Accelerating Cancer Research Using Semantics-Driven Technology

James Brenton, Jim Davies, Jeremy Gibbons, Steve Harris
MSES, Dec 2008
1. Phase III clinical trials

- randomized controlled studies to compare with existing treatments
- individual trials tend to lack diagnostic power
- need *meta-analysis* to combine results
- requires semantic metadata for integration
- *better science*

- software support for clinical trials (forms, services, databases…)  
- similar but different instances: a *software product line*
- custom software, but without hand-crafted implementations
- *less drudgery*
2. CancerGrid

- *metadata registries* for managing semantics
- *metamodel* of trial protocols, with tools for authoring protocols
- *model-driven generation* of software artefacts
- semantic information preserved throughout processing
- *breast cancer* clinical trials as exemplars

- UK Medical Research Council and Microsoft Research, 2005–2008
3. Semantics-driven technology

- **terminology services**
  collections of defined terms, and their relationships

- **metadata registries**
  observations, measurements, properties:
  terms, intended purposes, possible values

- **model repositories**
  schemas, services, forms, queries, reports;
  linked to metadata elements, propagating semantics

- "**semantic frameworks**"

- other application domains (eGov, logistics, libraries...)
4. Application to early-phase studies

- exploratory studies — from validation to discovery: first human experiments, small populations
- molecular and functional imaging to better understand the mechanisms of drug action
- four specific challenges arise:
  - metamodelling *marker studies* data
  - *federation* of metadata registries
  - semantic *integrity constraints*
  - *analytical techniques* and *tissue management*
- funded by Microsoft Research, 2008–2010
- hoping to accelerate clinically relevant discoveries in cancer care