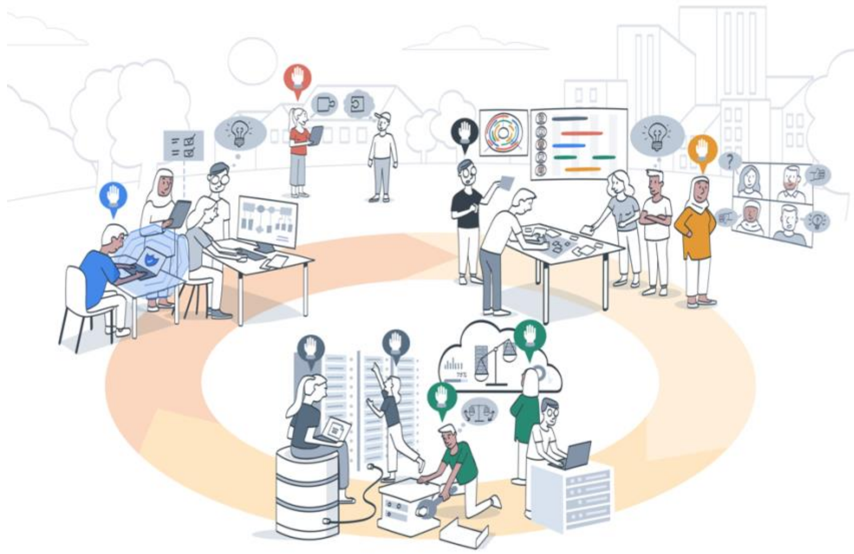
The Project Lifecycle Model Cheat Sheet



The Project Lifecycle Model Cheat Sheet

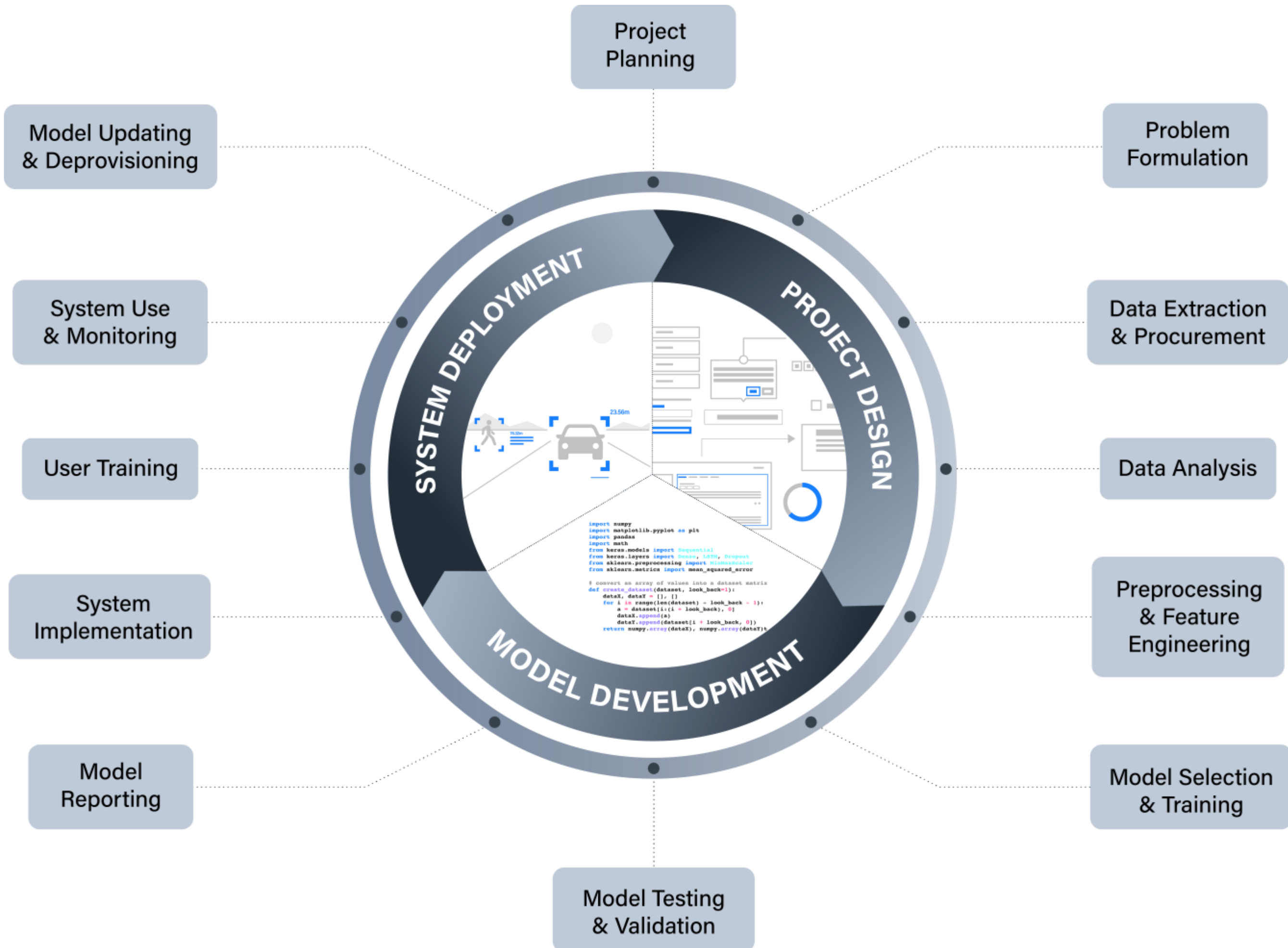
The project lifecycle model is a heuristic model that scaffolds reflection, deliberation, and practical decision-making across all stages of a project's lifecycle. The model is intended to be used by project teams (e.g. researchers and developers) to support the development of safe and ethical data-driven technologies.

It is a *heuristic model* because it represents the typical stages and tasks of a project that are undertaken to design, develop, and deploy a data-driven technology. However, like all models, it is an abstraction from the actual day-to-day practices that are carried out by a team:

> "All models are wrong, but some are useful."
>
> — George Box

> ⓘ About this Document
>
> This document is a cheatsheet for the project lifecycle model. It includes a graphic of the model, a table of the model's tasks, and a summary of the ethical significance of each task.
>
> For further details, please see the full module on the Turing Commons website.

The Project Lifecycle Model



The Project Lifecycle Stages

Project Lifecycle Stage	Information
Project Planning	**Description** The project planning task encompasses the preliminary activities that are intended to help determine the aims, objectives, scope, and processes associated with the project, including an assessment of the potential risks and benefits. **Ethical Significance** • Creates a space for anticipatory and reflective activities (qua AREA framework) that help create a stable foundation for the project. • Offers an opportunity for the team to agree on any "red lines" (e.g. contexts or domains in which a system should not be used, data types that are not permissible to collect or use). • Allows project team to set milestones and objectives that can be used throughout the project to determine if their original goals have been achieved.
Problem Formulation	**Description** This task involves the formulation of a clear statement about the overarching problem the target system or project seeks to address (e.g. a research statement or system specification) and a lower level description of the computational procedure that instantiates it (e.g. a functional mapping from input to output variables and explanation about why it is appropriate). Use of the term 'problem' is intended to focus attention on the fact that the project team is attempting to solve a problem, rather than just build a novel system. This helps to avoid the bias of 'Maslow's Hammer', in which you have a pre-existing solution (the hammer) and go looking for a problem (a nail) to solve, regardless of whether it is the right tool for the job.

<table>
<tr><th>Project Lifecycle Stage</th><th>Information</th></tr>
<tr><td></td><td>

Ethical Significance

The importance of this stage is split across the two interlocking understandings of the term “problem”:

- As a statement about a well-defined computational process (or a higher-level abstraction of the process), this task helps identify the validity and legitimacy of the project. For example, an algorithmic system that attempts to predict a candidates ‘employability’ (the target variable) on the basis of a model trained on biased data from historical hiring practices will be perceived as unjust.
- As a statement about how the system attempts to address a wider practical, social, or policy issue, this task helps the project team determine if their goal is valid and if the target system is sufficient to achieve their goal. It can also support stakeholder engagement and project communication activities.

</td></tr>
<tr><td>Data Extraction (Or Procurement)</td><td>

Description

By ‘data extraction’ we refer to both the design of an experimental method or decisions about data gathering and collection, based on the planning and problem formulation from the previous steps, as well as the actual extraction and storage of novel data or the procurement of existing data.

Ethical Significance

The well-known principle of ‘garbage-in, garbage-out’ summarises the importance of this task nicely.

As data-driven technologies, ML algorithms or AI systems depend on the data fed into them. However, due diligence at this stage is important for reasons other than statistical validity. Responsible data extraction is, among other reasons, vital for the design of accountable and trustworthy services, the development of safe, fair, and explainable algorithms, and the deployment of sustainable and privacy-preserving systems.

</td></tr>
</table>

Project Lifecycle Stage	Information
Data Analysis	**Description** Data analysis is typically split into two types: *exploratory* and *confirmatory* analysis: • Exploratory data analysis allows analysts to better understand the structure and content of the dataset, and identify possible associations between data types and variables. • Confirmatory data analysis is where initial hypotheses that are developed in the previous stage are evaluated using a variety of statistical methods (e.g. significance testing). **Ethical Significance** In the context of responsible research and innovation, data analysis is vital to the assessment of myriad biases that can negatively impact a project, many of which are most obvious at this stage in a project. Identifying and dealing with missing data is particularly important during this task. Although upstream stakeholder engagement activities can help mitigate the impact of this bias, identifying the scope of its impact and determining how effectively it can be addressed (e.g. using various methods imputation, collecting additional data), will largely depend on the quality of the data analysis task.
Preprocessing and Feature Engineering	**Description** Whereas data analysis can give rise to valuable insights (e.g. business intelligence), the data structures are not always appropriate to train ML algorithms. Therefore,`preprocessing and feature engineering` is required to clean, normalise, or otherwise refactor data into the features that will be used in model training and testing, as well as the features that may be used in the final system. **Ethical Significance** Features are dependent upon, but separate from, the raw data that are collected in the prior stages. They can be engineered by hand or by using algorithmic techniques to improve the performance of subsequent ML processes. However, the features that are used in the process of model training, for instance, do not only affect the

Project Lifecycle Stage	Information
	model's accuracy or predictive power; they also impact the ethical consequences of the project (e.g. reducing the explanatory potential of system, creating discriminatory outcomes). Therefore, selecting the *best* features is a vital, albeit often time-consuming and complicated task that can involve trade-offs about which parameter to optimise for (e.g. predictive power versus interpretability).
Model Selection & Training	**Description** Simply put, this stage involves the selection of a particular algorithm (or multiple algorithms) for training a model. There are many factors that feed into this decision, including (but not limited to): • Access to computational resources (some learning algorithms require vast levels of computational power) • Predictive performance of model (as compared to other models) • Properties of underlying data (e.g. is the size of the dataset sufficient) **Ethical Significance** There are, of course, many technical and logistical reasons for the responsible selecting and training of a model (e.g. ensuring parsimony, optimising performance). However, a key concept in the responsible development of a model is the inherent interpretability and post hoc explainability of the model and the behaviour of the system into which it is implemented. Although there are nuances and exceptions, it is generally the case that more complex models are harder to interpret and explain (e.g. linear regression versus convolutional neural networks). Selecting the right technique, therefore, depends on the ultimate use case of the model and system.

Project Lifecycle Stage	Information
Model Testing & Validation	**Description** Model testing and validation involves the assessment of a model against a variety of metrics, which may include the evaluation of the model's accuracy as applied to novel data (held out from the original training data). **Ethical Significance** Where a dataset is split into testing and training (i.e. internal validation), or where a model's performance is evaluated against wholly new data (e.g. external validation from a separate trial or project team), there are options to assess more than just the model's performance. Testing the generalisability of a model to a new domain or context can also help ensure the model is both sustainable and fair (e.g. has similar levels of accuracy or performance when validated externally).
Model Documentation	**Description** This task involves the documentation of both formal and non-formal properties of the model and the processes by which it was developed. This includes (but is not limited to): • Data sources and summary statistics • Model used (e.g. proprietary model purchased from vendor) • Evaluation metrics (e.g. model performance) **Ethical Significance** Clear and accessible documentation is an important form of responsible project governance for the following reasons: • In research projects it ensures reproducibility and replicability of results, as well as other values associated with open research, such as public accessibility.

Project Lifecycle Stage	Information
	• In development projects it ensures accountability and transparency of decision-making. • In all projects it can help affected individuals seek redress for any harms that may arise from the design, development, or deployment of data-driven technologies.
System Design and Implementation	**Description** System design and implementation is the process of putting a model into production, and implementing the resulting system into an operational environment. The system enables and structures interaction with the model, within the environment (e.g. a recommender system that converts movie ratings into recommendations for future watches for a specific user). **Ethical Significance** Regardless of how well the preceding stages have gone, unless the encompassing system is implemented effectively, the model's performance will be impacted. Here, we can note the importance of two forms of implementation: • Technical implementation: designing and building the hardware and software infrastructure (e.g. server, interfaces) that will host the model. Among other things, it is important to ensure the technical system is secure, performant, and accessible. • Social or organisational implementation: how the technical system is situated within broader social and organisational practices is also important when considering the project's goals and objectives (e.g. appropriately informed users, complementarity with organisational practices).
User Training	**Description** 'User training' includes any form of support or upskilling that is offered and carried out by the individuals or groups who are required to operate a data-driven technology (perhaps in a safety–critical

Project Lifecycle Stage	Information
	context), or who are likely to use the system (e.g. consumers). **Ethical Significance** User training is rarely carried out by the same team members who designed and developed the system. While developers may produce documentation for the model (see above), this is often insufficient as a form of user training—additional forms of formal training workshops or courses may be required depending on the complexity of the system. Insufficient or inadequate training can create conditions in which cognitive biases such as algorithmic aversion or over-reliance thrive (e.g. users do not trust the performance or behaviours of a trustworthy algorithmic system, or users trust the outputs of an untrustworthy system).
System Use & Monitoring	**Description** Depending on how an AI system has been designed, its deployment and use in an environment (physical or virtual) can create conditions for ongoing feedback and learning (e.g. robotic systems that employ reinforcement learning, digital twins linked to a monitored counterpart). Regardless, the use of metrics and evaluative methods are commonly used to monitor the performance of a system and ensure that it retains (or ideally improves on) the same level of performance that it had when first validated. **Ethical Significance** The potentially dynamic (and sometimes unpredictable) behaviour of machine learning models and AI systems means that ongoing monitoring and feedback of the system, either automated or probed, is important to ensure that issues such as model drift have not affected performance or resulted in harms to individuals or groups.
Model Updating & De-commissioning	**Description** If the use and monitoring of a model or system identifies vulnerabilities or inadequate levels of performance, it may be necessary to update the model through retraining (i.e. looping back through

Project Lifecycle Stage	Information
	some of the model development tasks) or deprovision the system if it is no longer fit-for-purpose. Where the latter option occurs, the project team or organisation may need to commence a new project lifecycle to address any gaps in their business or organisation that arise because of the deprovisioning of the present system. **Ethical Significance** An algorithmic model that adapts its behaviour over time or context may require updating or deprovisioning (i.e. removing from the production environment). While this can include elements such as improvements to the system's architecture (e.g. for speed or security), the more important component here is the model itself (e.g. the model parameters, the features used). An important issue to address is model drift, which can arise because of changes to the underlying data distribution used to train the original model (e.g. average values of house prices in a fluctuating property market) or because the semantic meaning of the features has changed due to shifting societal practices or norms (e.g. political or geographic boundaries).