Latar Belakang masalah :

* Sumber data yang beragam dan terdistribusi di berbagai organisasi/negara
* Data WW2 diperlukan untuk berbagai penelitian maupun public
* Data yang sudah ada tidak berada dalam bentuk yang mudah untuk dianalisa

Tujuan :

Data source : spreadsheet, xml, RDF, API/webpages, printed text, digital images.

Data modelling :

* Menggunakan CIDOC CRM sebagai dasar ontology
* Pendefinisian subclass khusus dengan tujuan agar pendefinisial custom property tidak mempengaruhi struktur CRM
* Menggunakan SKOS untuk melengkapi CRM

Challenges :

* Kompleksitas CRM
* Disambiguation pada entity linking

Istilah :

* CIDOC CRM : The **CIDOC CRM** represents an 'ontology' for cultural heritage information i.e. it describes in a formal language the explicit and implicit concepts and relations relevant to the documentation of cultural heritage. The primary role of the CIDOC CRM is to serve as a basis for mediation of cultural heritage information and thereby provide the semantic 'glue' needed to transform today's disparate, localised information sources into a coherent and valuable global resource.
* NEL : Entity linking is the process of connecting entity mentions in text to their knowledgebase counterparts. Prospective information extraction, retrieval, and knowledgebase population are some applications of entity linking. This job, however, is difficult due to entity ambiguity and name variants. Because of the large number of web applications that produce knowledgebase data, major entity linking research has been conducted.
* Event – based modelling
* 7 start LD publishing principles
* 5 star of vocabulary