

Proposal for a PhD research project

## **The notion of malfunction in engineering systems**

A technical artefact typically malfunctions if it does not satisfy the functional requirements laid out for it whereas it was claimed to meet these requirements (after delivery to the client) or just expected to meet them (for instance in the testing phase). Since a particular artefact performs its function through one or other physical mechanism characterizing this artefact, if an artefact malfunctions it does so by behaving physically in an aberrant way. Failure to meet the functional requirements without such physical aberrant behaviour will than not be malfunctioning but will be caused by a deviation in the circumstances in which the artefact operates or is being used. It is commonly understood that the functional requirements for an artefact fully specify the circumstances in which it will be required to perform its function.

This is the ideal story. It supposes that an exhaustive list of functional requirements, including an exhaustive list of the required conditions, was available during the design of the artefact. Any incompleteness in this lists opens the possibility of questioning judgements of malfunction. For traditional hands-on human-sized artefacts it may already be questioned whether it is even possible to come up with an exhaustive list of functional requirements and, in particular, with an exhaustive list of conditions that must be met, both regarding the environmental circumstances and the capabilities of the user(s) of the artefact, such that it is determinate whether or not an artefact is functioning properly. However, in modern technology we are dealing to an ever increasing extent with large-scale complex system-like artefacts that were never designed 'from scratch' to meet a list of functional requirements, but are rather continuously (re)designed at the level of their components. Additionally, we have complex artefacts, for example aeroplanes, the many parts of which are regularly (re)designed but very and which also perform their function within these large-scale systems, such that their environment is not fixed but the temporary outcome of a continuous process of partial (re)design. This situation makes it problematic to decide when an engineering system is malfunctioning, or which of the many things that go wrong within a technology can be termed as a malfunction, and what exactly it then is that is malfunctioning. A clarification of how the notion of malfunction is and can be used in modern technology is highly desirable. This is particularly so because of the risks that malfunctioning engineering system pose to people and goods. This situation not only poses ethical requirements to engineers involved in the design, operation and maintenance of technical artefacts and systems, but the legal apparatus that regulates the potential or actual harmful consequences of technology also directly affects engineering practice.

Possible research questions:

(A) Empirical:

1. How do engineers actually use the notion of malfunction? How is this use differ in other professions that touch upon technology (e.g. law)?

(B) Conceptual:

2. What (conceptual) limits are there to the judgement of malfunction in technical artefacts in general?

3. How can judgements of malfunction be grounded in engineering systems for which no defining list of functional requirements exist?