner about other tools such as R, Scala, PySpart, etc. if you prefer. As it may provide heads-up to the team on resource and staff allocation

Output and Deliverables

Define how you need the scoring to be executed. The frequency of output required from the project will influence the amount and nature of resources required from the client-side and your data science partner.

For example, some models might necessitate the involvement of a subject matter expert to conduct “sanity checks” to identify potential mistaken results. Meanwhile, in the case of batch and real-time scoring, it might make sense to retrain the model to improve accuracy or running time, which will require the involvement of data science expertise.

Specify the scoring requirements based on your preferences:

- ☐ Batch scoring
 - ☐ Perfect (no wrong/missing data)
 - ☐ Reliable (very little amount of wrong/missing data)
 - ☐ Sufficient (some wrong/missing data)
 - ☐ Insufficient (a large amount of wrong/missing data)
- ☐ Real-time scoring
- ☐ One-off output

It is essential to specify the deliverables that you require starting from types of documentation required such as pdf, doc, txt, ppt, other (specify) to mentioning the deliverables of development code, models, production code, etc. if you need the results to visualized on a dashboard, define the tools to specifically to be used like Tableau, or a custom dashboard, and others.

It is always preferable to specify what systems and pipelines the solution will be integrated into (if any). Your use case might benefit from integrating the product of the data science project into a system like Docker, as it will enable easier – even automatic – subsequent integration and allow the business to focus on the results instead of the technical details. However, this will likely increase costs, because most integrations require additional engineering

Deadlines

Although the above-mentioned set of items requires cooperation from your data science partner to finalize, sharing an idealized client-side vision of the delivery schedule provides pillars of project planning and scheduling. It is ideal to share specific details like:

- Start date of the project:
- The last date of delivery:
- Defining the necessary development milestones needs to be achieved during the target time period.

Now that you've understood the core elements/requirements to start any data science project, connect with your data science partner to plan your next steps together.

Need further assistance or looking for a potential partner to achieve your data science goals reach out to us!

SoulPage IT Solutions is a leading data science company, pioneering a state of strategies to help businesses in addressing challenges by integrating next-generation technologies like AI, Machine Learning, Deep Learning, RPA.

Our technological and strategic expertise can guide your business irrespective of the size you operate to scale ai outcomes for your business.

Whether you are foundational ready or transformationally ready, soulpage acts as an extended arm in driving technological advantage for your business. to know more about us visit www.soulpageit.com

