

Knowledge-driven drug development: the quest for a semantic repository of clinical studies

Kaushal Desai, AstraZeneca Pharmaceuticals

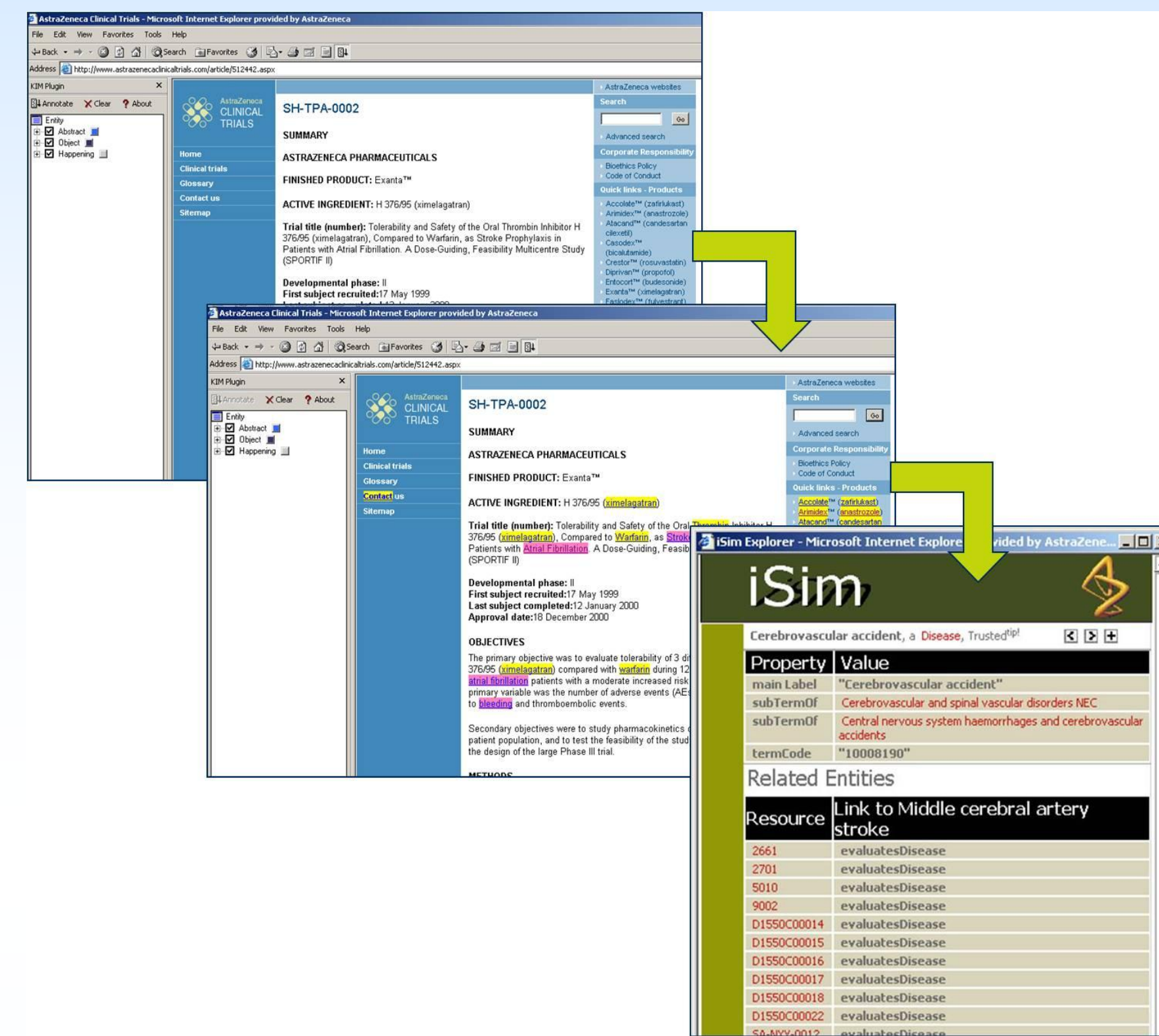
SCIENTIFIC RATIONALE

- Knowledge exploitation drives innovation in drug development.
- Knowledge generated in clinical studies is often stored in disconnected 'silos' in structured databases as well as 'unstructured' study documents.
- Challenges associated with **access and re-use** of knowledge therefore, limit cross-compound, cross-study information exploitation.
- A semantic repository of clinical studies could provide the capability to –
 - Respond efficiently to regulatory queries and avoid excessive, repetitive integration effort .
 - Enable information exploitation through greater control over and utilization of organization knowledge.
 - Deliver automated reasoning capability to proactively identify optimal study designs based on successful historic outcomes.

SEMANTIC ANNOTATION OF CLINICAL STUDIES

Deliverables

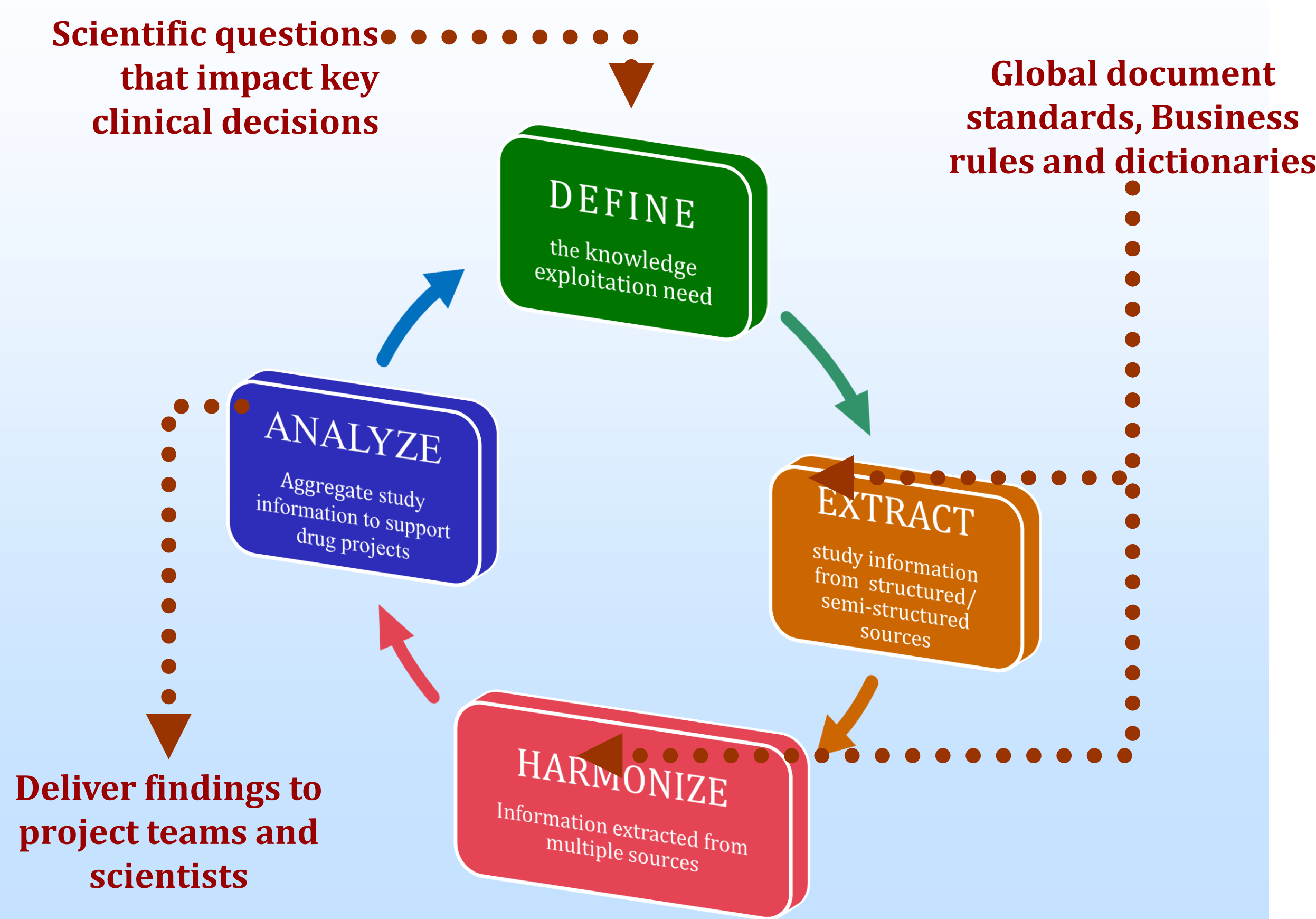
- The ability to search clinical studies across multiple underlying databases and document stores, based on **planning** (study objectives, study design etc.), **execution** (milestone dates, events, number of subjects etc.) and **outcome** (adverse events, efficacy, PK results etc.) related criteria.



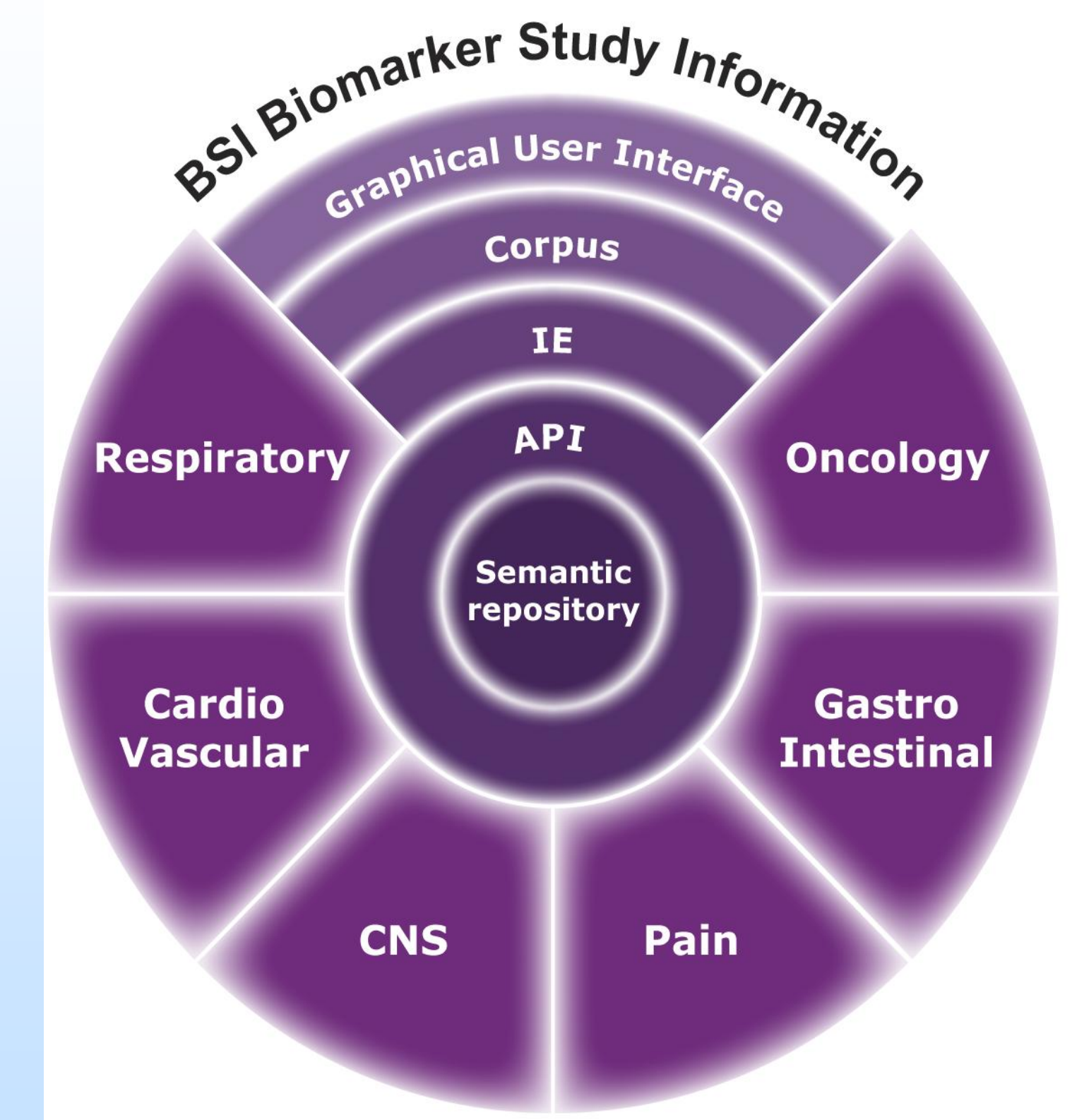
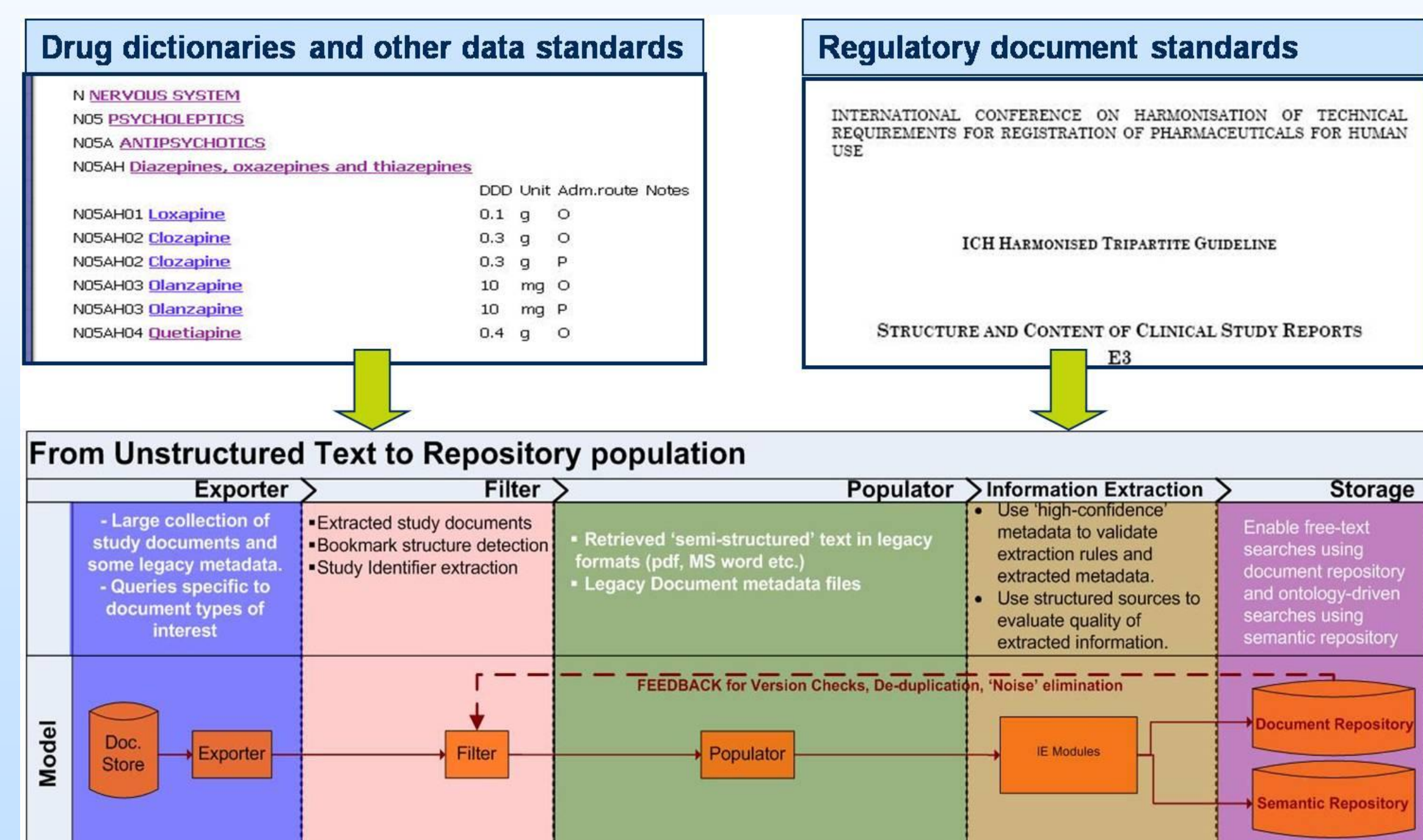
BENEFITS FROM THE SEMANTIC APPROACH

- **How well do semantic tools enable the utilization of information and knowledge ?**
 - Within 1.5 years of deployment, the information integration benefits resulted in more than 7 years of work savings on various drug project and information exploitation teams.
 - Benefits from reasoning on clinical study information are yet to be realized.
- **What are the measurable benefits of using an RDF/OWL framework ?**
 - Abstraction from drug-related 'facts' to drug-classes and abstraction from adverse event-related 'facts' to System organ class would be cumbersome in case of a 'relational' architecture.
 - Integrated search on studies in a global organization without forcing a single 'static' information model on source systems would be difficult with a 'relational' solution.
 - Future scalability and addition of 'new knowledge' would be difficult without use of uniformly accepted standards.

CONCEPTUAL SOLUTION



RDF/OWL FRAMEWORK FOR INTEGRATION



REFERENCES

- Kaushal Desai, Borislav Popov. *Semantic annotation of clinical studies for knowledge-driven drug development*. European Semantic Technology conference, Austria, Vienna 2008.
- McCallum A., *Information Extraction: Distilling Structured data from unstructured text*. ACM Queue, November 2005.
- Gardner SP. *Ontologies and semantic data integration*. DDT, July 2005.

ACKNOWLEDGEMENTS

- **Business Team:** Kaushal Desai (Project lead), James Weatherall, Bosse Andersson.
- **IS Collaborators:** Elisabet Soderheim (project manager), Frederik Lundbeck (technical lead), Bob Keesler, Tobias Heed, Mattias Almen).

External Partner

- **Ontotext Labs (Sirma Group)**
KIM (Knowledge and Information Management) platform is a software platform for automatic ontology population and open-domain dynamic semantic annotation of unstructured and semi-structured content for Semantic Web and knowledge management applications.