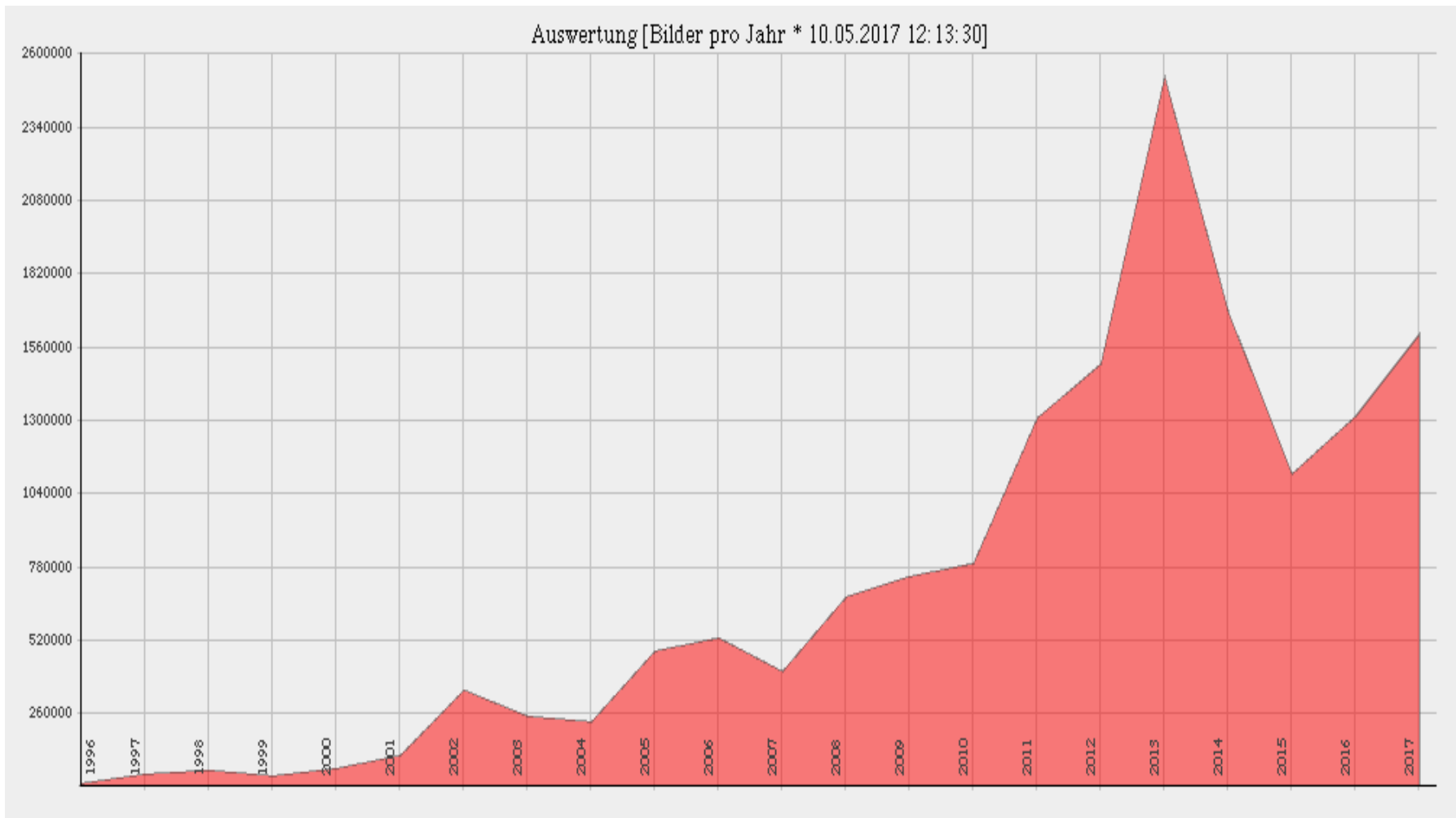


iMAGINE


if autotagging becomes real

KEYSTONE
Switzerland
K. Csoknyay

Challenge



Today

Muster/Bedingung  AP_IN IPTCRule AP_IN

Status







Typ

Priorität

Sequenz

nach erster bedingter zutreffender Regelaktion stoppen

Aktionen

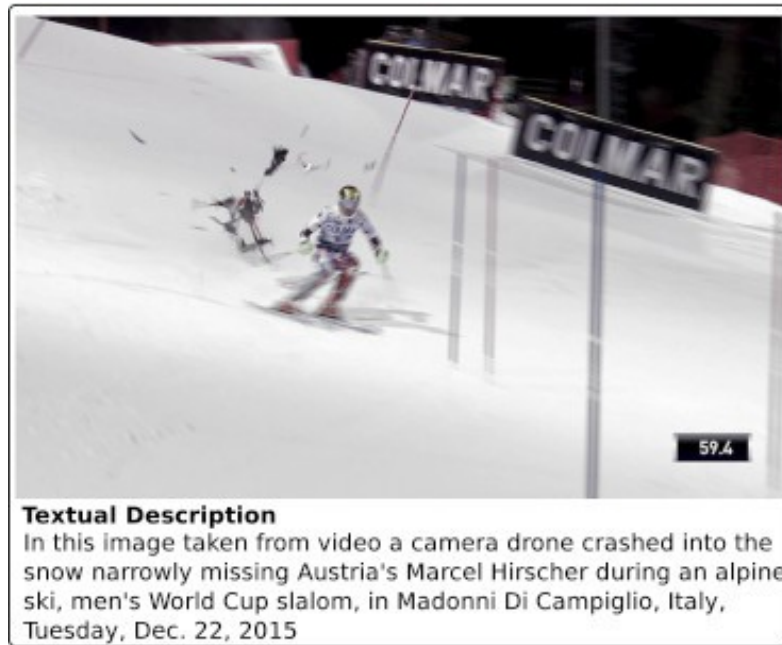
     

Sequence	Create	Zielfeld	Mode	Typ	Quellenfeld	Wert	Bedingung
10	1	i110_credit	template			KEY	
20	1	i100_country_code	template			[ext_i101_country_name]	
30	1	i101_country_name	template				
40	1	udf18	template			112	
50	1	i115_source	prepend			AP	
60	1	i25_keywords	template				
70	1	selection	template			news	
80	1	isOnline	replace			true	
90	1	udf17	replace			EDITORIAL	
100	1	udf14	replace			CHR	
110	1	udf06	template			[i105_headline]	
120	1	i105_headline	template				
130	1	udf17	regex	value	i15_category	c\$ S SPORT	
140	1	i5_object_name	replace	script		i5_object_name.toUpperCase()	!empty(i5_objec...
150	1	udf07	template			[ext_i5_object_name]	

IMAGINE - Project

- CTI Research Project
 - Participants are KEYSTONE and HTW Chur
- Goals
 - automated instead of manual tagging
 - more extensive and consistent metadata
 - training for manual tagging
- Setup
 - HTW provides the technical solution
 - KEYSTONE provides the test corpus

IMAGINE - Overview



Named entities
 Persons, Organizations
 Locations



named entity linking



data enrichment

Named entities & Context Information

International Ski Federation
 Madonna di Campiglio
 Franco Kasperhofen am
 SkiResort, ...
 o, ...

Marcel Hirscher
 Skier, Athlete, WinterSportPlayer, ...
 Birthplace: Austria, Salzburg (state), Annaberg-Lungötz; ...

keyword extraction

Keywords
 camera drone, crash, Marcel Hirscher, alpine ski, World Cup, Madonna Di Campiglio, Italy

story clustering

Stories
 Italy: European Cup, ...
 Drone: camera
 Skiing: alpine ski, World Cup, ...

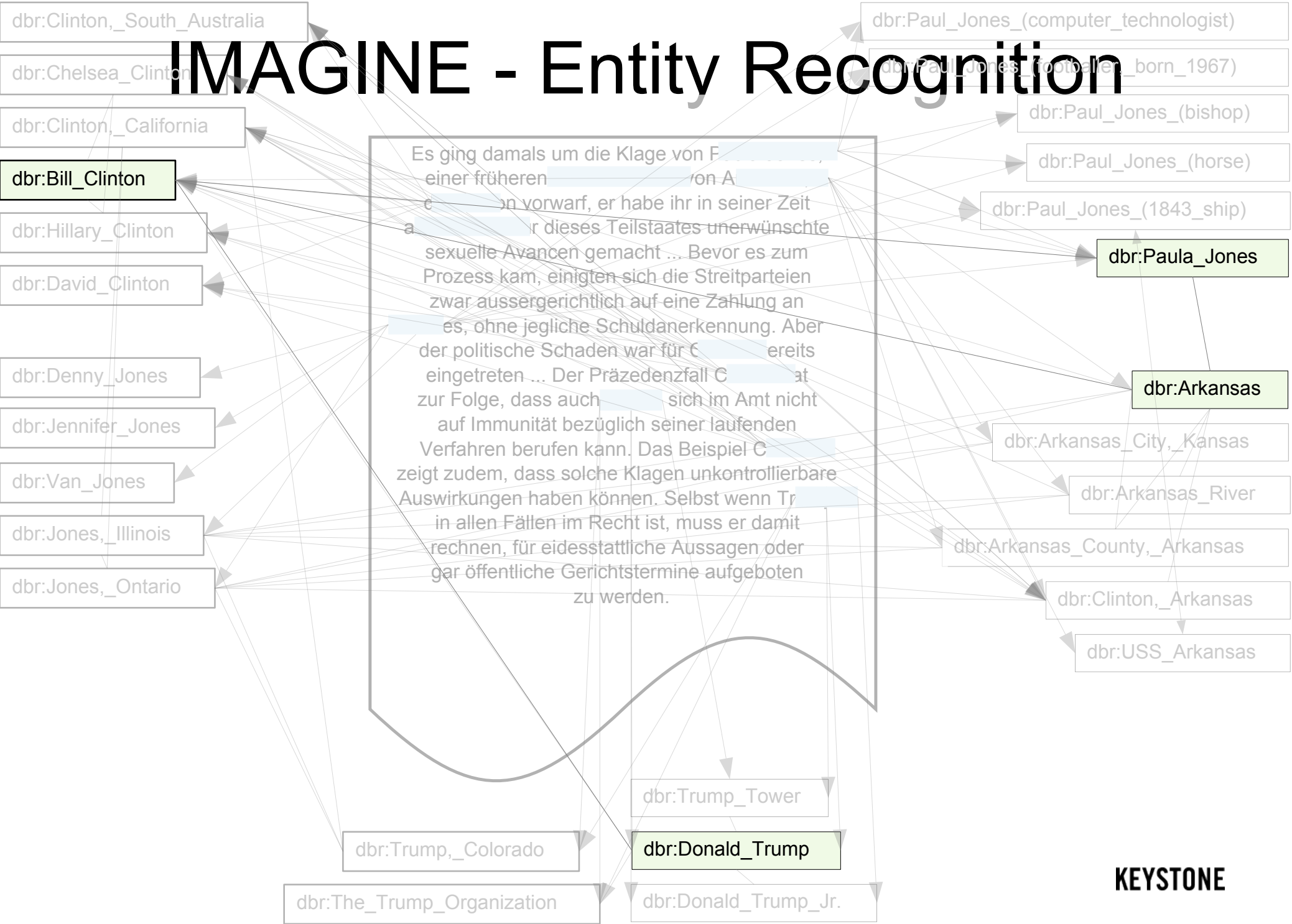
classifier

KEYSTONE classification

alpine skiing, slalom, accident (general)

IMAGINE - Entity Recognition

Es ging damals um die Klage von F [redacted] einer früheren [redacted] von A [redacted] c [redacted] n vorwarf, er habe ihr in seiner Zeit a [redacted] r dieses Teilstaates unerwünschte sexuelle Avancen gemacht ... Bevor es zum Prozess kam, einigten sich die Streitparteien zwar aussergerichtlich auf eine Zahlung an [redacted] es, ohne jegliche Schuldanerkennung. Aber der politische Schaden war für C [redacted] ereits eingetreten ... Der Präzedenzfall C [redacted] at zur Folge, dass auch [redacted] sich im Amt nicht auf Immunität bezüglich seiner laufenden Verfahren berufen kann. Das Beispiel C [redacted] zeigt zudem, dass solche Klagen unkontrollierbare Auswirkungen haben können. Selbst wenn Tr [redacted] in allen Fällen im Recht ist, muss er damit rechnen, für eidesstattliche Aussagen oder gar öffentliche Gerichtstermine aufgebeten zu werden.

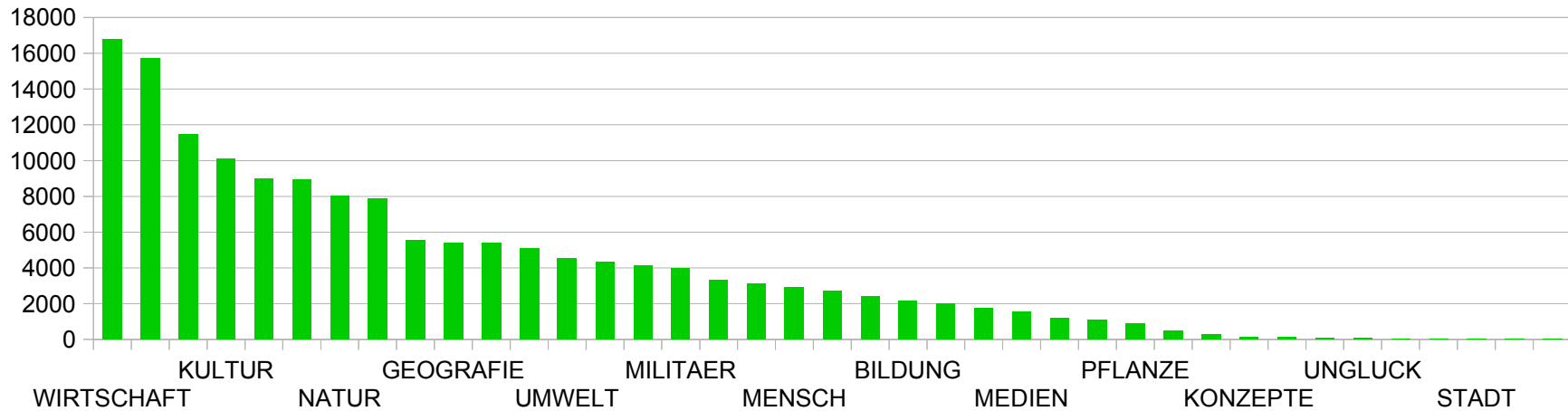


IMAGINE - Story Clustering

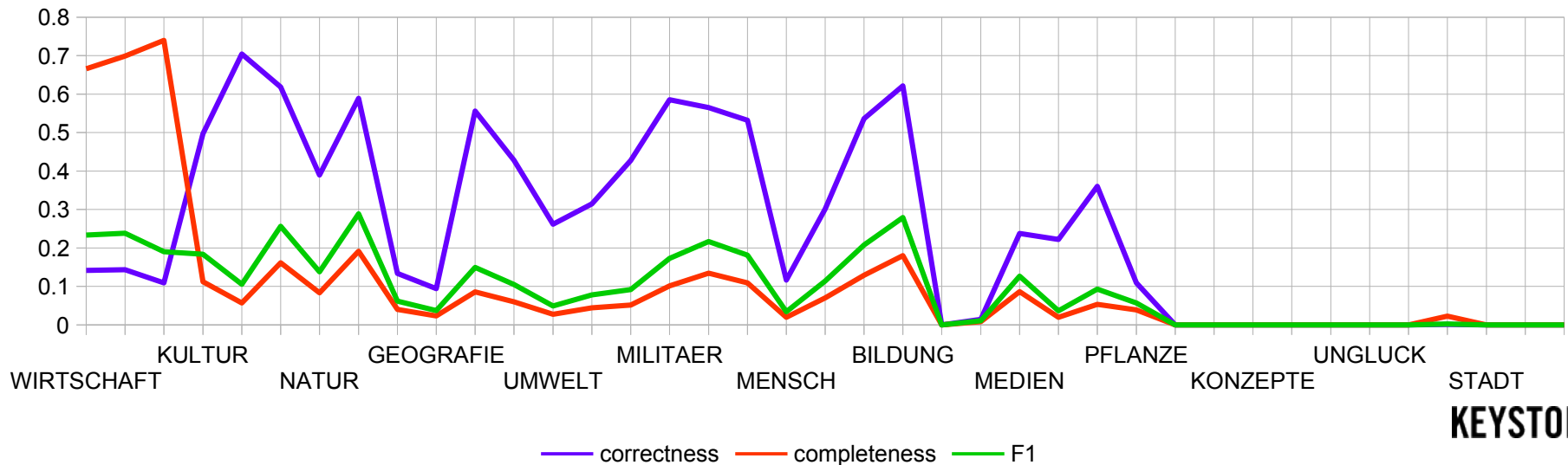


IMAGINE - CV Mapping

Categories in Testdata



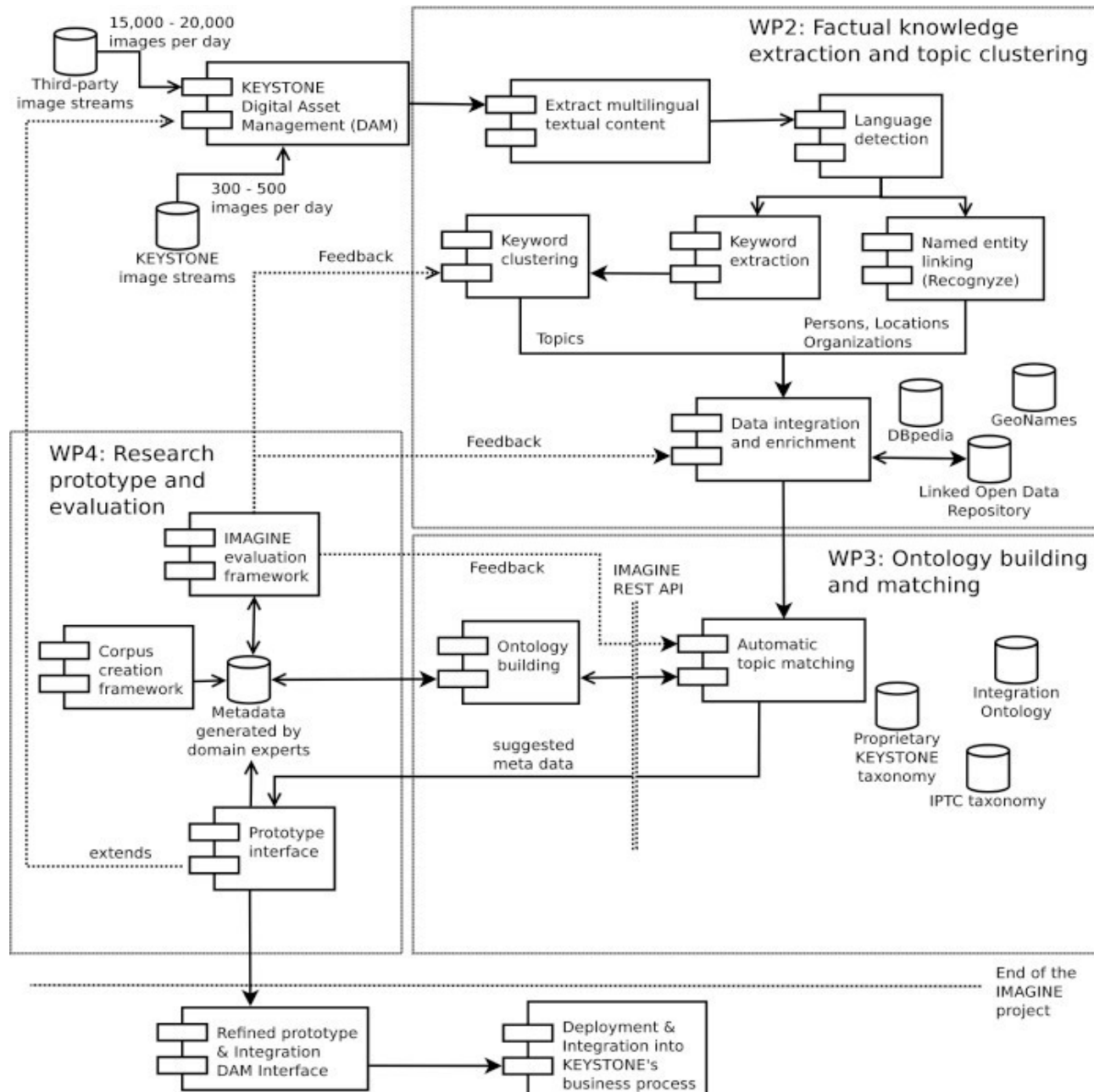
Completeness, Correctness and F1 with Top3



IMAGINE - Result

```
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<imagine xmlns="http://imagine.semanticlab.net/ns/1/" xmlns:xs="http://www.w3.org/2001/XMLSchema" xmlns:xsi="http://www.w3.org/2001/XMLSchema-
instance">
  <associations>
    <association>production,</association>
    <association>farming,</association>
    <association>canton,</association>
    <association>contract</association>
  </associations>
  <entities>
    <entity url="http://dbpedia.org/resource/Jardins" type="geo">Jardins, Jardins</entity>
    <entity url="http://dbpedia.org/resource/Switzerland" type="geo">Switzerland</entity>
    <entity url="http://dbpedia.org/resource/Avusy" type="geo">Avusy</entity>
    <entity url="http://dbpedia.org/resource/Les,_Catalonia" type="geo">Les</entity>
    <entity url="http://dbpedia.org/resource/Kullorsuaq_Island" type="geo">the</entity>
    <entity url="http://dbpedia.org/resource/Christian,_West_Virginia" type="geo">Christian</entity>
    <entity url="http://dbpedia.org/resource/Cocagne,_New_Brunswick" type="geo">Cocagne, Cocagne</entity>
    <entity url="http://dbpedia.org/resource/Geneva" type="geo">Geneva, Geneva</entity>
  </entities>
  <topics>
    <topic url="http://keystone.ch/ontology/topics/KULTUR"/>
    <topic url="http://keystone.ch/ontology/topics/LANDWIRTSCHAFT"/>
    <topic url="http://keystone.ch/ontology/topics/TIER"/>
  </topics>
  <keywords/>
  <clusterDocuments>
    <clusterDocument>276598821</clusterDocument>
    <clusterDocument>276598765</clusterDocument>
    <clusterDocument>276599262</clusterDocument>
    <clusterDocument>276599307</clusterDocument>
    <clusterDocument>276599277</clusterDocument>
  </clusterDocuments>
</imagine>
```

IMAGINE - Architecture



IMAGINE - Conclusion

- Our Conclusion
 - entity recognition is very promising
 - story clustering is interesting
 - keyword mapping is promising
- To Consider
 - without IMAGINE only a part of the daily production may be tagged completely
 - with IMAGINE all daily incoming images may be tagged automatically. For an example: if only 50% may be processed automatically by a lower fail rate it would be better than by manually tagging...

IMAGINE - Outlook

- HTW will finish their work in the following weeks
- KEYSTONE will implement the interfaces by the end of this year
- First tests with our DAM will start next year
- Improve precision and recall remain work in progress
- Use the enriched metadata for our search engines
- Combine IMAGINE with CBIR to get better visual results

IMAGINE - References

- <http://www.htwchur.ch/digital-science/forschung-und-dienstleistung/institut-sii/projektuebersicht/Imagine.html>
- <https://www.weblyzard.com/research/>
- <http://www.keystone.ch/>
- https://www.kti.admin.ch/kti/en/home.html?_organization=711