# AI Best Practices

## version

**William Jones**

February 22, 2023

# Contents

# Welcome to ai-best-practices's documentation!

## Executive Summary

Artificial Intelligence (AI) and Machine Learning (ML) technologies are key to many modern engineering projects due to their ability to solve many problems that are difficult or impossible with other methods. While most engineers will find themselves enjoying a significant overlap between these techniques and their existing skill set, they are also liable to find that AI and Machine Learning is its own field with its own unique demands and (often hidden) pitfalls. While there are many resources available for self teaching, it is generally assumed the practitioner is either an absolute beginner to engineering, or already a seasoned expert in AI and ML. In this document, we provide a practical guide to AI and Machine learning for electronic systems engineers who already have a strong base of knowledge in electronic systems but no specialized expertise. This guide will be practice focused, with the goal of helping engineers to make good decisions and avoid problems. The guide will cover, among many others areas:

- Preliminary problem assessment
- Data collection
- Pipelining
- Testing and Validation
- Privacy and Security
- Ethical and Legal considerations

## Target Audience

The target audience for this guide is electronic systems engineers with little to no specific expertise in AI and Machine Learning techniques. It will be assumed, because of this background, that readers will have a base level of competence in programming and mathematics, though the guide will err on the side of caution in respect of assumed knowledge. We may, for example, assume our reader has knowledge of concepts in basic calculus, but are unlikely to assume that they are able to remember any specific formula. We are also assuming that the primary interest in practically deploying these techniques rather than understanding the theory and context of their development. For readers interested in a more theoretical treatment, we list several texts in the resources section appropriate to a range of different levels of background knowledge.

## Using this Guide

This guide provides a preamble discussing the key issues in each section, followed by a questionnaire for each section dictating, for each AI risk and privacy level, key steps to take. Where AI and Privacy levels are both specified, you must fulfill this requirement if you meet either of these specifications.

In terms of structure, this guide specifies several different categories of requirements. Creating an effective AI project is an iterative process, and it is not intended that the ordering of the requirements specifies a strong order in which to tackle the requirements, beyond the obvious (e.g. we must collect the data before we can do anything with it). The only exception to this is the Preliminary Assessment that forms the first step of any project, and establishes the AI and Privacy risk levels discussed below.

### AI Risk Levels

One of the key concepts leveraged in this guide in order to provide an appropriate level of risk management to AI is the classification of different AI applications into different risk categories. This framework strongly mimics the approach proposed in the upcoming EU Artificial Intelligence act, and defines four categories of risk for AI:

- Unacceptable risk. Applications that are a clear threat to the safety, livelihoods or rights of people.
- High risk. Applications that are critical to the safety or wellbeing of people, that require special mitigations to adequately cover the risks they pose
- Limited risk. Applications with lower impacts to safety or wellbeing, but which still require some level of mitigation

- Minimal or no risk. Applications with minimal to zero risk.
- We define these terms in more detail, with examples, in the appendix.

The framework this guide provides dictates different requirements according to each category.

## Privacy Risk Levels

The other key concepts used in this guide are privacy and security requirements. As data driven methods, AI and Machine Learning algorithms have the potential to pose significant risks in these areas. We define three levels of privacy and security requirements:

- High risk. Applications that deal with personal data that may pose a risk to the safety or wellbeing of people
- Limited risk. Applications that deal with personal data with lower impacts to safety or wellbeing, but which still require some level of mitigation
- Minimal or no risk. Applications that either do not deal with personal data, or do so with minimal to zero risk.

# Preliminary Assessment

In order to usefully realize the content of this guide, the AI and Privacy risks that the project poses need to be defined. To do this, we need to undertake several preliminary steps to define the project.

The very first step is to establish that AI and Machine Learning algorithms are an effective and suitable approach to the proposed problem. The goals of the project should be reviewed in the context of what AI and Machine Learning algorithms can achieve, and compared to other (non AI) approaches. Simpler approaches may lead to comparable or better results, without the overhead of managing an AI and Machine Learning project.

The next step is to scope out the goals and requirements of the proposed product in the context of the business constraints it is subject to. Given the challenges of validating AI approaches, it is imperative to define clear, specific goals for the project. When the goals and requirements of the proposed project have been established, well defined metrics of what success in achieving these goals will look like need to be created.

Once these goals and metrics are defined, it is then appropriate to undertake a risk assessment of the project. This risk assessment should, without exception, inform a fair assessment of the AI risk and privacy risk category that the proposed projects outcomes fall into.

| Requirement | Evidence | AI Risk | Privacy Risk | Complete |
|---|---|---|---|---|
| Assess alternative approaches | • | Mandatory | Mandatory | • |
| Risk Assessment | • | Mandatory | Mandatory | • |
| Risk Categories | • | Mandatory | Mandatory | • |

# Data Collection

In all approaches that learn from data, the quality and availability of data are of paramount importance.

The first step is to establish what kind of data needs to be collected in order to solve the problem of interest, and where and how it is going to be obtained at a high quality. It’s important to consider how much data is likely to be needed to solve the problem while maintaining realistic expectations of what data is cost effectively obtainable. It’s also important to pick data that both accurately represents your quality of interest, and is amenable to computation. AI methods are fundamentally statistical methods that deal with quantitative data, and subjective measures (e.g. free text answers) are substantially more difficult to manage. Whatever is decided, it is important to consider and record factors that may impact data quality, or introduce sampling biases.

The process of establishing a data collection pipeline should be an iterative process: it is unlikely the first solutions to the above questions will be the best ones. Data exploration should be undertaken to establish interesting patterns, and to identify problems in the data (and remedies thereto) as early as possible. Note that it is possible to introduce biases by doing this carelessly. Even with a carefully designed data pipeline, data will, almost universally, require preprocessing before it is handed off to any AI pipeline. It may, for example, require cleaning, transformation, and/or labeling. This step may be revisited in light of later processing stages, but should still be carried out at this stage. In all instances, it is imperative that the processes used to achieve this are tracked and can be replicated. Raw

collected data should be permanently stored where practicable for limited risk projects, or always for high risk projects.

The risk level and privacy requirements of the AI application may dictate additional steps. All projects should undertake a risk assessment on the potential risks and harms the data collection process may cause, but High risk projects must further provide an appropriate mitigation strategy to reduce identified harms, and human oversight. High risk projects must also take adequate steps at the data collection stage to screen data for biases that may cause discriminatory outcomes of the AI and Machine Learning algorithm. Projects with high privacy requirements must take extra steps at this stage to ensure that data is kept properly secured and anonymized, in addition to the ones specified in the “Privacy and Security” section.

| Requirement | Evidence | Risk Requirement | Privacy Requirement | Complete |
|---|---|---|---|---|
| Establish the goals that the data collection process is achieving | • | Mandatory | Mandatory | |
| Select target data | • | Mandatory | Mandatory | |
| Target data quality analysis | • | Mandatory | Mandatory | |
| Establish data quantization process | • | Mandatory | Mandatory | |
| Create data quality checks | • | Mandatory | Mandatory | |
| Create data collection process | • | Mandatory | Mandatory | |
| Create data pre-processing pipeline | | Mandatory | Mandatory | |
| Establish infrastructure for storing raw data | | High | None | |
| Data biases and discrimination analysis | | High | None | |
| Data collection harm risk assessment | | Mandatory | Mandatory | |
| Data collection harm risk mitigation strategy | | High | Moderate | |
| Data is anonymised | | High | High | |
| Data transmitted during the collection process must be encrypted | | High | High | |
| Data must be secured after collection | | High | High | |

# Indices and tables

- **genindex**
- **modindex**
- **search**