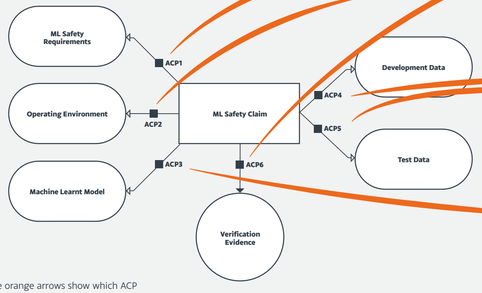
Assurance of Machine Learning for use in Autonomous Systems

Pattern

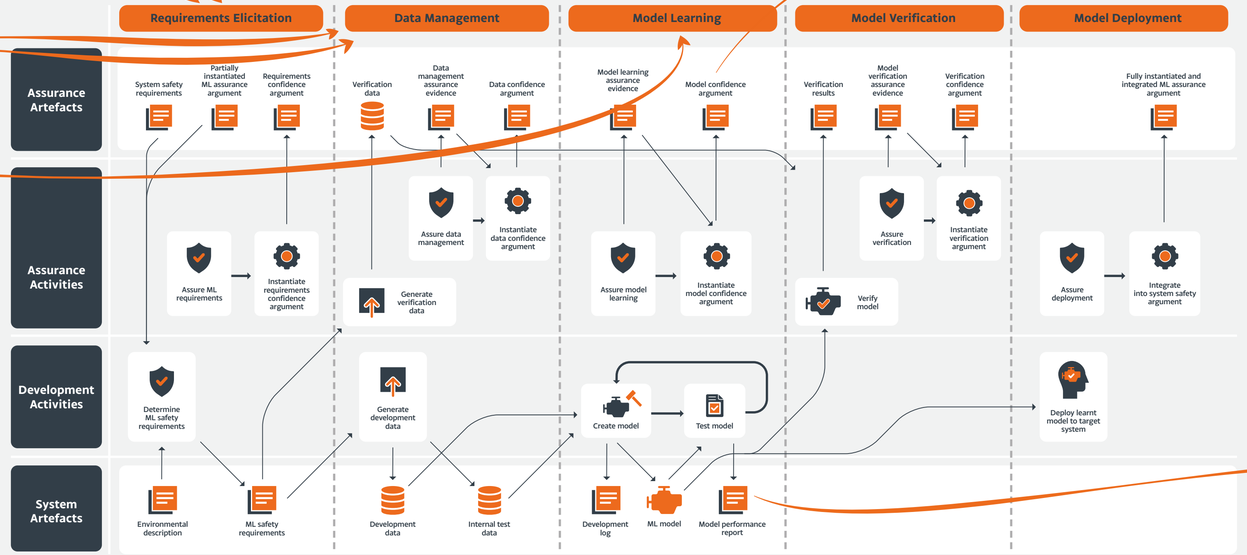
The confidence arguments for the ACPs shown on the pattern diagram are developed using the AMLAS process (shown on the right).



The orange arrows show which ACP relates to which phase of the machine learning lifecycle in the process.

An example of the outcome of this, for ACP3, is shown on the right hand page.

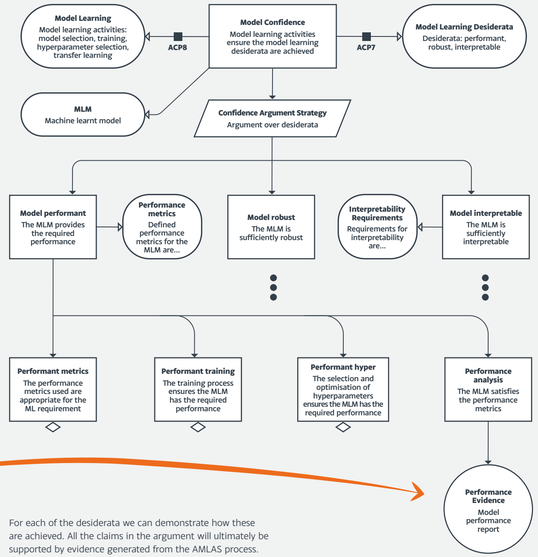
The AMLAS process



An example

Our patterns give the structure of the confidence arguments, based around the process activities and sets of identified desiderata (desired properties). The diagram below shows the model confidence argument that would be used to support assurance claim point 3. For model learning, the desiderata we identified were:

- **Performant** – the MLM should provide the required performance
- **Robust** – the MLM should perform well in circumstances where the inputs encountered at run-time are different to those present in the training data
- **Interpretable** – the MLM should be able to produce artefacts that support the analysis of its output, and thus any decision based on it.



For each of the desiderata we can demonstrate how these are achieved. All the claims in the argument will ultimately be supported by evidence generated from the AMLAS process.