The CERA-2 Meta Database and Needs for a Common Information Structure

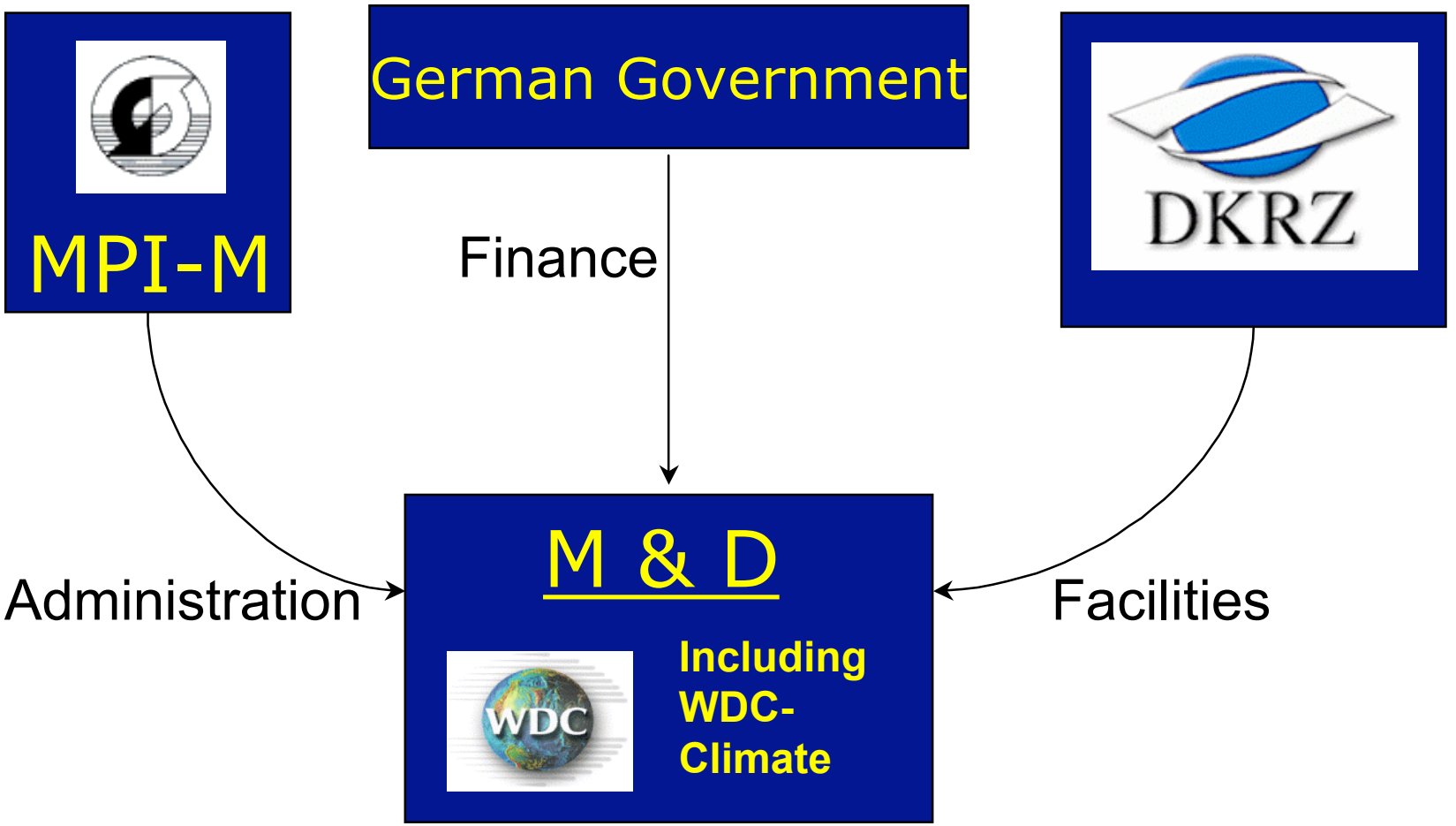
Frank Toussaint, Michael Lautenschlager
WDC-Climate, M&D, MPI for Meteorology

Berlin 14. / 15. October 2008

Contents

- World Data Center for Climate (WDCC) at MPI Meteorology
- WDCC: Hardware, Software & Database
- WDCC: The Metadata and Data Model
- Metadata on Data – discovery MD, use MD, provenance MD
- METAFOR
  - a possible MD structure
  - automatic data collection

World Data Center Climate at MPI for Meteorology

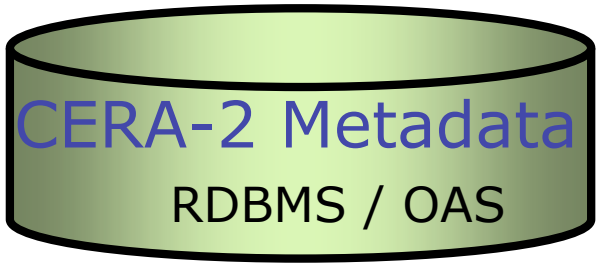


Partners, e.g.:

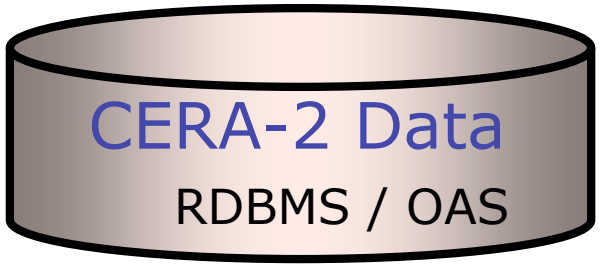
British Atmospheric Data Centre (BADC)
German Weather Service (DWD)
MetOffice
National Center for Atmospheric Research (NCAR)

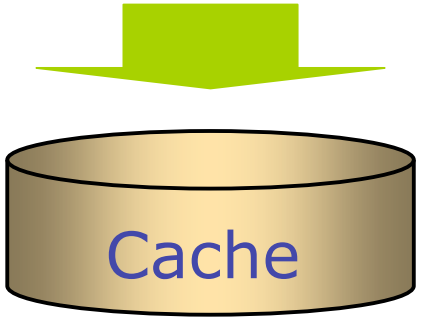
WDCC: Hardware, Software Structure & Database

1. Metadata level:



2. Data level:





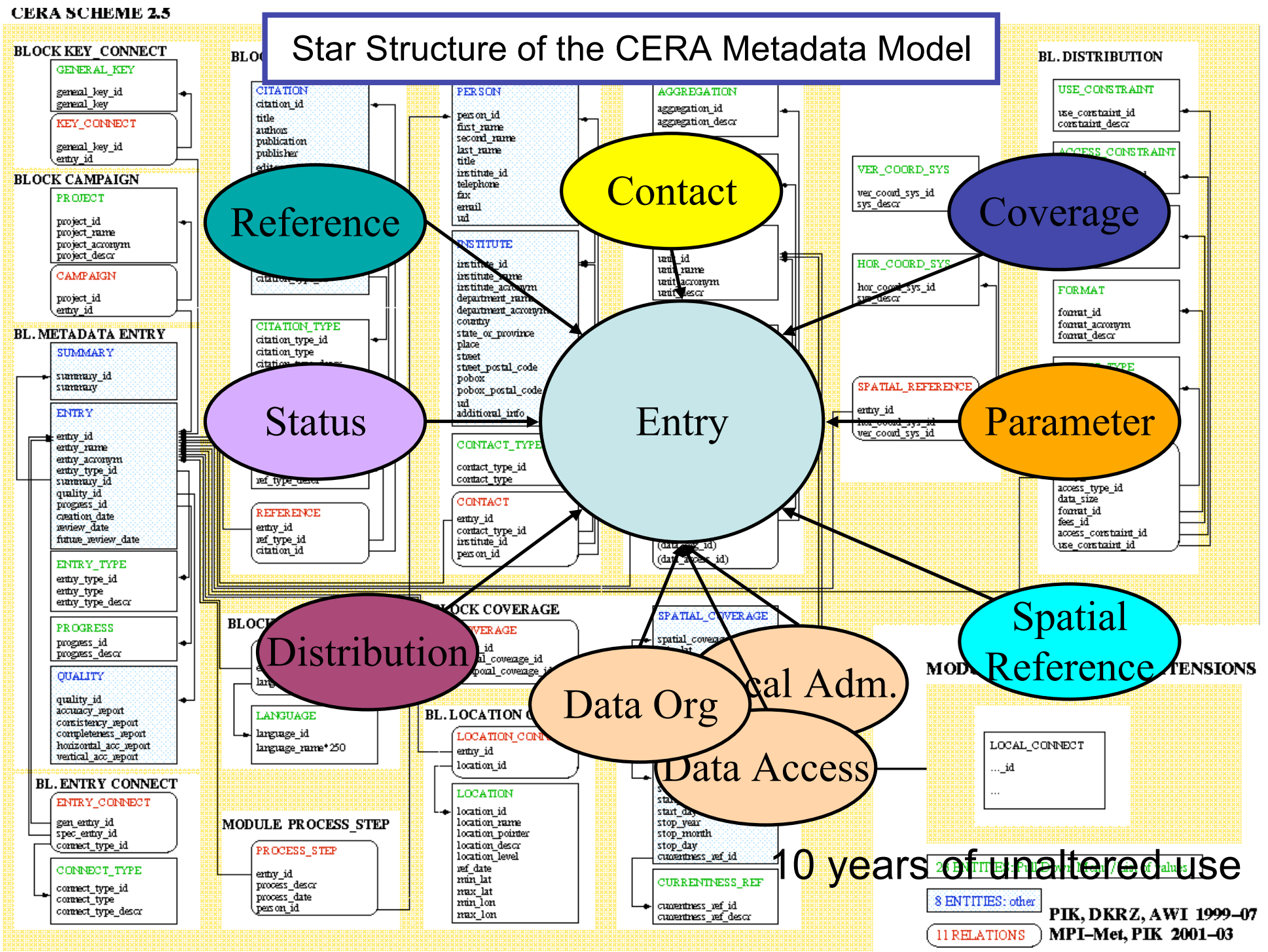
370 Terabytes of 5 Petabytes tape data are web accessible

Users:

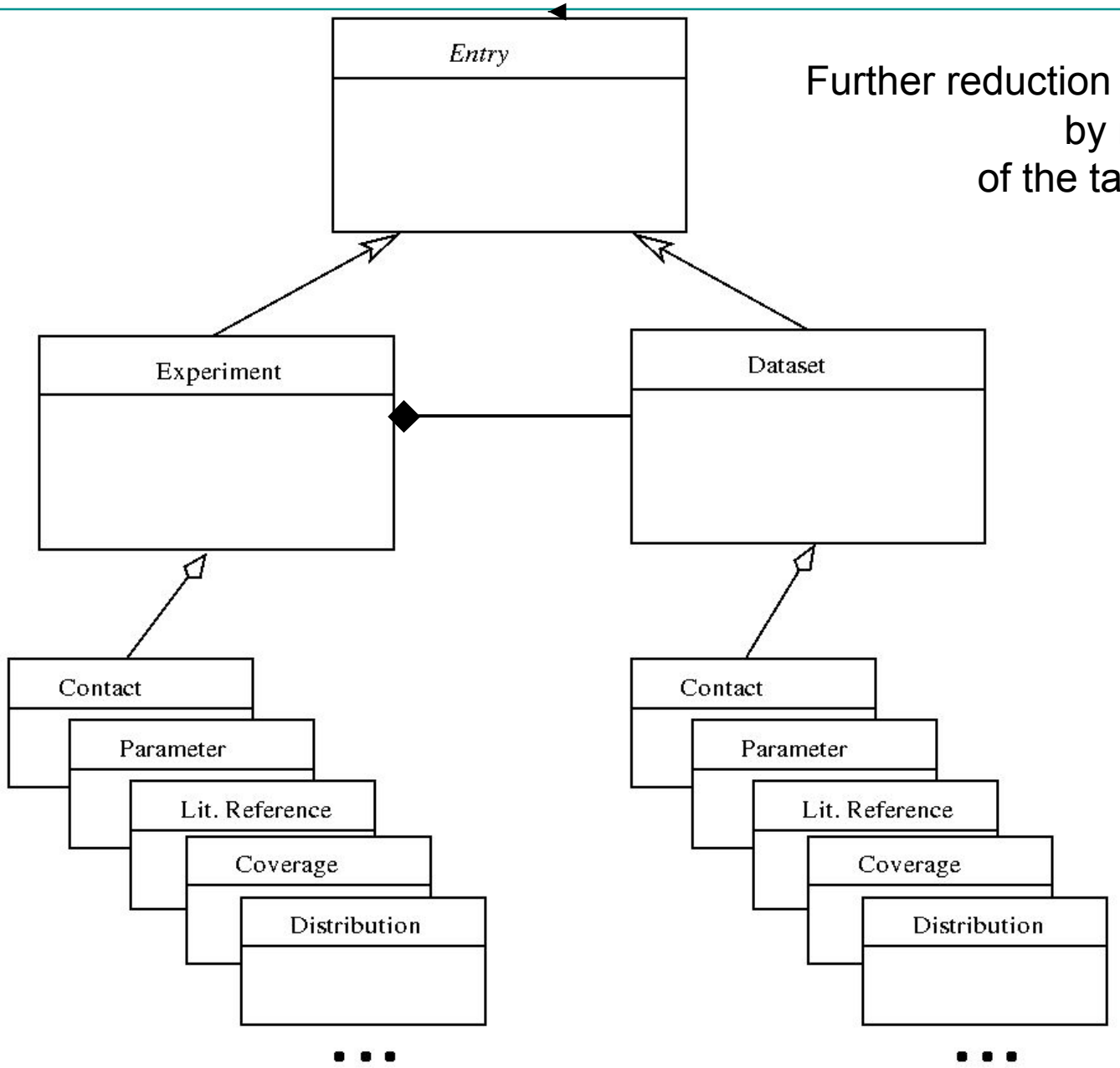
- experienced climate modelling and analyzing / earth system research (interdisciplinary)
- national / international

Software:

- GUI (Java architecture)
- Middleware (Java)
- Oracle RDBMS, BLOBs


CERA SCHEME 2.5
Star Structure of the CERA Metadata Model
BLOCK KEY_CONNECT
GENERAL_KEY
general_key_id
general_key
KEY_CONNECT
general_key_id
entry_id
BLOCK CAMPAIGN
PROJECT
project_id
project_name
project_acronym
project_descr
CAMPAIGN
project_id
entry_id
BL. METADATA ENTRY
SUMMARY
summary_id
summary
ENTRY
entry_id
entry_name
entry_acronym
entry_type_id
summary_id
quality_id
progress_id
creation_date
review_date
future_review_date
ENTRY_TYPE
entry_type_id
entry_type
entry_type_descr
PROGRESS
progress_id
progress_descr
QUALITY
quality_id
accuracy_report
consistency_report
completeness_report
horizontal_acc_report
vertical_acc_report
BL. ENTRY CONNECT
ENTRY_CONNECT
gen_entry_id
spec_entry_id
connect_type_id
CONNECT_TYPE
connect_type_id
connect_type
connect_type_descr
CITATION
citation_id
title
authors
publication
publisher
CITATION_TYPE
citation_type_id
citation_type
REFERENCE
entry_id
ref_type_id
citation_id
PERSON
person_id
first_name
second_name
last_name
title
institute_id
telephone
fax
email
ud
INSTITUTE
institute_id
institute_name
institute_acronym
department_name
department_acronym
country
state_or_province
place
street
street_postal_code
pobox
pobox_postal_code
ud
additional_info
CONTACT_TYPE
contact_type_id
contact_type
CONTACT
entry_id
contact_type_id
institute_id
person_id
AGGREGATION
aggregation_id
aggregation_descr
BL. DISTRIBUTION
USE_CONSTRAINT
use_constraint_id
constraint_descr
ACCESS_CONSTRAINT
FORMAT
format_id
format_acronym
format_descr
access_type_id
data_size
format_id
fees_id
access_constraint_id
use_constraint_id
VER_COORD_SYS
ver_coord_sys_id
sys_descr
HOR_COORD_SYS
hor_coord_sys_id
SPATIAL_REFERENCE
entry_id
ver_coord_sys_id
Reference
Contact
Coverage
Status
Entry
Parameter
Distribution
Data Org
cal Adm.
Data Access
Spatial Reference
BLOCK COVERAGE
SPATIAL_COVERAGE
LANGUAGE
language_id
language_name*250
BL. LOCATION
LOCATION_CONN
entry_id
location_id
LOCATION
location_id
location_name
location_pointer
location_descr
location_level
ref_date
min_lat
max_lat
min_lon
max_lon
MODULE PROCESS_STEP
PROCESS_STEP
entry_id
process_descr
process_date
person_id
start_year
stop_year
stop_month
stop_day
currentness_ref_id
CURRENTNESS_REF
currentness_ref_id
currentness_ref_descr
MODULE EXTENSIONS
LOCAL_CONNECT
..._id
...
10 years of unaltered use
8 ENTITIES: other
11 RELATIONS
PIK, DKRZ, AWI 1999–07
MPI–Met, PIK 2001–03

Cera-2 Objects: Experiment & Dataset



Further reduction of structures
by multiple use
of the table structure

WDCC: Need for Documentation – the Way to METAFOR

- different levels of MD:
  - use MD
  - discovery MD
  - provenance MD

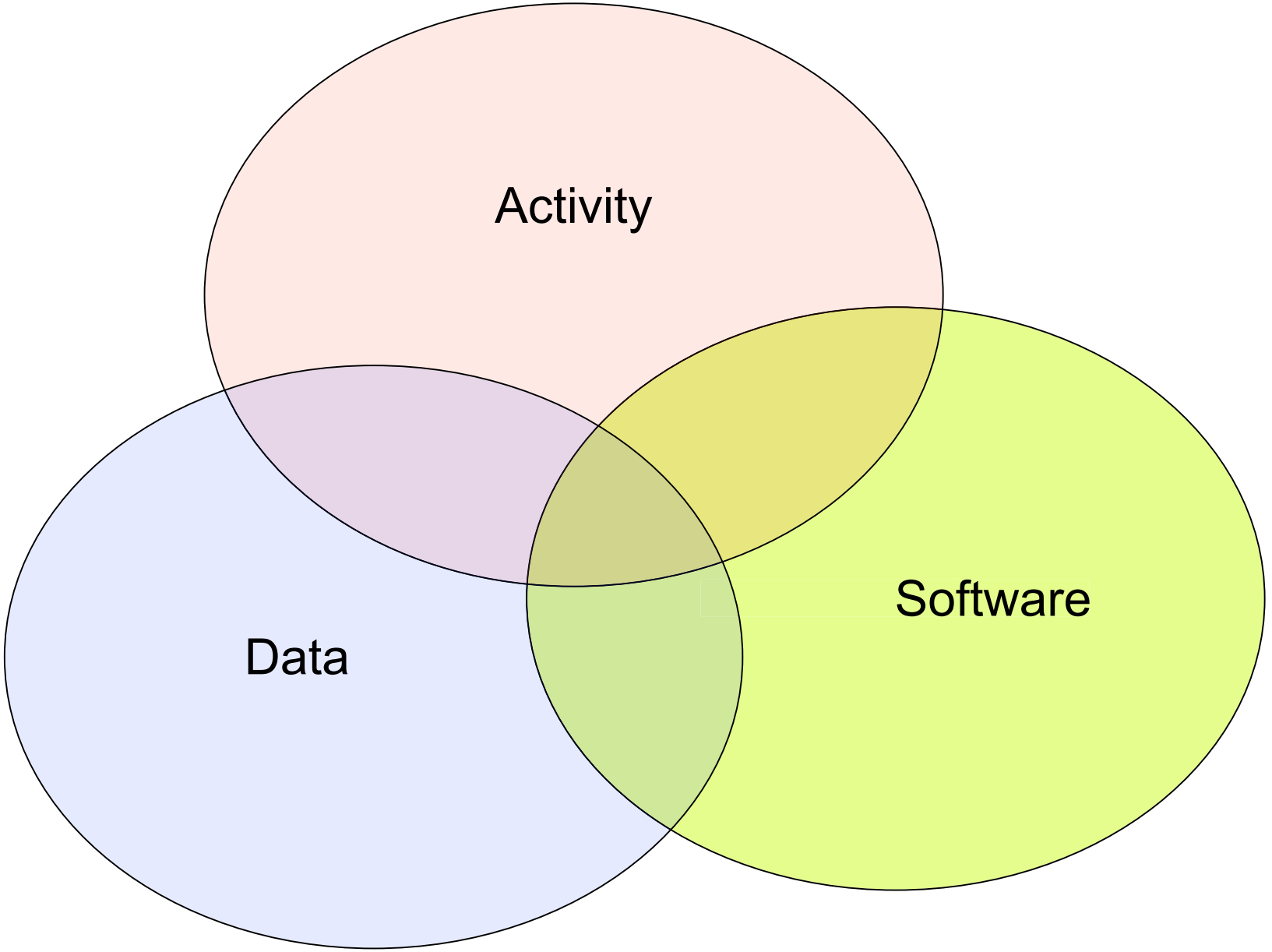
- There were requests for provenance MD, but there was little interest in filling in forms.

- It was not before a problem came up, that it came to a project on MD of SW, data & the connecting activity:

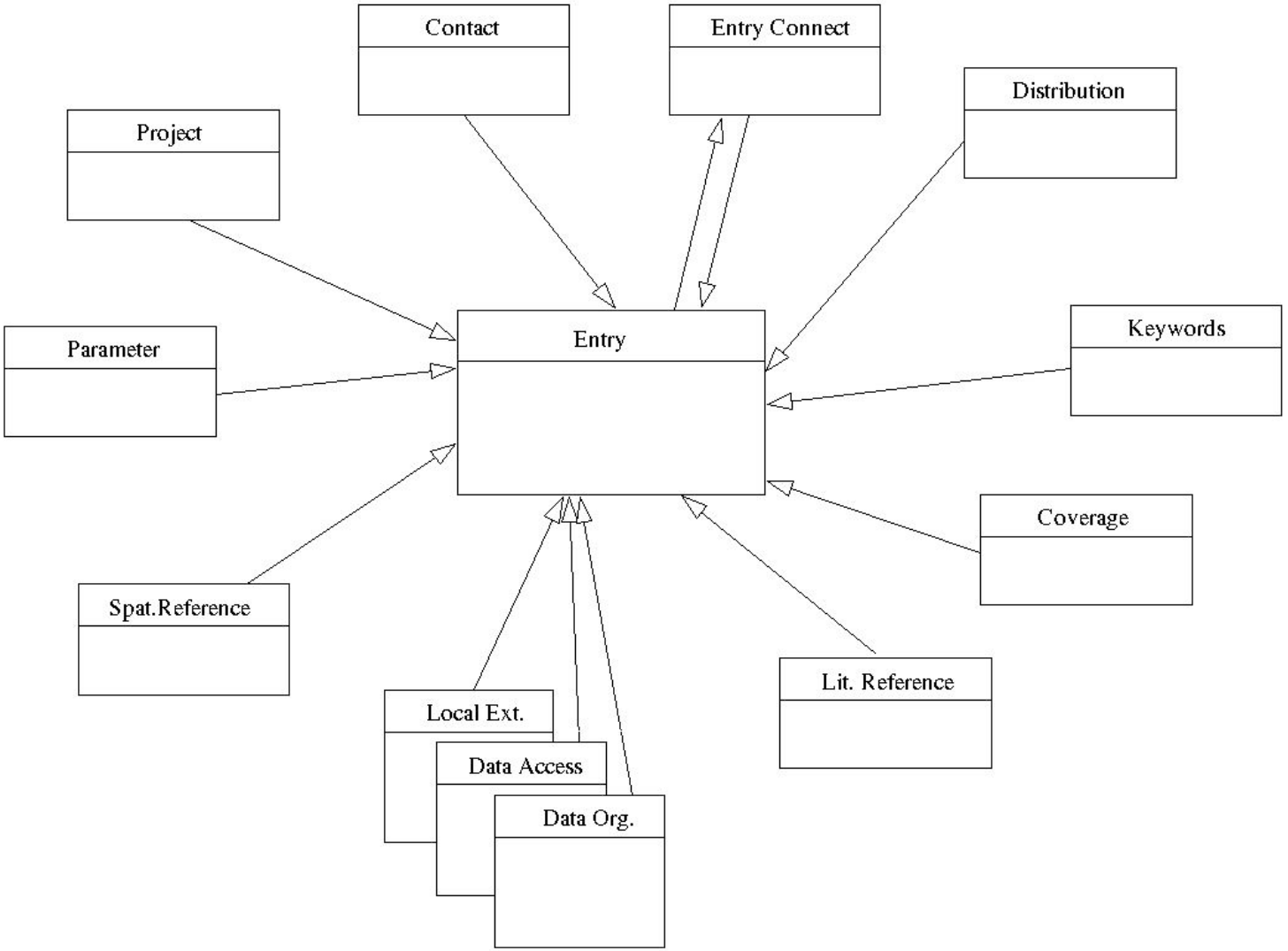
- This problem was a client's question:
  To which value was a certain parameter set
  during last year's model run ???

  ➔ Need for more comprehensive MD in a
  Common Information Model (CIM)

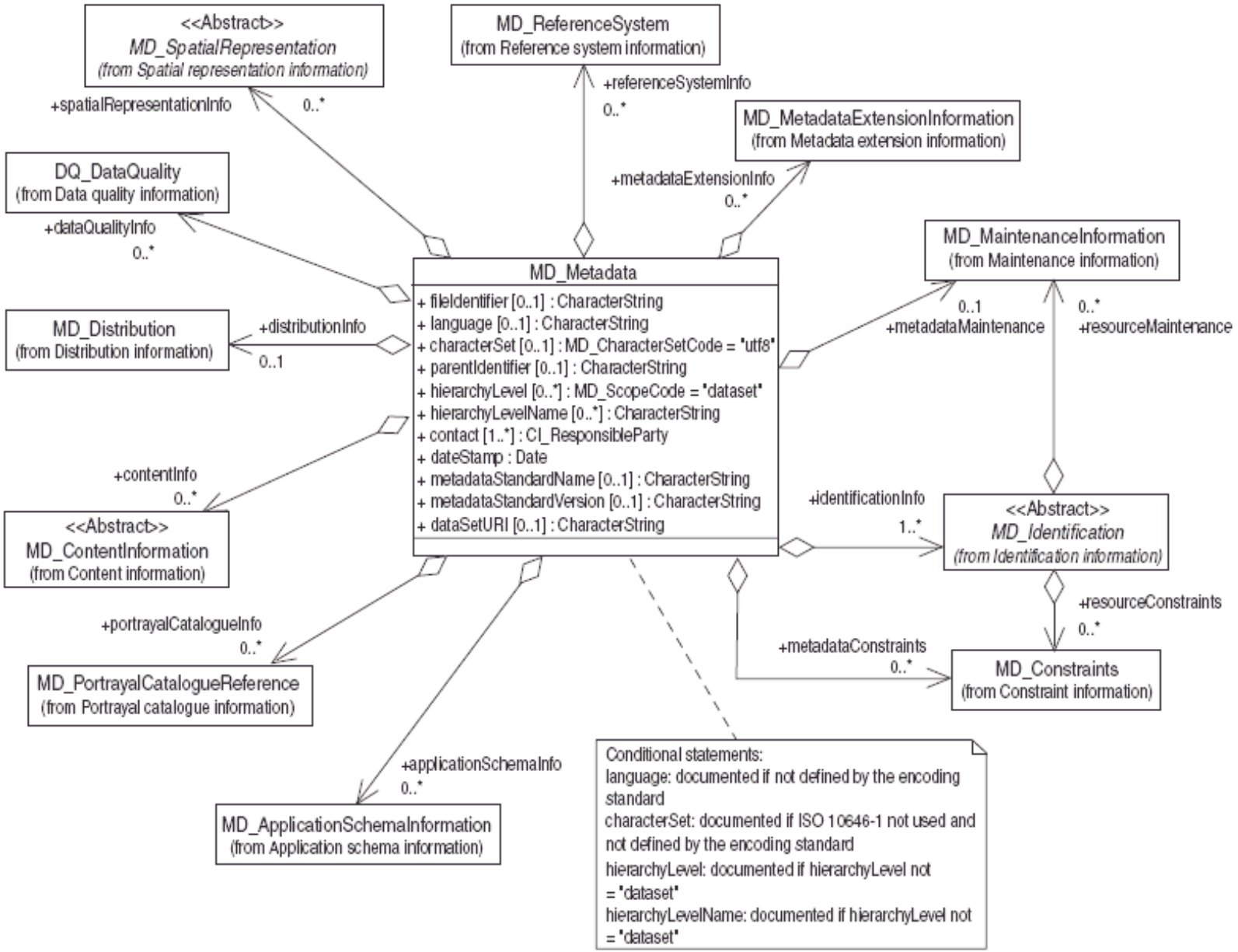
METAFOR: The Three Main MD Areas of the CIM



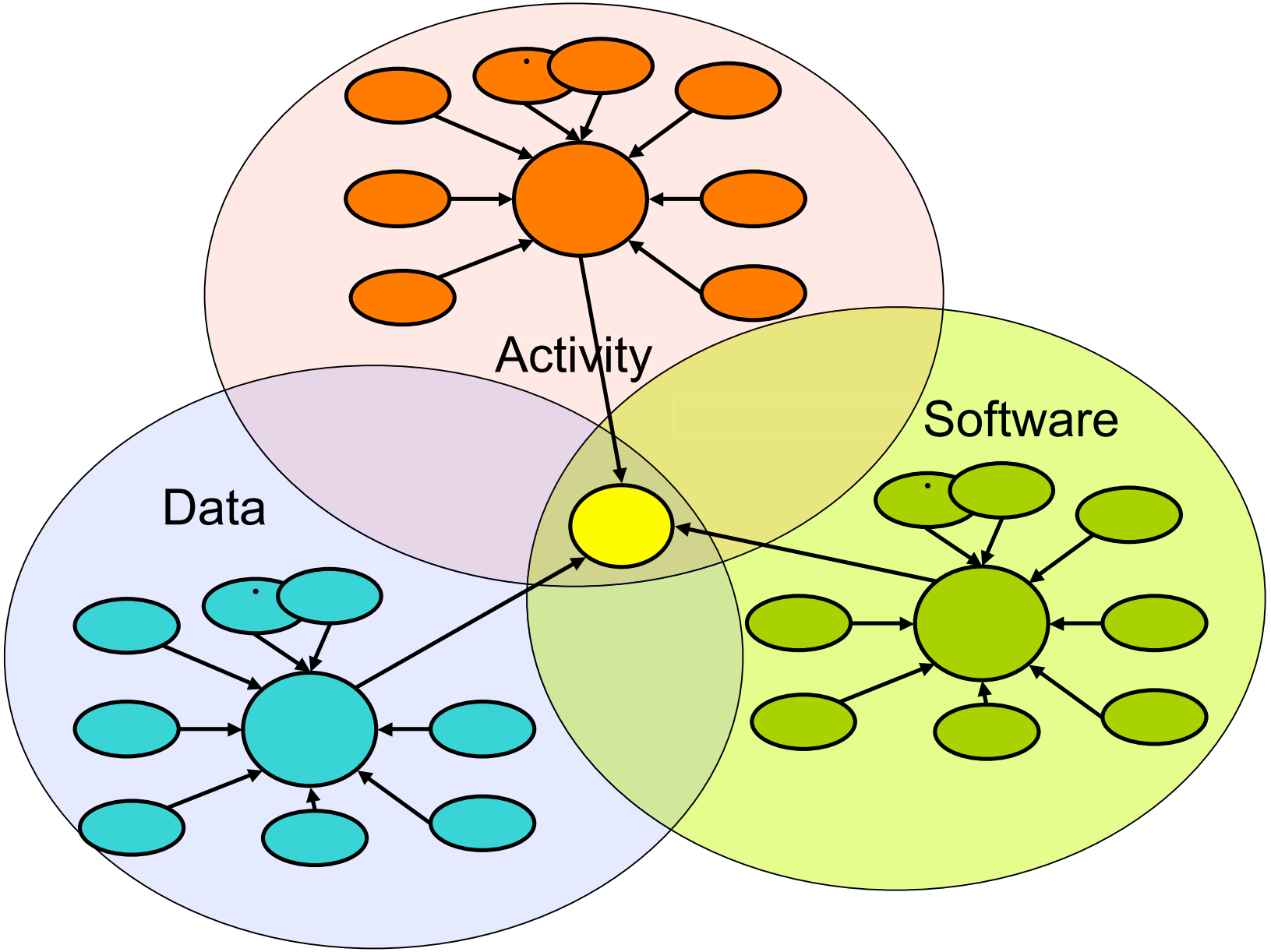
Cera-2 DB Model

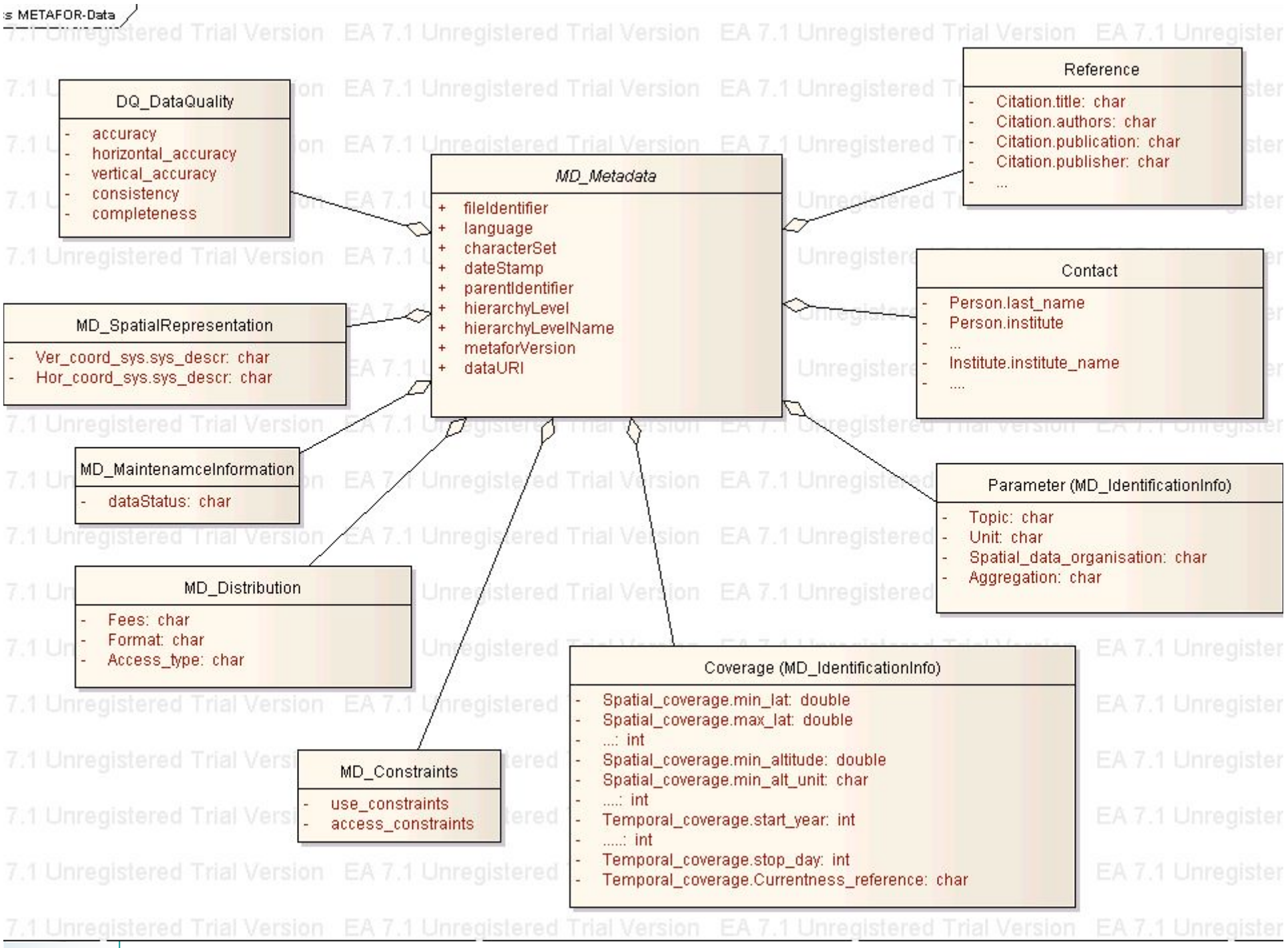


Star structure of the C3-Grid ISO 19115 Data Model

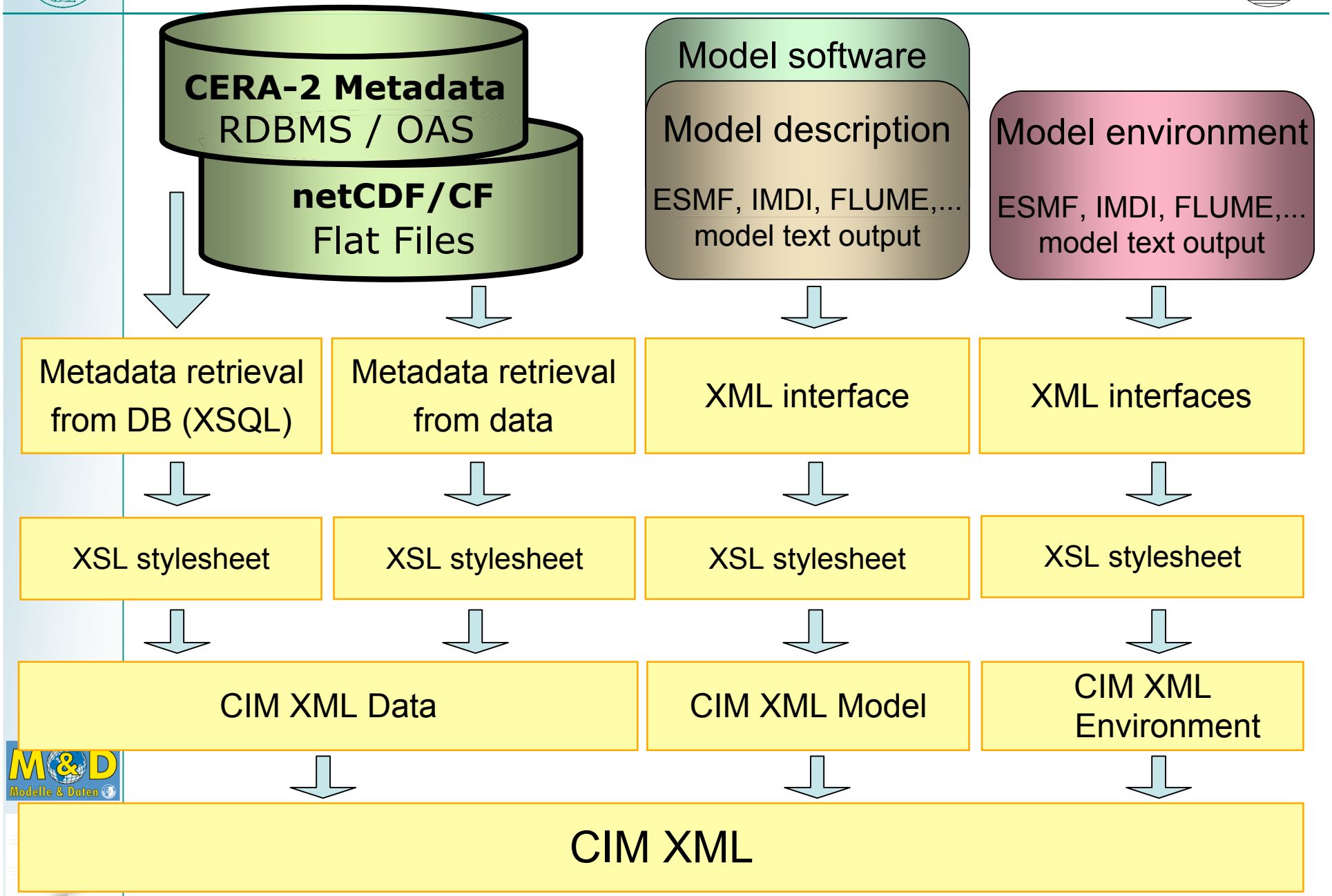


METAFOR: How Could a CIM Structure Look Like ?

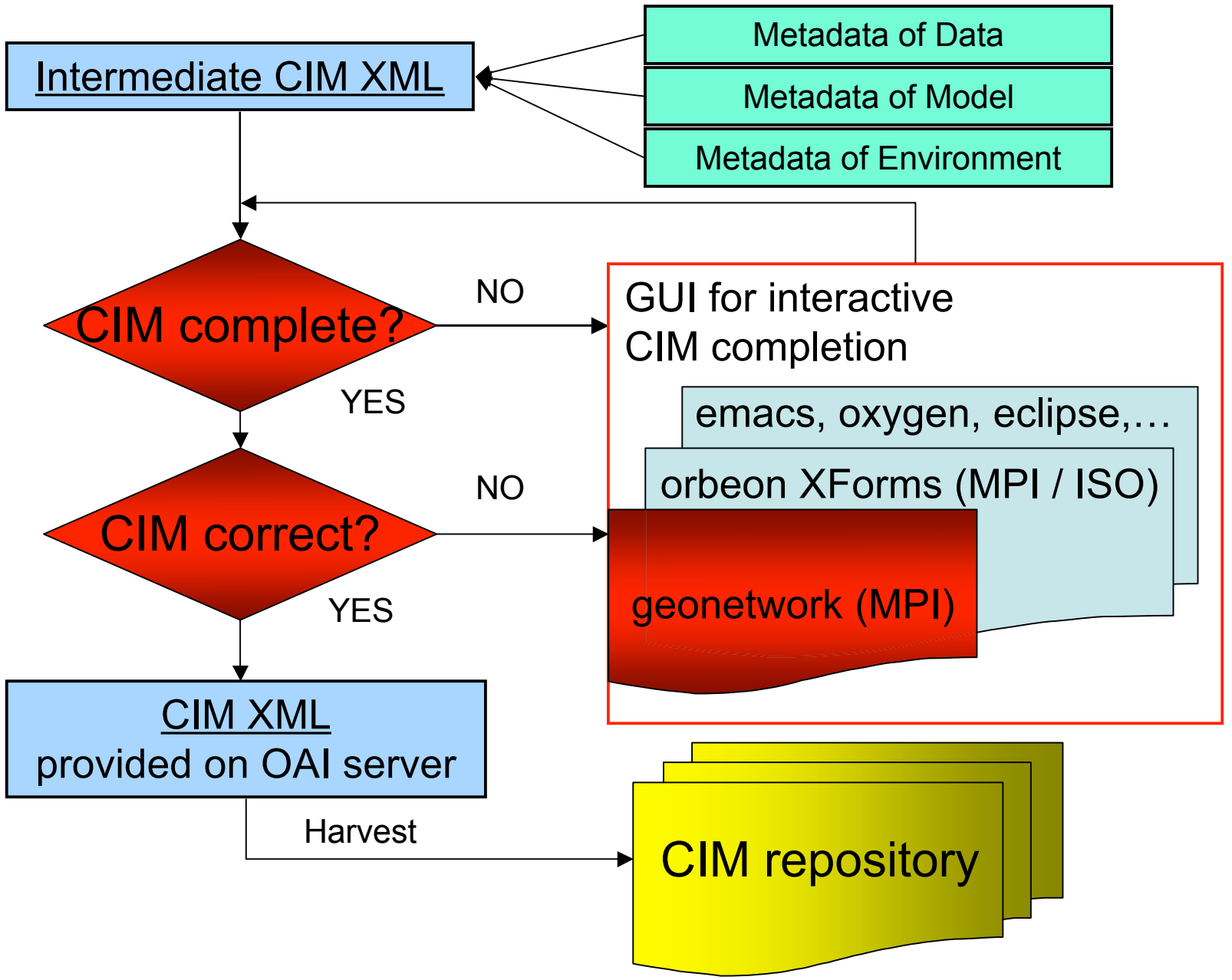




METAFOR: Model Output Data (Metadata Flow)



WP6 Reference Architecture




Thank you !
Questions?