Utilizing CIDOC-CRM , CRMsci and CRMinf to model Modern and Historical Earthquakes.

- *Iosif Oikonomakis,*
- *Pavlos Fafalios,*
- *Michalis Mountantonakis,*
- *Yannis Tzitzikas.*

https://users.ics.forth.gr/~tzitzik/publications/Tzitzikas_2024-MTSR.pdf

I. Oikonomakis, P. Fafalios M. Mountantonakis, and Y. Tzitzikas,
Modeling Modern and Historical Data as a Knowledge Graph: A Case Study for Earthquake Data, MTSR Conference, November 2024, Athens

25-09-2024 ,
for CIDOC CRM SIG

www.ics.forth.gr/isl

- Information Systems Laboratory, FORTH-ICS, Heraklion, Greece
- Computer Science Department, University of Crete, Heraklion, Greece
- Technical University of Crete, Chania, Greece

In One - Slide

- **Motivation** : Representing historical scientific data is a complex task.
- **Challenges** : Perplexities arise, such as :
  a) conflicting information between sources
  b) variable detail and precision
  c) inconsistent scientific data and date representation
- **Approach** : We face this challenge by focusing on modern and historical earthquakes. We decided to use CIDOC-CRM, CRMsci and CRMinf with a few extensions to create a model, CRM-EQ. We model :
  a) Basic information, information sources and consequences using CIDOC-CRM.
  b) observations, evaluations and measurements using CRMsci.
  c) beliefs using CRMinf.
  d) comparisons and model shortcuts using our own extensions.
- **Implementation** : We implemented CRM-EQ in RDF, along with some examples of earthquakes instances and queries regarding them.

Outline

- **Motivation**
- **Approach**
- **Related work**
- **Methodology**
  1. Collecting initial **questions**.
  2. Finding **available sources**.
  3. **Data Collection**
  4. **Modeling**
  5. **Implementation**
- **Conclusion**

Motivation : Problems in historical scientific data representation

1. Historical data are **sparse** and **vague**.
2. Historical scientific information before the modern era (1900-today) is **infrequent** and often **found in books**.
3. Information is expressed at **various levels of detail and precision**.
4. Details of a historical event are often **conflicting between different sources**.
5. Historical scientific data can **vary in notation and measurement methodology**.

Approach : Modeling Modern and Historical Data

- There arises a need for a **model** that facilitates representation of **both precise and imprecise data.**

- We focus on the domain of **historical earthquakes**.

Case Study : Earthquakes

- *Why earthquake data?*

- **Earthquakes** can be analyzed **both** as **historical events** and as **scientific observations.**
- **Cataloguing** historical earthquake data in a way that facilitates representation, reasoning and analytics **benefits the study** and prediction **of catastrophic earthquake events**.
- The **disparity between modern** seismographic data **and historical information** is profound. Modern data is precise, catalogued frequently and offers little space for uncertainty. Data before 1900 is often imprecise, indefinite and can be conflicting between sources.

Case Study : Earthquakes in Crete

- *Why earthquakes in Crete?*

- Crete is a region with **frequent seismic activity**. Crete lies above a convergent boundary where the Aegean Sea and African tectonic plates meet.
- Historically, Crete is one of the **first places** where signs **of civilization** have been recorded, along with Egypt and Mesopotamia. Therefore, there is both **historical interest** in earthquakes and their impact on civilization.
- Studying **ancient buildings** and the **damage** they have suffered **provides** us with **data about ancient earthquakes**.
- Living in such a highly seismic region and **witnessing catastrophic earthquakes**, like the one on 21st of September 2021, which destroyed the village of Arkalochori, **instills** in ourselves **the importance of researching earthquakes**.

Methodology

1. Collect initial **questions**.
2. Find **available sources**.
3. **Gather** and Inspect **data** provided by sources.
4. Evaluate **modeling** approaches.
5. **Implementation** of chosen modeling approach.

1) Initial questions – what questions would we like to answer ?

- **Basic information** Analytical Queries :
  - Average and maximum of earthquake magnitude in Chania in a year.
  - Earthquakes with intensity > 6 before 1800. For each of their key properties (location, magnitude, depth) provide their precision, and the provenance of the information.

- **Provenance based** Analytical Queries :
  - Earthquakes mentioned in source X, that come from source Y.
  - Earthquakes in source X , whose information is wrong according to source Y

2) Find available sources. – Modern earthquakes

- **Modern earthquake data is readily available** on the internet. We downloaded data from 1900 to today from the seismic catalogues of the United States Geological Survey (USGS) and the University of Athens (UOA).
- The limit was set to earthquakes of magnitude 3 or above, as lower magnitude earthquakes are mostly not felt. We wanted to include earthquakes with epicenter near Crete but not on the island, so we arbitrarily included earthquakes up to 30km from Crete's shores.

2) Find available sources. – Ancient earthquakes

Earthquake data before 1900 is not easily accessible compared to modern earthquake data. We **collected** data **manually** from 2 sources.

I. Eleftherios Platakis' article on 'Κρητικά Χρονικά' vol.4 p.463-526 , published in 1950.

II. Gerassimos A. Papadopoulos book, 'A Seismic History of Crete.' , published in 2011.

3) Gather and Inspect data provided by sources. - Modern Earthquake Data

- **Data about modern earthquakes exists in organized databases** and can be downloaded in many different forms.

- However, **upon inspection**, some **minor problems** did **arise**. For example, the column ‘magnitude type’ in USGS data contains abbreviations (e.g. ‘mb’ , ‘mw’) whose correspondence to a certain type of magnitude is not explained . Furthermore , different abbreviations may refer to the same type of magnitude (e.g. ‘mw’ and ‘mww’).

3) Gather and Inspect data provided by sources. - Ancient Earthquake Data

- Gathering ancient earthquake data proved much more **challenging**. One must :

a) **Discover potential sources of data**, which are usually physical books or digital pdfs.

b) Examine those sources and **understand the way** earthquake **data is collected** and **presented.**

c) Gather the data manually and **record** it in a way that **facilitates modeling** and representation, while simultaneously **does not alter** or omit **information** of the original source.

3) Perplexities in collected Data

- The data we gathered presented us several challenges we had to face in regards with modelling the data. Those challenges are :

  a) **Conflicting information between sources :** Damage caused, deaths, date and intensity of an earthquake may differ between different source material.

  b) **Variable detail and precision :** Imprecise information is common, like uncertain chronologies, varying and relative intensities and epicenters located somewhere in a general area.

  c) **Inconsistent representation :** Notation and measurement units differ between sources. Dates switch from Julian Calendar to Gregorian Calendar, intensities switch from Rossi-Forel to Modified Mercalli and magnitude type notation changes between organizations.

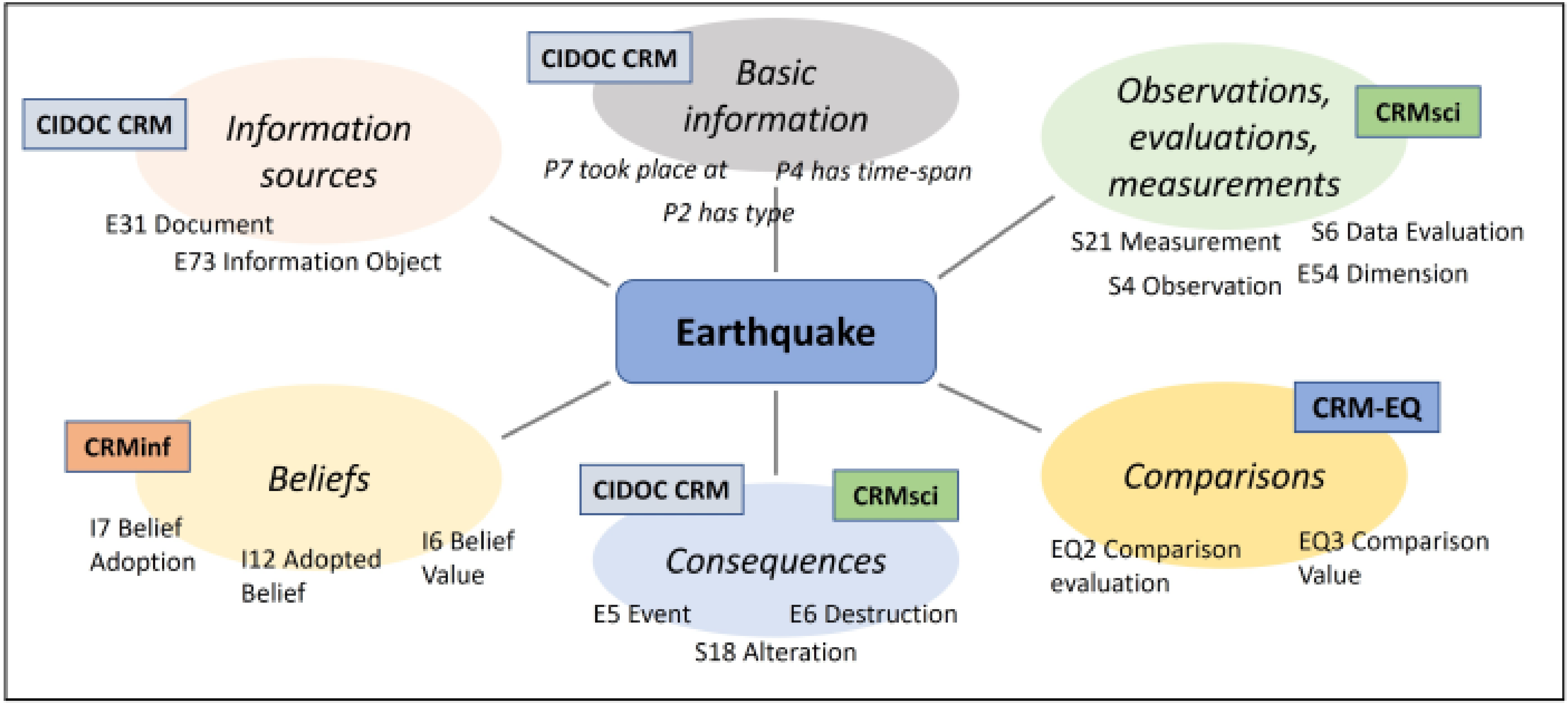
3) Example of uncertain data

Around date (century)	Around 2100. Devastating earthquake destroys almost completely Knossos.
Possible Intensity	1246. Earthquake, according to Raulin, destroys the walls of the city of Chania. Possible intensity 9.
Possible Intensity, Possible year	1303. According to Stavrakis, taken from Zampelios, this earthquake was terrifying. The walls of Chandax (modern day Heraklion) were ruptured and the fort at the entrance of the Heraklion port (Koules) was destroyed. The fort was rebuilt. Possible intensity 8-9. The chronology might be incorrect. It [this account] probably is about the earthquake of 1304. 1304. August 8. Morning. Hieronymus Donatus in his letter to Petros Kontarinos reports a death toll of 4000 due to this earthquake.
Relative magnitude	1861. November 26. Time 00.40 – 02.40 According to Schmidt, two tremors in Chania, one of which at 00.40 and the second at 02.40. Really Strong. 1861. November 27. Time 04.00 – 07.00 According to Schmidt two tremors in Chania at 4.00 and 07.00 . Stronger than the one reported on the 26th of November.

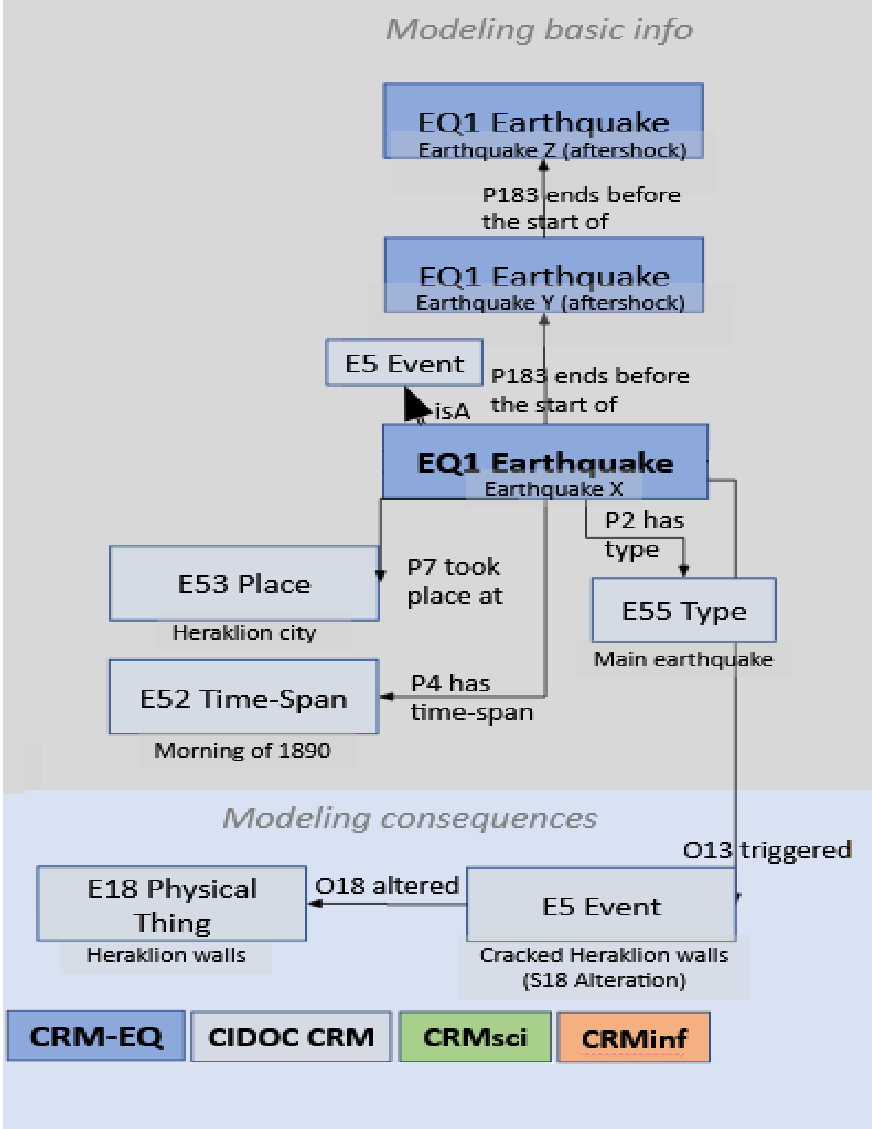
4) Evaluate modeling approaches – Existing ontologies and our requirements.

- Most existing ontologies about earthquakes focus on emergency response, emergency management and engineering.
- Ontologies closer to our purpose are the warehousing ontology by Nimmagadda et al (2007) and the Earthquake Observation Ontology by Uematsu et al (2023).
- However, these are not extensive enough for our case as there is no planning for :
  1. Earthquakes before 1900.
  2. Data ambiguity.
  3. Historical sources and conflicts between them.
  4. Older scientific notations/measurement units (i.e Rossi-Forrel intensity , Julian Calendar dates).

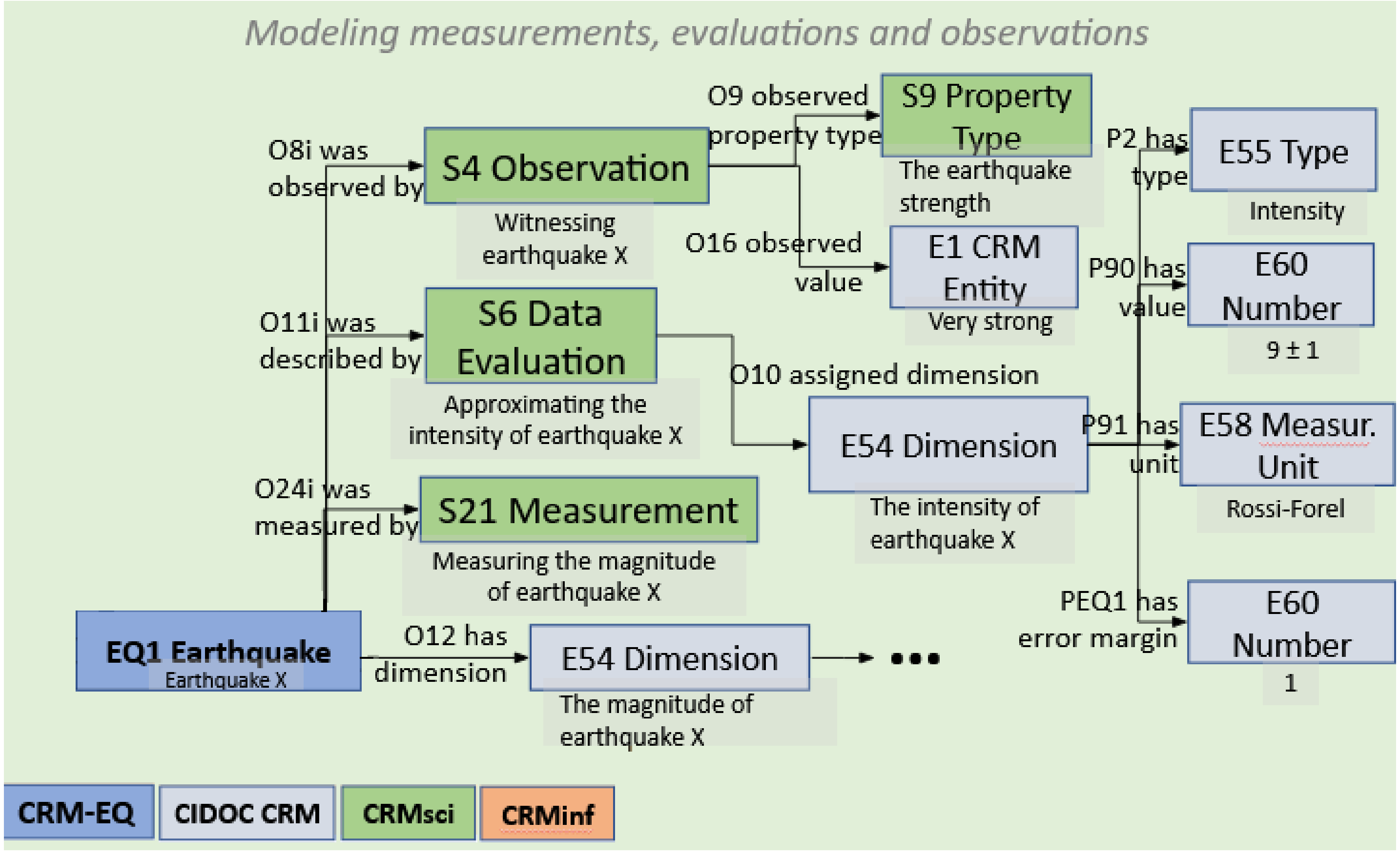
4) Our Modeling - Overview



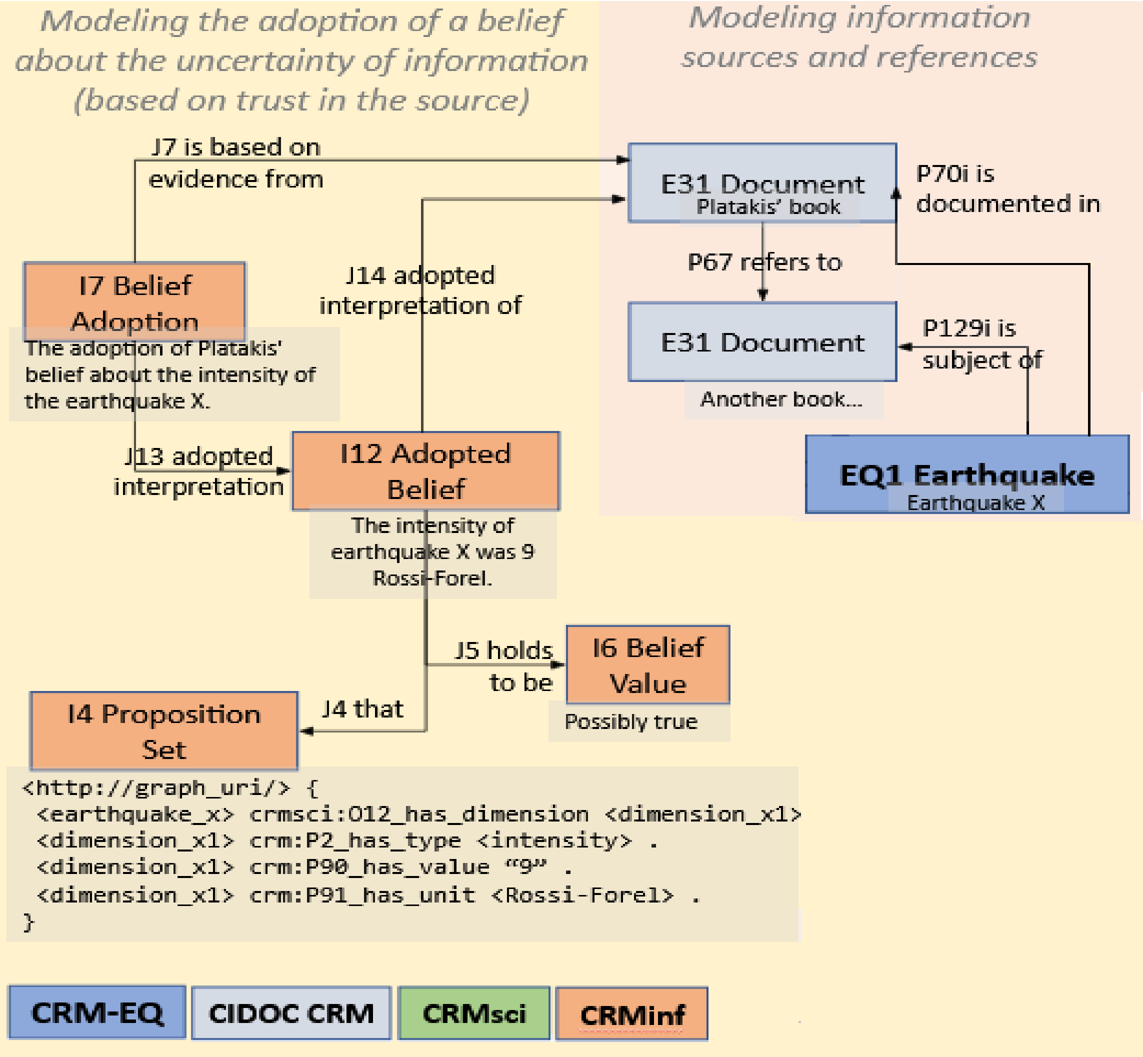
4) Earthquake Model – Basic info and consequences



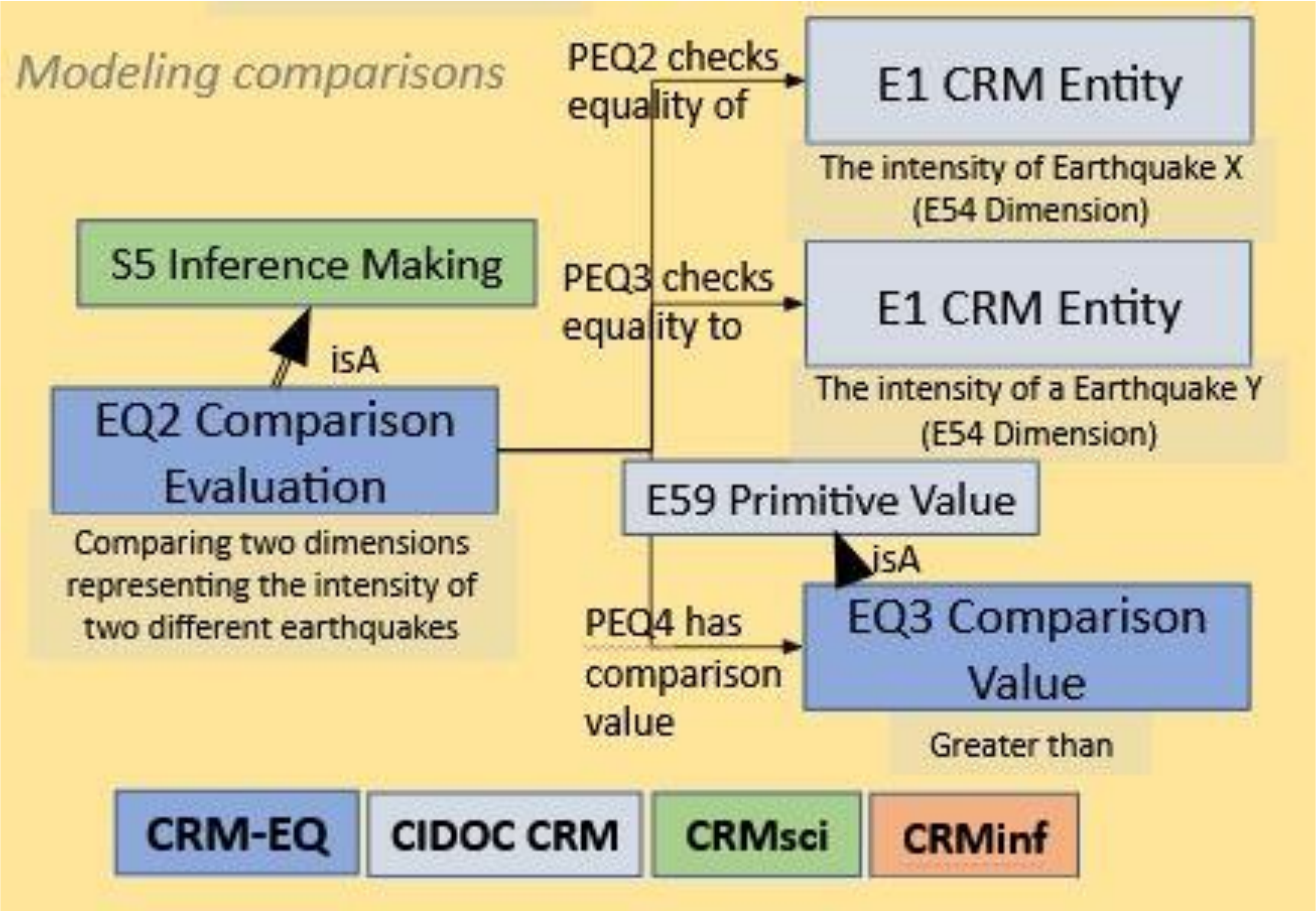
4) Earthquake Model – Measurements and Observations



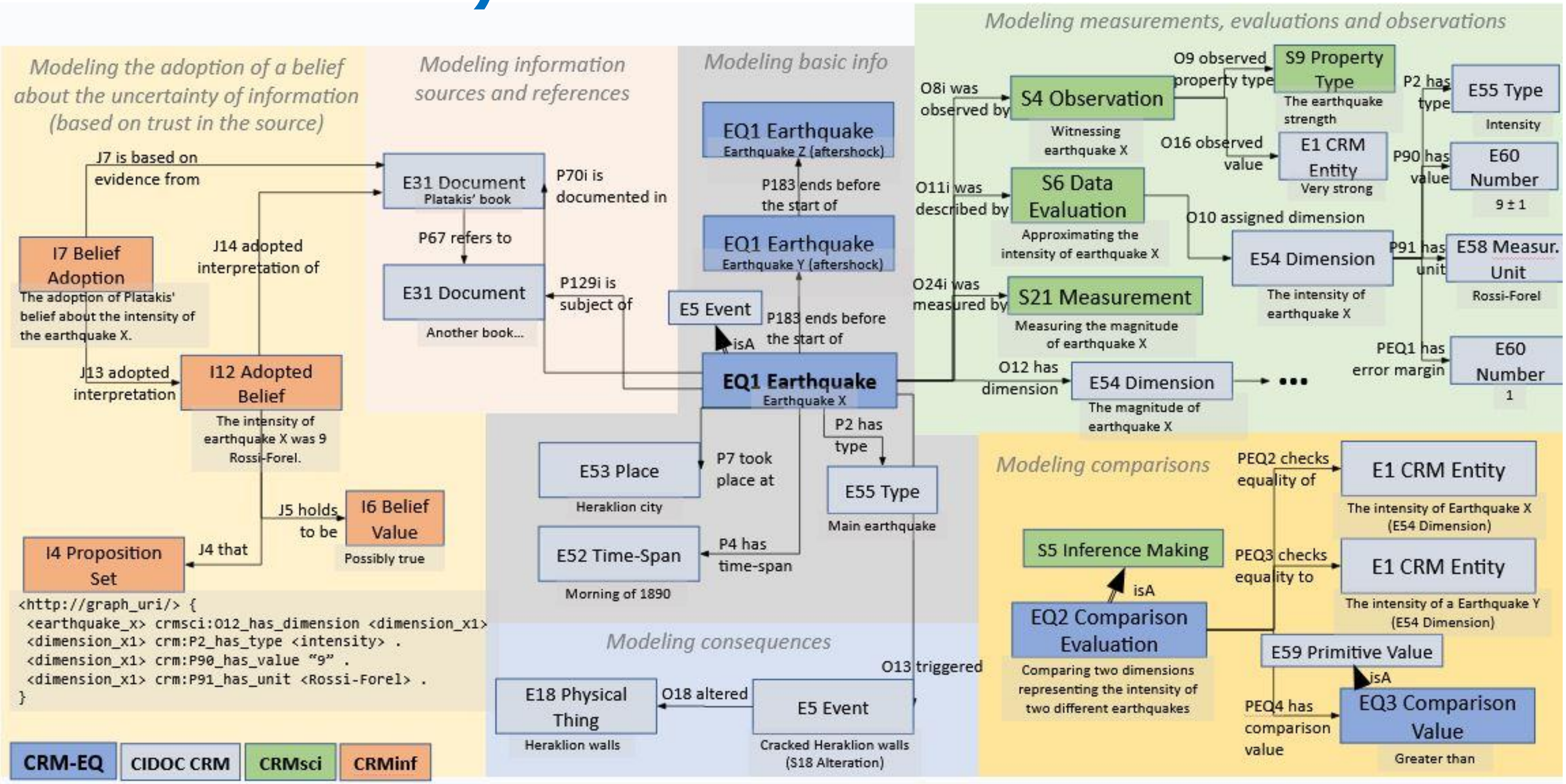
4) Earthquake Model – Sources and Uncertainty Beliefs



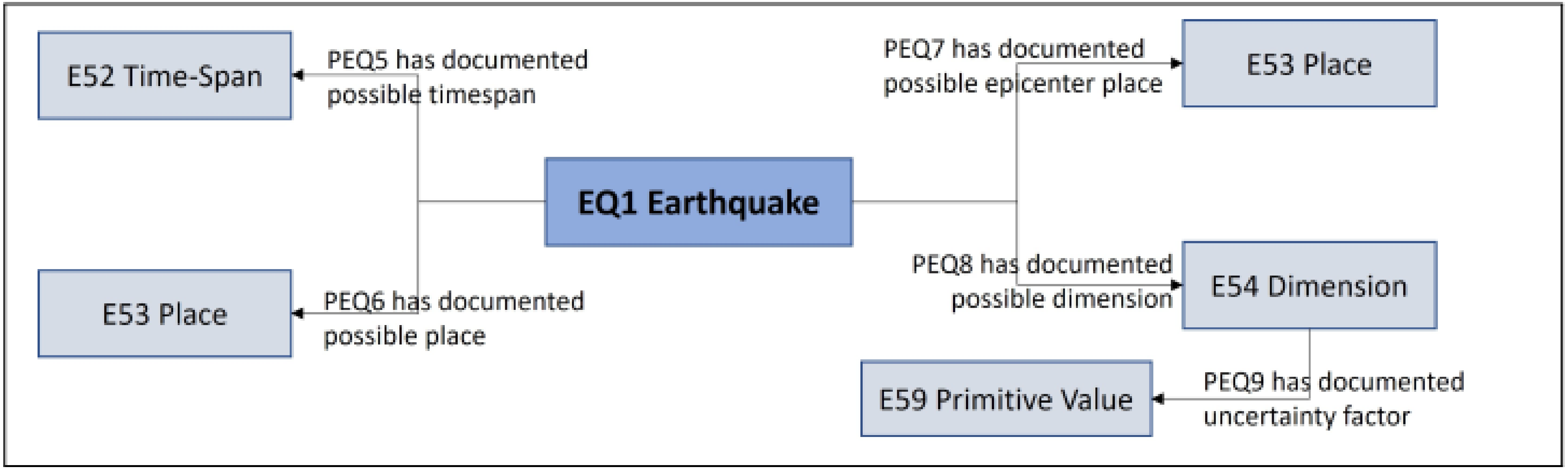
4) Earthquake Model – Comparisons



4) The CRM-EQ Model



4) Shortcuts of CRM-EQ



Shortcut	Sub-Property of
PEQ5 has documented possible timespan	*P4 has time-span*
PEQ6 has documented possible place	*P7 took place at*
PEQ7 has documented possible epicenter place	*P7 took place at*
PEQ8 has documented possible dimension	*O12 has dimension*
PEQ9 has documented uncertainty factor	-

4) Examples

- In the next slides we present some earthquake modeling examples from Platakis' article.

Color Coding

Models	Objects	Properties
CIDOC - CRM	E31 Document	P14 carried out by
CIDOC - CRMsci	S18 Alteration	O10 assigned dimension
CIDOC - CRMinf	I2 Belief	J1 used as premise
Earthquakes Model	EQ1 Earthquake	PEQ9 is in statement

4) Example 1

Earthquake 1

Recorded in different locations, *Relative Epicenter*

Greek Source

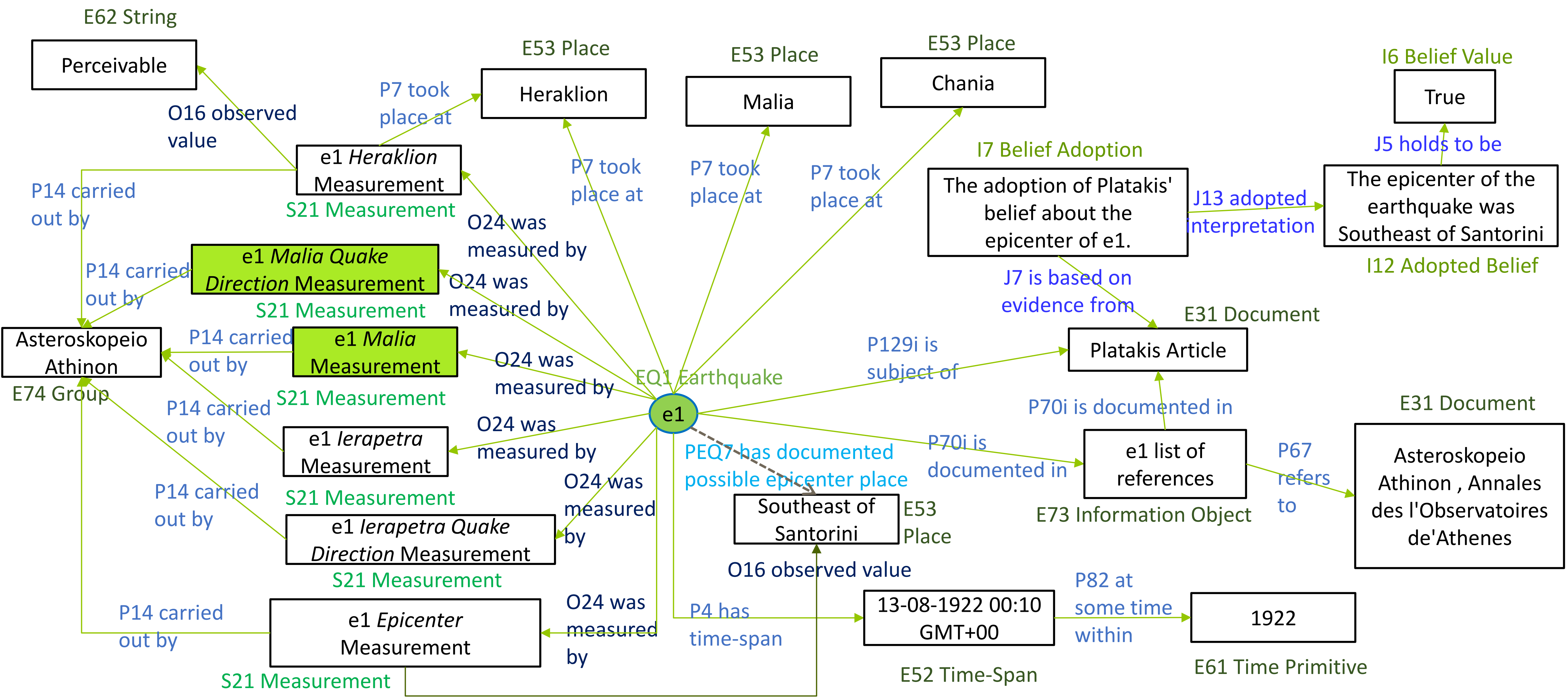
1922. Αὐγούστου 13. Ὥρα 00.10.
Ἡράκλειον: — (αἰσθητός, —, —).
Μάλια: κυματοειδὴς (4, 2 ἕως 3, W)
Ἱεράπετρα: ὁριζόντιος (4, 10, SE).
Τὸ ἐπίκεντρον νοτιοανατολικῶς τῆς Σαντορίνης.

English Translation

1922. August 13. Time 00.10.
Heraklion: - (perceivable, -, -).
Malia: wave (4,2 to 3,W).
Ierapetra: horizontal (4,10,SE).
The epicenter is southeast of Santorini.

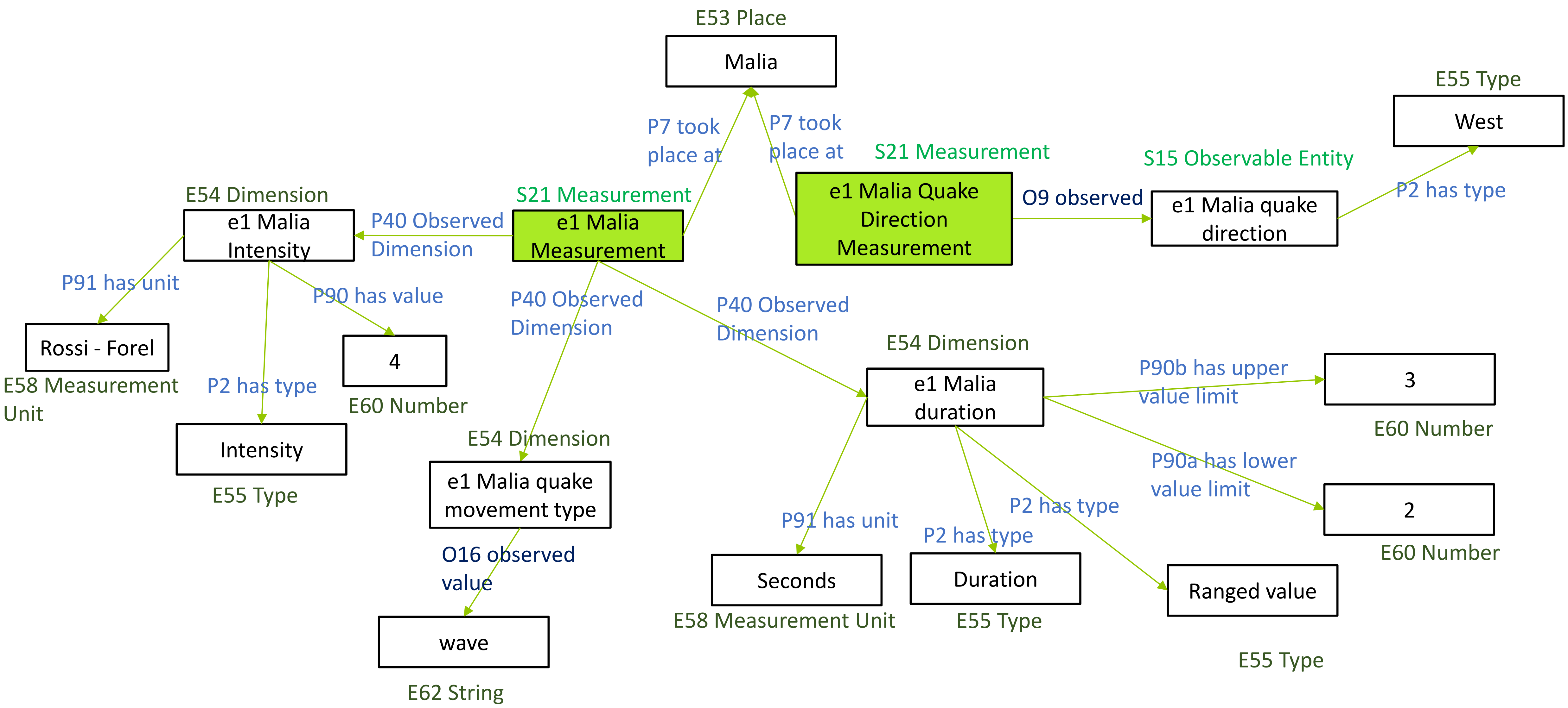
Information in parentheses is (Intensity , Duration , Quake Direction)

4) Modeled Example 1 , Slide 1



CSD FORTH INSTITUTE OF COMPUTER SCIENCE ISL

4) Modeled Example 1 , Slide 2



4) Example 2

Earthquake 2

Possible Intensity

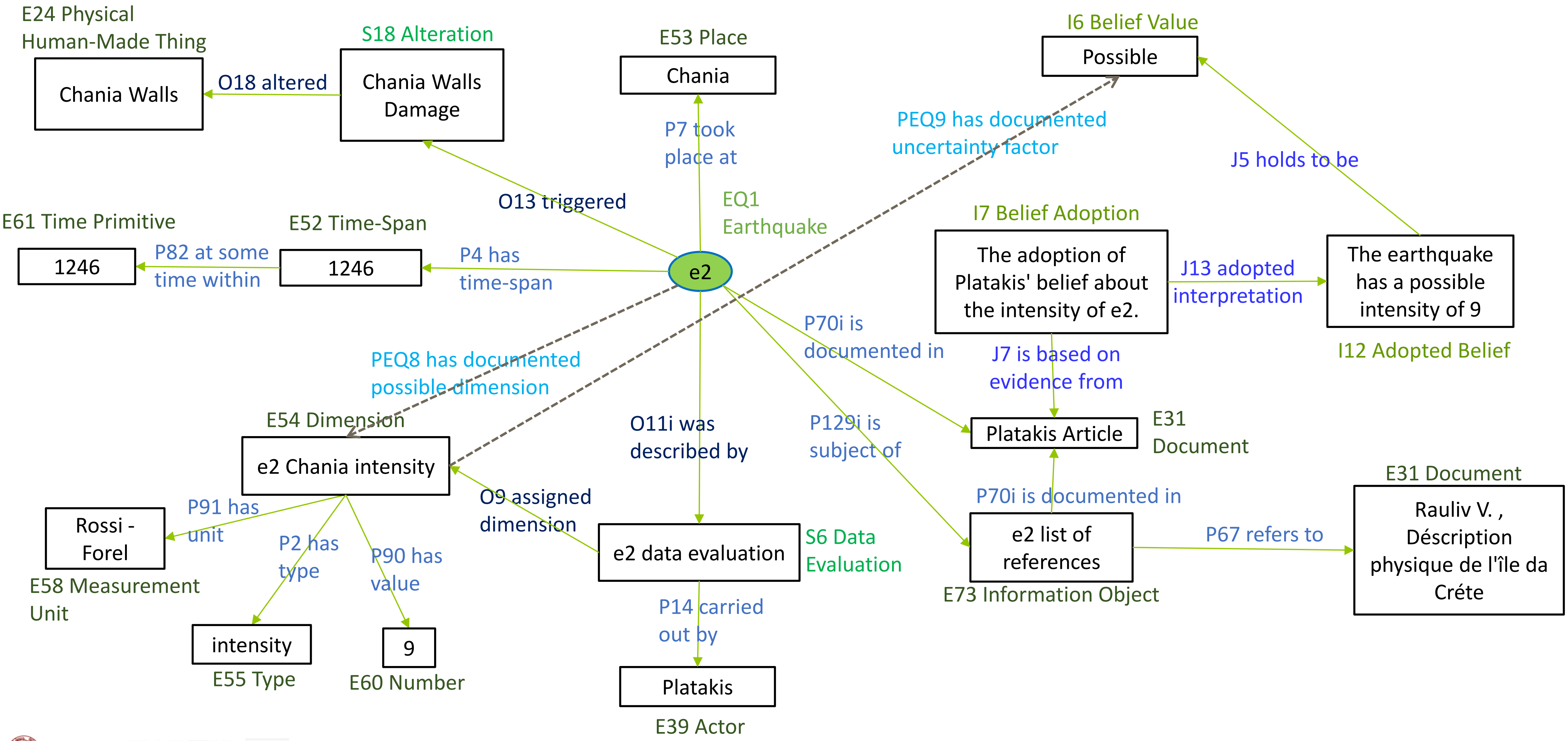
Greek Source

1246. Σεισμὸς κατὰ τὸν Raulin [58] καταστρέψας τὰ τείχη τῆς πό-λεως τῶν Χανίων. Πιθανὴ ἔντασις 9

English Translation

1246. Earthquake, according to Raulin, destroys the walls of the city of Chania. Possible intensity 9.

4) Modeled Example 2



4) Examples 3 and 4

Earthquake 3

Earthquake Comparison,

possible intensity, uncertain date

and

Earthquake 4

Earthquake Comparison

Greek Source

eq3 1303. Κατὰ Σταυράκην (σ. 107), ἀντλοῦντα ἐκ τοῦ Ζαμπελίου[59], ὁ σεισμὸς οὗτος ἦτο φοβερός. Διερράγησαν τὰ τείχη τοῦ Χάνδακος καὶ κατεστράφη ἐπίσης τὸ παρὰ τὴν εἴσοδον τοῦ λιμένος Ἡρακλείου φρούριον (ὁ μεγ. Κοῦλες) ἀνοικοδομηθὲν πάλιν[60].

Πιθανὴ ἔντασις 8–9. Ἐσφαλμένη ἴσως ἡ χρονολογία τοῦ σεισμοῦ τούτου. Πρόκειται μᾶλλον περὶ τοῦ σεισμοῦ τοῦ 1304.

eq4 1304. Αὐγούστου 8. Πρωΐα.

Ὁ Hieronymus Donatus[61] εἰς τὴν πρὸς τὸν Πέτρον Κονταρηνὸν ἐπιστολήν του ἀναβιβάζει τὸν ἀριθμὸν τῶν νεκρῶν ἐκ τοῦ σεισμοῦ τούτου εἰς 4.000.

English Translation

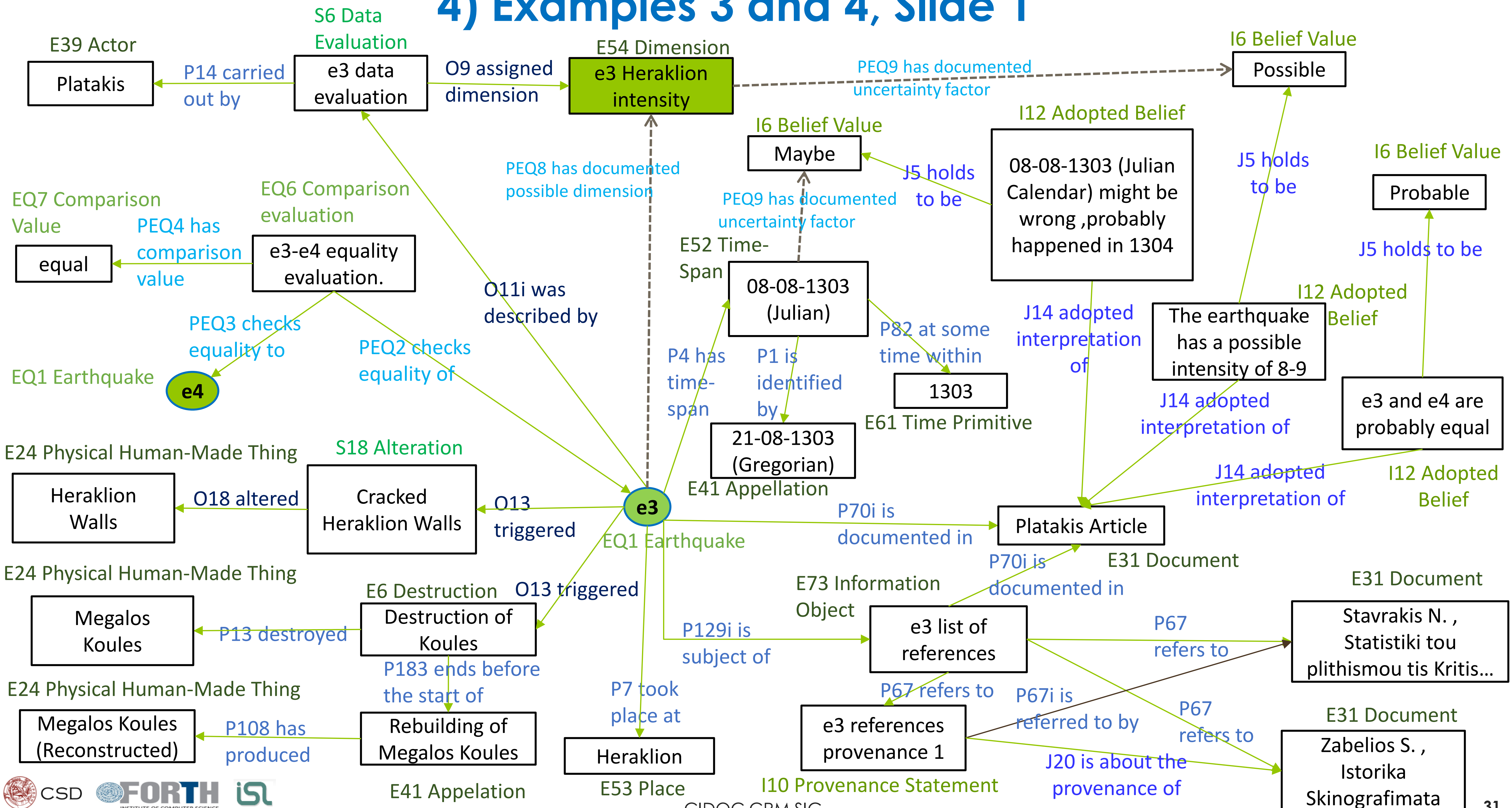
eq3 1303. According to Stavrakis, taken from Zampelios, this earthquake was terrifying. The walls of Chandax (modern day Heraklion) were ruptured and the fort at the entrance of the Heraklion port (Koules) was destroyed. The fort was rebuilt.

Possible intensity 8-9. The chronology might be incorrect. It [this account] probably is about the earthquake of 1304.

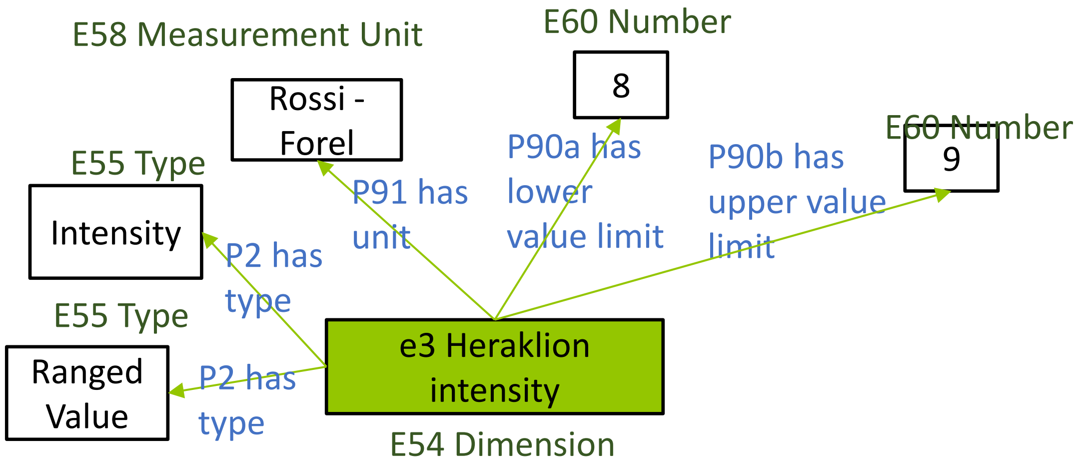
eq4 1304. August 8. Morning.

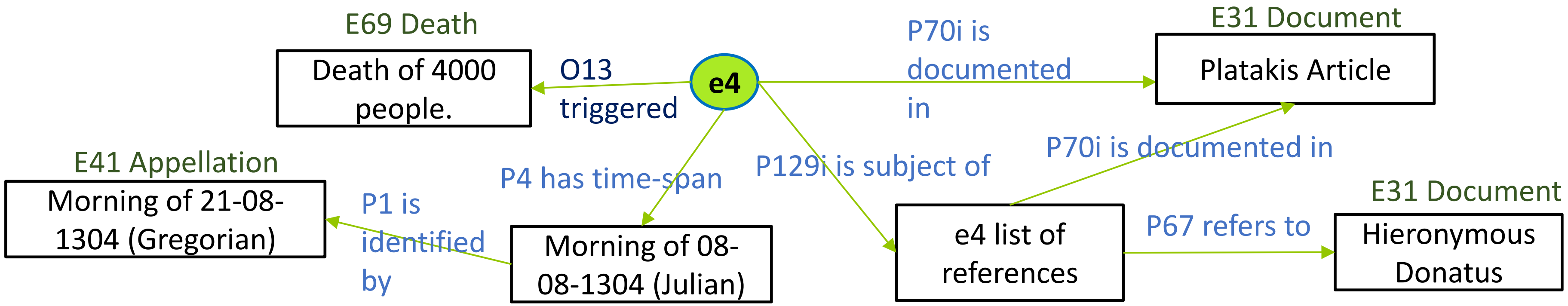
Hieronymus Donatus in his letter to Petros Kontarinos reports a death toll of 4000 due to this earthquake.

4) Examples 3 and 4, Slide 1



4) Examples 3 and 4, Slide 2





5) Implement chosen modeling approach.

- We created the CRM-EQ ontology using RDF representation. Data and information are publicly available at :

  https://demos.isl.ics.forth.gr/crm-eq

- We formulated competency questions to test our implementation

5) Competency question example

- Natural language : *"Give me all earthquakes with intensity >= 6 before 1800, whose focus falls into Crete. For each earthquake, give the following information: date, place, intensity, uncertainty factor (if any), information sources"*

- SPARQL :

```
SELECT ?eq ?date ?place ?intensity ?uncertainty ?source_title
WHERE {
  ?eq a eq:EQ1_Earthquake ; crm:P4_has_time-span ?timespan ; crm:P7_took_place_at ?place .
  ?place crm:P89_falls_within eq:Crete.
  ?timespan crm:P82_at_some_time_within ?date . FILTER (year(?date)<1800) .
  ?eq crmsci:O12_has_dimension ?intDim .
  ?intDim crm:P2_has_type eq:intensity ;
          crm:P90_has_value ?intensity . FILTER(?intensity >= 6) .
  OPTIONAL { ?intDim eq:PEQ9_has_documented_uncertainty_factor ?uncertainty } .
  OPTIONAL { ?source crm:P70_documents ?eq ; crm:P102_has_title ?source_title } }
```

- Note : the underlying engine/triplestore should have inference enabled for the rdfs:subClassOf, rdfs:subPropertyOf and owl:sameAs relationships

Conclusion - Summary

- We presented an approach for handling historical data which are sparse and vague.
- We focused on a case study for earthquake data for the island of Crete.
- We discussed the requirements and challenges for constructing a model for the mentioned data.
- We presented the proposed representation for making it feasible to provide an answer to the competency questions. We have showcased the feasibility of the approach by implementing it using RDF.
- As a future work and research, we plan to produce and make publicly available a complete dataset about all earthquakes in Crete we have available and to investigate various analytic and visualization services that are suitable for uncertain data.

Questions

- Thank you for attending my presentation! I will now accept questions!

MTSR notice

This model will also be presented in the MTSR Conference in Athens, November, 2024 .

https://users.ics.forth.gr/~tzitzik/publications/Tzitzikas_2024-MTSR.pdf

I. Oikonomakis, P. Fafalios M. Mountantonakis, and Y. Tzitzikas,
Modeling Modern and Historical Data as a Knowledge Graph: A Case Study for Earthquake Data, MTSR Conference, November 2024, Athens

Related Work

a) **General Emergency Management ontologies**
   - SOFERS emergency response system (Liu et al, 2014)
   - Emergency Case Ontology Model (Yang et al, 2009)

b) **Earthquake Emergency Management Ontologies**
   - EEM (Spalazzi et al, 2014)
   - EDER (Zhong et al , 2017)

c) **Earthquake Engineering ontologies**
   - Earthquake Engineering Projects and Experiments ontology. (Hasan et al, 2015)
   - OntoBSRA (Xu et al, 2022)

d) **Earthquake Emergency Training**
   - EDSS (Chou et al, 2018)

e) **Data warehousing**
   - Ontology based data warehouse modeling (Nimmagadda et al, 2007)

f) **Earthquake Observation Ontology**
   - Earthquake ontology and LOD (Uematsu et al, 2023)

References

1. Zhong, S., Fang, Z., Zhu, M., & Huang, Q. : A geo-ontology-based approach to decision-making in emergency management of meteorological disasters. Natural Hazards, 89(2), 531–554. (2017) https://doi.org/10.1007/s11069-017-2979-z.
2. Chou, C.C., Jeng, A.P., Chu, C.P., Chang, C.H., Wang, R.G.: Generation and visualization of earthquake drill scripts for first responders using ontology and serious game platforms. Advanced Engineering Informatics 38, 538–554 (2018)
3. Hasan, R., Farazi, F., Bursi, O.S., Reza, M.S.: A faceted lightweight ontology for earthquake engineering research projects and experiments. In: Experimental Research in Earthquake Engineering: EU-SERIES Concluding Workshop. pp. 11–19. Springer (2015)
4. Liu, Y., Chen, S., Wang, Y.: Sofers: Scenario ontology for emergency response system. Journal of Networks 9(9), 2529 (2014)
5. Nimmagadda, S.L., Dreher, H.: Ontology based data warehouse modeling and mining of earthquake data: prediction analysis along Eurasian-Australian continental plates. In: 2007 5th IEEE International Conference on Industrial Informatics. vol. 1, pp. 597–602 (2007). https://doi.org/10.1109/INDIN.2007.4384825

References

6. Spalazzi, L., Taccari, G., Bernardini, A.: An internet of things ontology for earthquake emergency evaluation and response. In: 2014 International Conference on Collaboration Technologies and Systems (CTS). pp. 528–534. IEEE (2014)
7. Uematsu, H., Takeda, H.: Earthquake ontology and LOD (2023), http://wwwkasm.nii.ac.jp/papers/takeda/23/uematsu23ijckg.pdf
8. Xu, M., Zhang, P., Cui, C., Zhao, J.: An ontology-based holistic and probabilistic framework for seismic risk assessment of buildings. Buildings **12**(9), 1391 (2022)
9. Yang, P., Wang, W., Dong, C.: Application of emergency case ontology model in earthquake. In: 2009 International Conference on Management and Service Science. pp. 1–5. IEEE (2009)
10. Papadopoulos, G.A.: A Seismic History of Crete. The Hellenic Arc and Trench - Earthquakes and Tsunamis: 2000 BC - 2011 AD. Ocelotos, Athens (2011)
11. Platakis, E.: Οι σεισμοί της Κρήτης από των αρχαιοτάτων μέχρι των καθ' ημάς χρόνων., Κρητικά Χρονικά vol. D. Andreas G. Kalokairinos (1950)