KB-Driven Information Extraction to Enable Distant Reading of Museum Collections

AIUCD Conference
Rome, January 26-28, 2017

CC BY NC ND

G.A. Cignoni & E. Meloni – HMR Project

collections:
tangible cultural heritage

BY NC ND

G.A. Cignoni & E. Meloni – HMR Project

collections and catalogues

value of collections

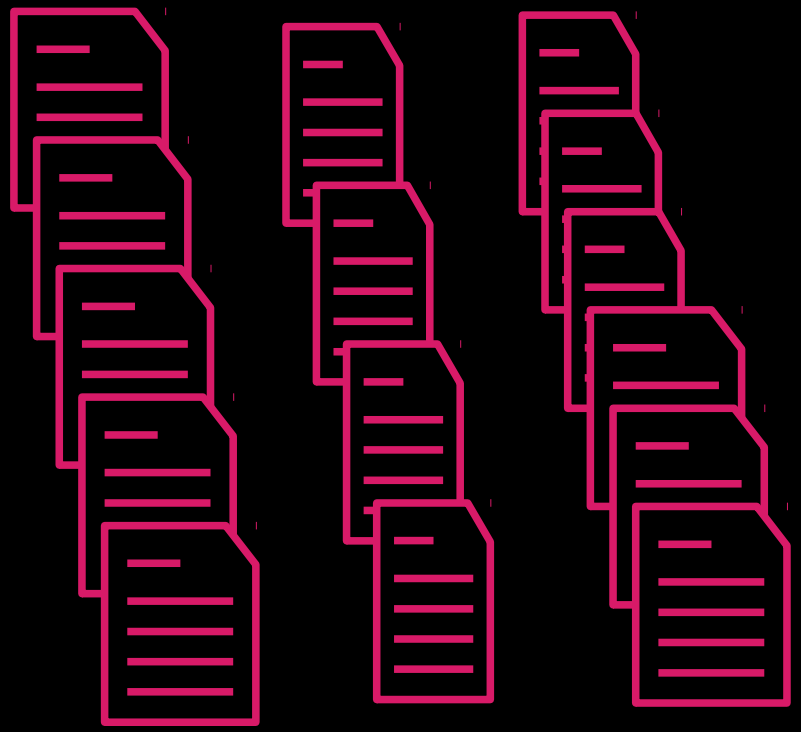
research and disclosure
storytelling through exhibitions and events

catalogues (& archives)

data or knowledge?
sparse lists of records

a paradigm shift

a shared knowledge base
facts, relations, views



CC BY NC ND

collections and catalogues

value of collections

research and disclosure
storytelling through exhibitions and events

catalogues (& archives)

data or knowledge?
sparse lists of records

a paradigm shift

a shared knowledge base
facts, relations, views

CC BY NC ND
G.A. Cignoni & E. Meloni – HMR Project

collections and catalogues

value of collections

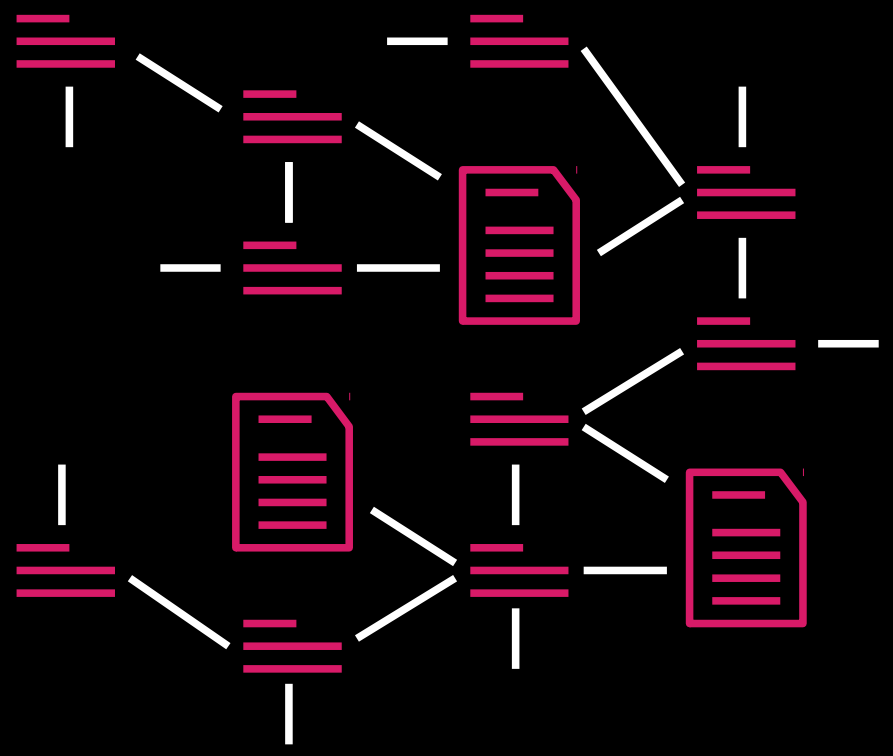
research and disclosure
storytelling through exhibitions and events

catalogues (& archives)

data or knowledge?
sparse lists of records

a paradigm shift

a shared knowledge base
facts, relations, views



CC BY NC ND

a case study: history of information science & technology (HIST)

BY NC ND

G.A. Cignoni & E. Meloni – HMR Project

CHKB

a computing history knowledge base

contents provided by experts

a running prototype...

... still an empty KB

already seen here :)

BY NC ND

problems & opportunities

a common project, a demanding endeavour

museums already have their catalogues
yet another {DB, web site, virtual museum...}

not only collections

lot of information is on the web
literature, retrocomputing clubs, blogs...

quality of information

the web is often unreliable
inaccurate data, biased opinions, hoaxes

CC BY NC ND

the proposed solution
(viable, to be implemented)

BY NC ND

goals and means

collect all the knowledge about HIST

- complete and enrich the catalogues
- harvest the web using the KB ontology (OBIE)

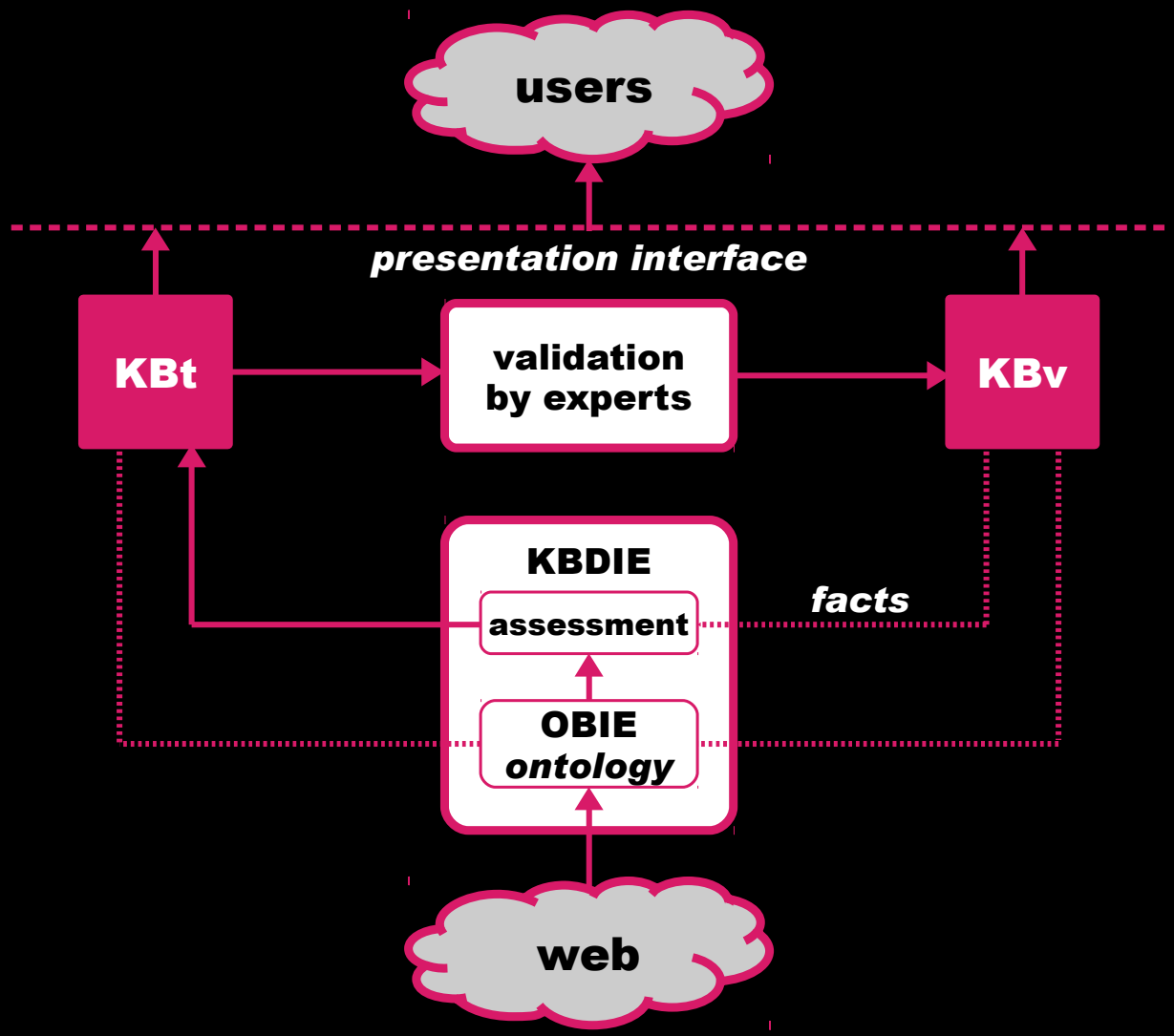
avoid most of the garbage

- initially inject the KB with a good set of facts
- use KB content to assess OBIE results (KBDIE)

maintain the KB authoritative

- put in the KB only facts validated by experts
- just “train” the KBDIE engine (less demanding)

BY NC ND

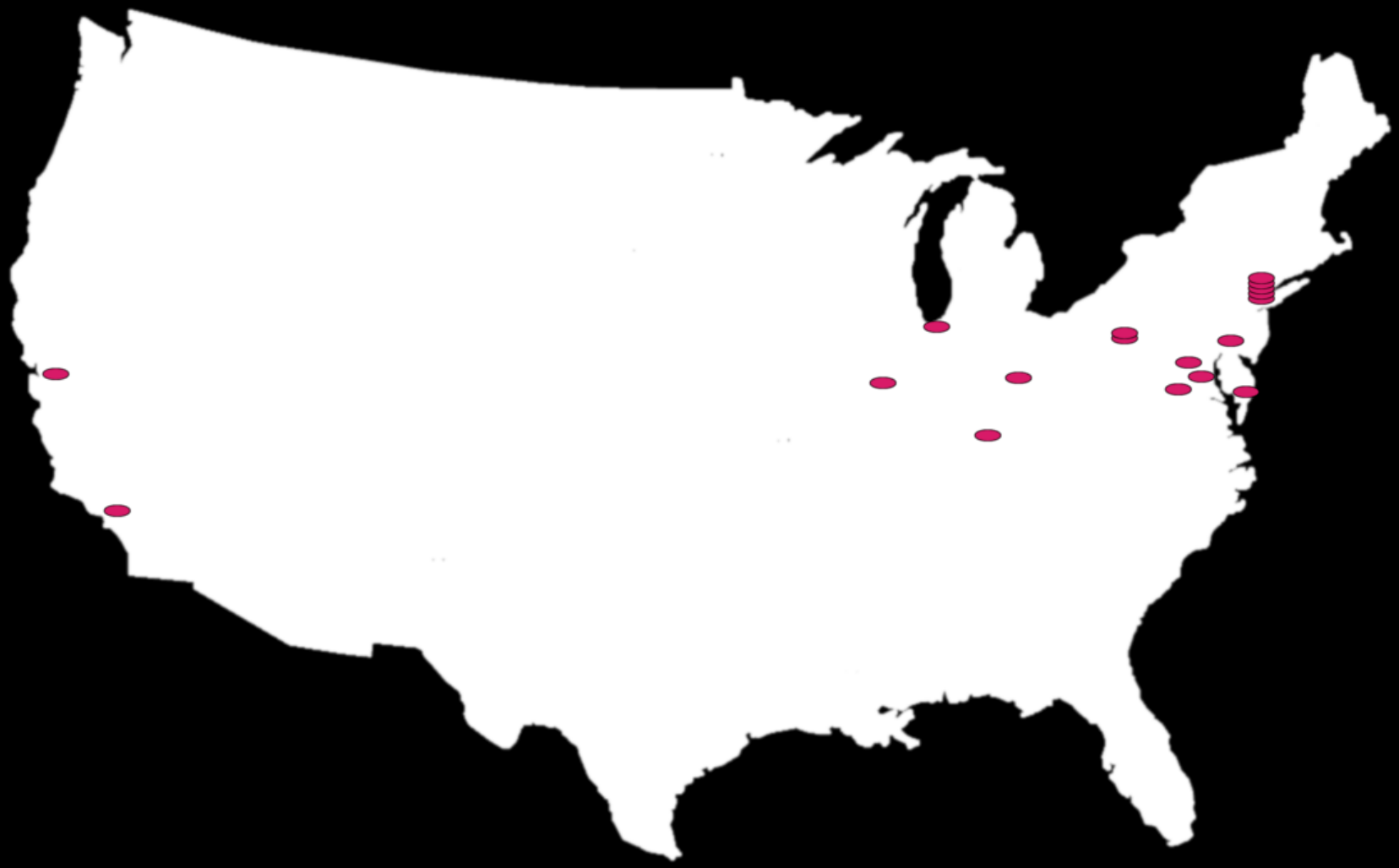

users
presentation interface
KBt
validation
by experts
KBv
KBDIE
assessment
facts
OBIE
ontology
web

BY NC ND

benefits
of a shared KB
(HIST examples)

BY NC ND

UNIVAC (USA) vs...



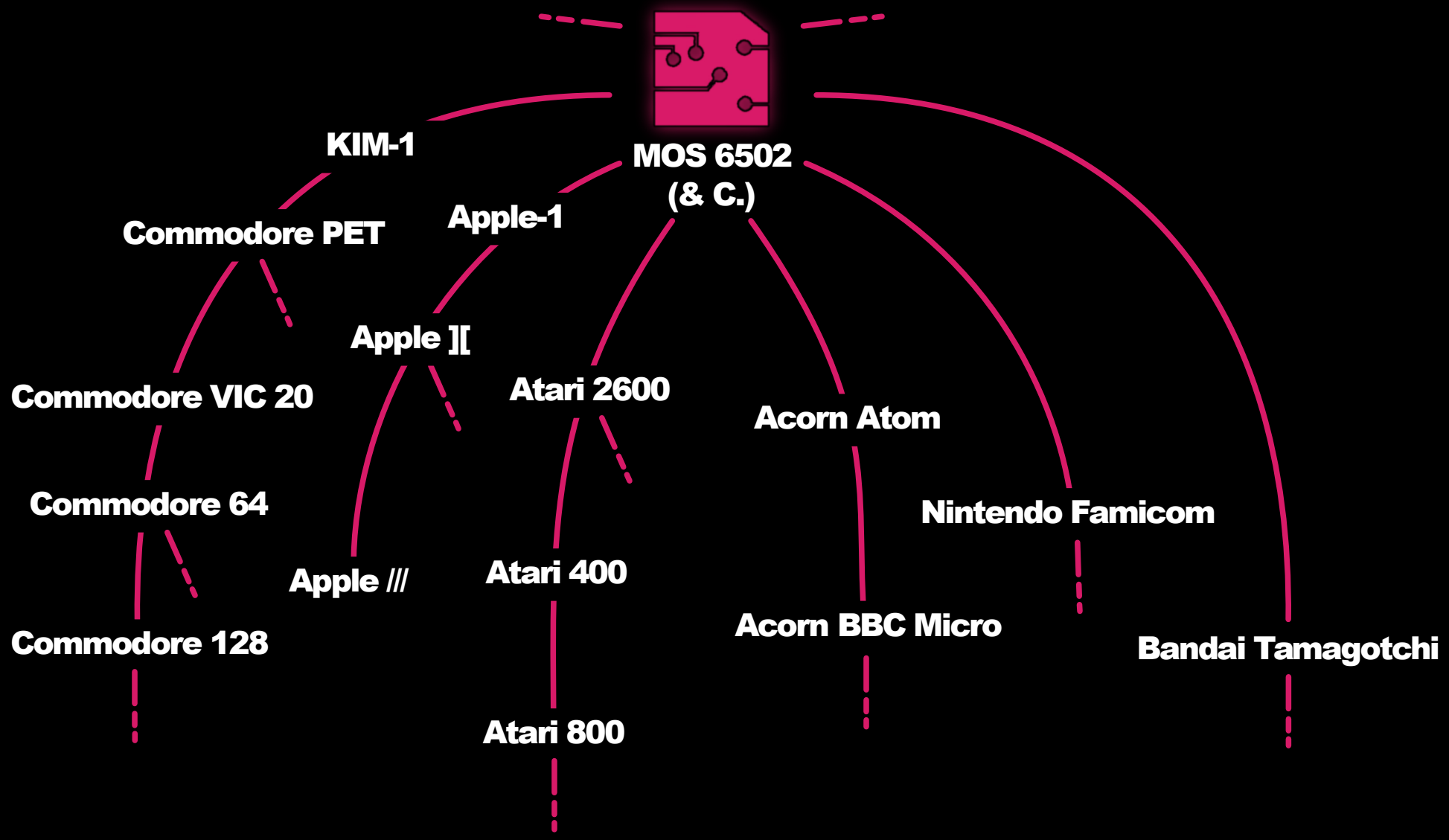
BY NC ND

G.A. Cignoni & E. Meloni – HMR Project

... Ferranti (UK)

BY NC ND

the 6502 lineage



BY NC ND

G.A. Cignoni & E. Meloni – HMR Project

summing-up

full exploitation of available technologies

far beyond old style catalogues
and the current “valuable chaos” of internet

future work

further development of the CHKB prototype
an EU project (fingers crossed)
VAST-LAB PIN, Hyperborea srl & others

BY NC ND