

An Ontology for Biostatistics



KNUT M. WITTKOWSKI

The Rockefeller University Hospital
Center for Clinical and Translational Science
New York, NY 10021, kmw@rockefeller.edu

NIH Data Sharing Policy

"To ensure that others can use" data, **meta-data** is needed that documents "

- definitions of variables, details about codes
- the methodology and procedures used to collect the data,
- variable field locations, [...].



The precise content of **documentation** will vary by the

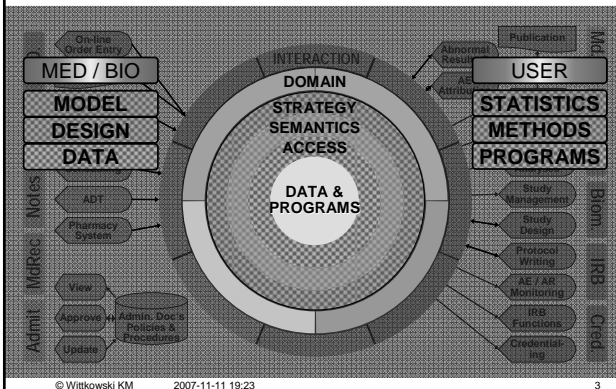
- scientific area,
- type of data collected,
- study design, and
- characteristics of the dataset."

http://grants.nih.gov/grants/policy/data_sharing/data_sharing_guidance.htm

© Wittkowski KM 2007-11-11 19:23

2

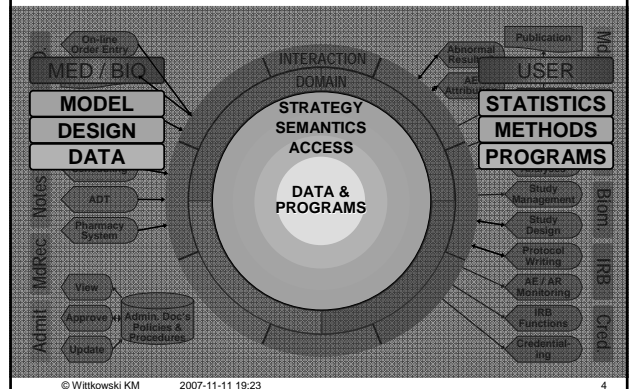
Layers of Meta-Data



© Wittkowski KM 2007-11-11 19:23

3

Layers of Biostatistics Meta-Data



© Wittkowski KM 2007-11-11 19:23

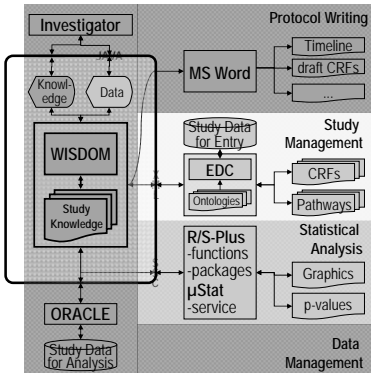
4

WISDOM for C&T Design

Goal:

to make knowledge about the design of the study available to the software an investigator uses for

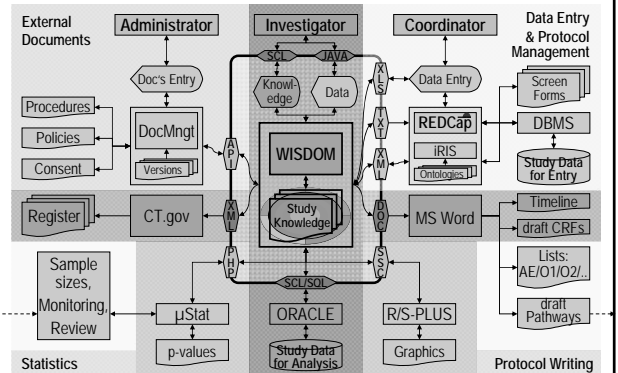
- Protocol Writing
- Study Management
- Data Management
- Statistical Analysis



© Wittkowski KM 2007-11-11 19:23

5

WISDOM as a Knowledge Broker



© Wittkowski KM 2007-11-11 19:23

6

WISDOM Knowledge Acquisition

© Wittkowski KM 2007-11-11 19:23 7

WISDOM: Data Base Creation

© Wittkowski KM 2007-11-11 19:23 8

WISDOM for REDCap

© Wittkowski KM 2007-11-11 19:23 9

WISDOM as a μStat User Interface

© Wittkowski KM 2007-11-11 19:23 10

NIH Data Sharing Policy

"To ensure that others can use" data, meta-data is needed that documents

- "definitions of variables, details about codes"
- "the methodology and procedures used to collect the data,"
- "variable field locations, [...]."

The precise content of only the DOMAIN documentation will vary by the

- scientific area,
- type of data collected,
- study design, and
- characteristics of the dataset."

MODEL
DESIGN
DATA
DOMAIN
MODEL
DESIGN
DATA

http://grants.nih.gov/grants/policy/data_sharing/data_sharing_guidance.htm

© Wittkowski KM 2007-11-11 19:23 11

Summary / Conclusions

For WISDOM, e.g., to assist clinical/translational scientists with

- Experimental Design
- Protocol Writing
- Data Management (Oracle)
- Data Acquisition (REDCap)
- Safety Monitoring
- Statistical Analyses (μStat)
- Data Sharing

A consensus ontology has to be developed for representing biostatistics knowledge at the layers

MODEL
DESIGN
DATA

© Wittkowski KM 2007-11-11 19:23 12