doi: 10.1017/S143192760909309X
Microsc Microanal 15(Suppl 2), 2009

Applying Engineering Information Management Principles to Microscopy

N.H.M. Caldwell,* C.A. McMahon,** M.J. Darlington,** P. Heisig,* D.M. Holburn* and P.J. Clarkson*

* Department of Engineering, University of Cambridge, Trumpington Street, Cambridge, CB2 1PZ, United Kingdom
** Department of Mechanical Engineering, University of Bath, Claverton Down, Bath, BA2 7AY, United Kingdom

In the engineering industry and elsewhere, corporate business models are shifting away from product delivery paradigms to product-service paradigms, where ongoing through-life service support is preeminent. In the UK, a Grand Challenge project (www.kimproject.org) has been established to investigate and develop good practice with regard to the design and use of information, knowledge-support systems, and business methods necessary for organisations to succeed and profit within this new environment [1]. One outcome of this is a series of principles for through-life management of engineering information, derived from the research insights and experiences of project members in both academia and industry, and supported by detailed explanations and implementation guidance.

Although the Principles are not yet in their final form, there are clear parallels between the requirements for effective through-life management of engineering information (as created in the design of products and services) and scientific information, such as the micrographs, X-ray spectra, and other analyses produced by microscopists. Both fields involve the creation, recording and retention of substantial quantities of information, which is used and reused for different purposes over potentially long periods of time. Industrial applications of microscopy, e.g. metrology, material analysis, semiconductor inspection, nanotechnology, etc., directly generate engineering information.

There are eleven Principles, all of which can be mapped to guidance for individual microscopists, organisations employing microscopy, and/or instrument and software tool manufacturers. Some examples of how the principles can be applied to microscopy now follow:

1. *The Principle of Parsimony: Create, record and retain information only if necessary.*
   Microscopists should record only the images and other data that they need. Beam damage associated with optimal imaging, vacuum damage, and sample degradation may mean that there is only one opportunity to acquire best-quality images from a particular specimen location. Using sample mapping routines provided in microscopy software may allow return to "virtually" marked areas of interest. Capturing every possible image is neither desirable nor feasible.
2. *The Principle of Granularity: Record information in a storable information object at a granularity appropriate for use and reuse.*
   The fundamental information object in microscopy is the micrograph with its associated annotation (scale marker, instrument parameters, etc.). Even when performing serial sectioning and 3-d reconstruction, the individual image slices should be the smallest chunk of information recorded. The reconstructed representation is itself an aggregate information object.
3. *The Principle of Identity: Give an information object a unique and persistent identifier.*
   Micrographs (and other analyses) are routinely stored as electronic data files. They need identifiers that will enable them to be uniquely "named" – formal file naming systems should use contextualising characteristics, e.g. microscopist's username, timestamp, an abbreviation of the

specimen's nature, type of analysis, etc., to make unique identifiers human-readable. Identifiers should persist, thus archiving datasets to new storage locations should not render them invalid.

4. *The Principle of Uniqueness: Create an information entity once only and explicitly reference it everywhere else.*
   Data should not be unnecessarily duplicated; in an ideal world, original data would be reprocessed at need rather than each type of post-processing creating another variant dataset.
5. *The Principle of Usability: Design an information entity explicitly to achieve its intended goals.*
   For instance, micrographs, spectra, etc. must be fit for purpose. They must be annotated with instrumental parameters, descriptions of processing performed, and compliance with regulations. Software should employ pull-down menus and contextualised autocompletion of template forms to make this process as streamlined and efficient as possible.
6. *The Principle of Reusability: Design an information entity explicitly to maximise its potential for reuse wherever appropriate.*
   The associated annotation of data must be searchable and provide sufficient context information that others can use and assess the data. In micrographs, overlaid annotation should be separable from the image itself, so that post-processing routines do not have to handle micron markers etc.
7. *The Principle of Evaluation: Assess and assign the value of an information object throughout its life from creation to disposal.*
   The value of data to an individual or an organisation is not constant. Choosing which data to keep and which to discard requires reflection on data usage patterns and potential needs.
8. *The Principle of Portability: Create an information entity and its annotations systematically using representations supporting perpetual reuse.*
   Using proprietary formats to encode data leads to vendor lock in and the risk that the data will become unreadable in the future. Representing data using open-source standards (such as XML) ensures that data will remain (machine and human) readable and reusable forever.
9. *The Principle of Robustness: Use robust methods to capture, create and manipulate information entities.*
   Data recording procedures should be unambiguous. Software should support the process, automatically generating annotation and identifiers, reducing administrative burdens on people.
10. *The Principle of Discovery: Actively employ the information repository as a resource for learning and discovery.*
    Existing data sets provide obvious opportunities for exemplars for teaching and publication. New algorithms applied to pristine original data may generate new insights.
11. *The Principle of Design: Design all aspects of information management to satisfy the organisation's current and future needs.*
    The organisation must develop strategies to the creation, annotation, storage, retrieval, manipulation, dissemination and obsolescence of microscopical information, not just one aspect.

The authors welcome feedback on the Principles for both engineering information and microscopy.

References

[1] A. Ball et al., Int'l. Jour. Digital Curation 1 (2006) 56.
[2] The KIM project is funded by the UK Engineering and Physical Sciences Research Council (EPSRC) and the Economic and Social Research Council (ESRC) under Grant Numbers EP/C534220/1 and RES-331-27-0006. The contribution of the KIM project team to developing the principles for through-life management of engineering information is gratefully acknowledged.

https://doi.org/10.1017/S143192760909309X Published online by Cambridge University Press