

CHIST-ERA Projects Seminar 2021

Topic VADMU

Alejandro Benito-Santos

Full-Time Researcher

Data visualization and visual analytics group (VisUSAL)
Department of Computer Science & Automation
University of Salamanca (Spain)

April 13th, 2021



MINISTERIO
DE ECONOMÍA, INDUSTRIA
Y COMPETITIVIDAD



National
Science
Centre
Poland

FWF
Der Wissenschaftsfond:



IRISH
RESEARCH
COUNCIL
An Chomhairle um
Thaighde in Éirinn



Programme co-funded by the
EUROPEAN UNION

PROgressive Visual DEcision-Making in Digital Humanities

The PROVIDEDH project aims to give Digital Humanities scholars a space to explore and assess the completeness and evolution of digital research objects, the degree of uncertainty that the models applied to the data incorporate and to share their perspectives and insights with the project's broad range of stakeholders.

Data Visualisation
and Visual
Analytics tools

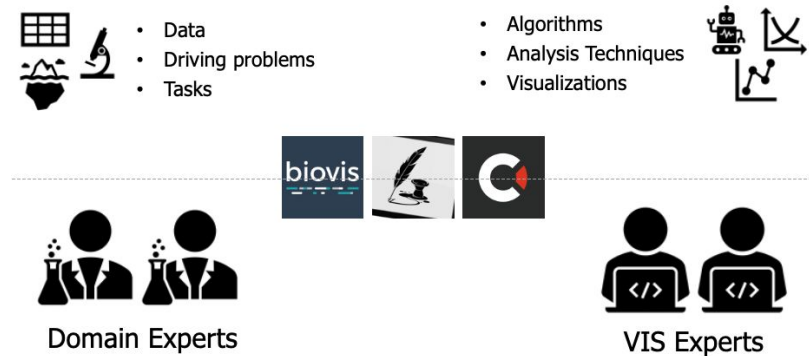
In the context of
computer-assisted
humanities research

Open the black box
with Visual Analytics.

Uncertainty Management
and Representation

Visualization for the Digital Humanities

- Visualization for the Digital Humanities (VIS4DH) is a type of interdisciplinary visualization research in which domain and visualization experts collaborate to study traditional humanities questions:
 - **Domain experts** provide **data, driving problems** and **tasks**.
 - **Visualization experts** provide **algorithms** and **visualization/analysis techniques**.
- It presents common elements with other interdisciplinary collaborations (e.g., BioVis or VizSec)
- But also many other features that make it unique....



Visualization for the Digital Humanities (VIS4DH)

- DH are an **exciting** field of **experimentation** for VIS researchers:
 - Datasets normally are **rich** and **underexplored**.
 - Text + Image Visualization
 - DH **pursue different aims** to those of other disciplines that are commonly supported by visualization.
 - The main **aim of humanities research** is not to produce **definite answers** to questions, but rather to **enable the formulation of more and better questions**.
- VIS4DH has grown in **popularity** in recent times.
 - 5th edition of the VIS4DH Workshop (<https://vis4dh.dbvis.de/>)

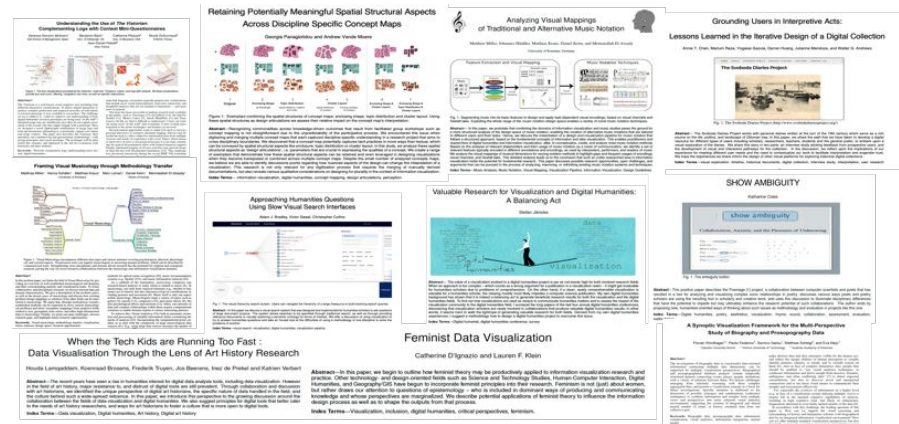


Fig: Some papers presented at VIS4DH.

Visual Analytics, Uncertainty and Digital Humanities

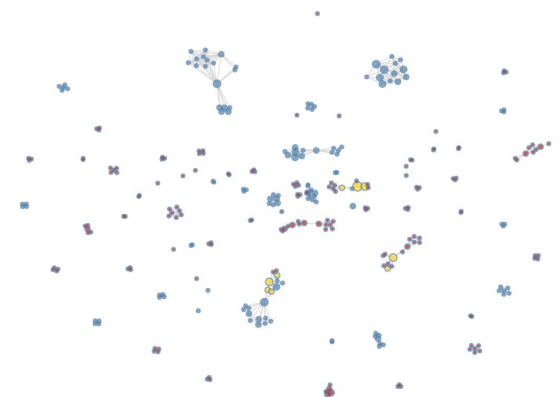
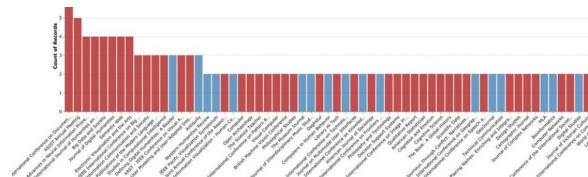
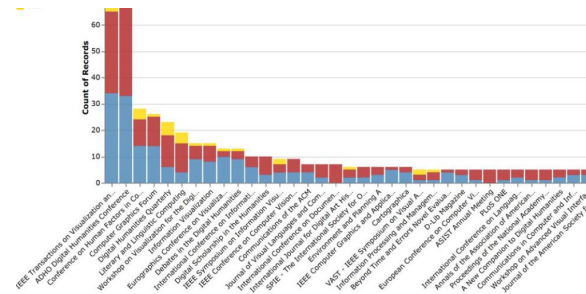
“As digital visualization tools have become **more ubiquitous**, humanists have adopted many applications such as **GIS mapping, graphs, and charts** for **statistical display** that were developed in other disciplines. But, I will argue, such graphical tools are a kind of **intellectual Trojan horse**, a vehicle through which assumptions about what constitutes information swarm with potent force.”

“**Creating bar charts with ambiguity and degrees of uncertainty** or other variables in them **might cause champions of legibility and transparency** some unease, but the **shift away from standard metrics** to metrics that express interpretation is an **essential move for humanists** and/or constructivists across disciplines.”

J. Drucker, «Humanities Approaches to Graphical Display», DHQ, vol. 5, n.º 1, pp. 1–21, 2011.

VIS4DH - Challenges

- Disparity of research aims.
 - Existing **evaluation techniques often are not fit** for this kind of research.
 - **Previous VIS research** to get inspiration from is **scarce**.
 - It is sometimes hard to produce “valuable” visualization research in these collaborations.
- The discipline is highly fragmented.
 - There is still a **large gap** between VIS4DH and DH4VIS scholars.
 - Different venues & formats.
 - Low overlap between communities.



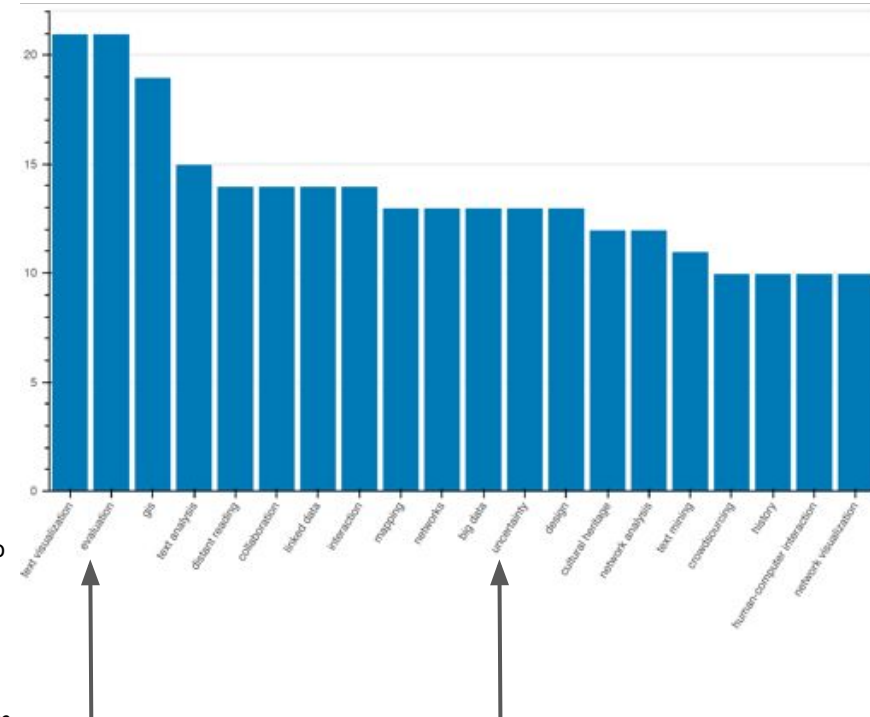
A. Benito-Santos y R. Therón, «Pilaster: A Collection of Citation Metadata Extracted From Publications on Visualization for the Digital Humanities», en 2020 IEEE 5th Workshop on Visualization for the Digital Humanities (VIS4DH), oct. 2020, pp. 24-29, doi: 10.1109/VIS4DH51463.2020.00009.

A. Benito-Santos and R. Therón, "GlassViz: Visualizing Automatically-Extracted Entry Points for Exploring Scientific Corpora in Problem-Driven Visualization Research," in *2020 IEEE Visualization Conference (VIS)*, Oct. 2020, pp. 226–230, doi: 10.1109/VIS47514.2020.00052.

Challenges in Digital Humanities Visualization

- Better evaluation techniques for tools and results.
 - How do we know VIS4DH tools work well?
 - Only users can tell!
 - We need to involve humanities scholars deeply into the design of interactive systems.
- Better categorization, assessment and communication of **uncertainty**.
 - In the humanities research process
 - In the data analysis pipeline (algorithms, models, data formats, visualization)
- Expose algorithms and other computational methods to the final user.
 - Open the black box and allow finer control.
 - Allow users to know what is going on behind the scenes and also to increase their algorithmic/computational literacy.

Top 20 most-used keywords (795 in total)



#2 evaluation

#12 uncertainty

Two key scientific questions:

1. Is it possible to **assess the degree of uncertainty** that a particular humanities digital research object incorporates, **tracking** such degree during the evolution of the dataset and how it is affected when different computational models are applied to it?
2. Is it possible to **convey this evolution of uncertainty** by means of interactive multimodal interfaces that progressively adapt to the moment in which they are used to enable a better decision-making?

- University of Salamanca (SPAIN)
— [Visual Analytics](#) Leader
Project Coordinator: Roberto Therón
- Austrian Academy of Sciences (AUSTRIA)
— [Open Innovation in Digital Humanities](#) Leader
PI: Eveline Wandl-Vogt
- Trinity College Dublin, Faculty of Arts Humanities and Social Sciences (IRELAND)
— [Engagement and Exploitation](#) Leader
PI: Jennifer Edmond
- Poznan Supercomputing and Networking Center (POLAND)
— [Infrastructures and Platform Leader](#)
PI: Cezary Mazurek



Approach

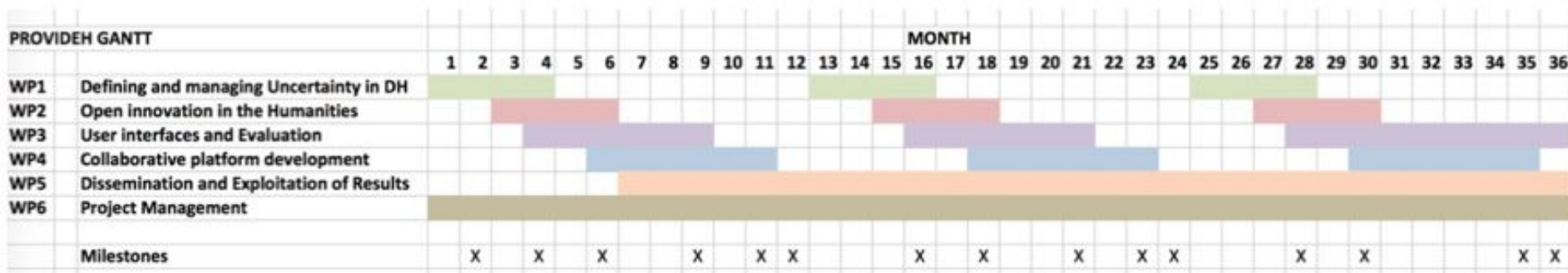
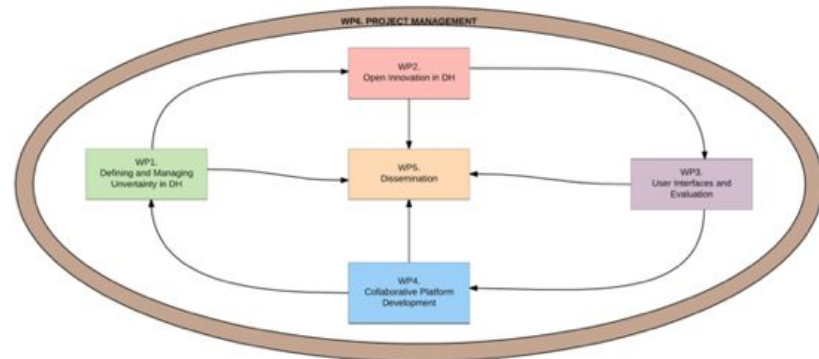
PROVIDEDH is guided by a User Centered Design process integrated with agile development.

Furthermore, the plan contemplates three iterations (one per year), so we make sure that we can respond to user feedback.

The agile development will include the rapid design and development of micro-prototypes.



Programme co-funded by the
EUROPEAN UNION

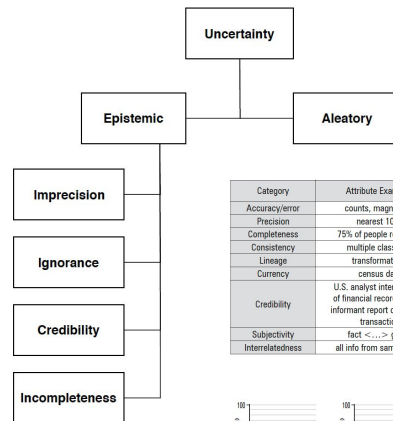


Modeling uncertainty in DH

Uncertainty Taxonomies

Existing uncertainty models

- There are many existing uncertainty visualization taxonomies:
 - We first relied on previous models originating in visualization and GIScience.
 - GIScience is **conceptually closer** to **humanities/social sciences**.

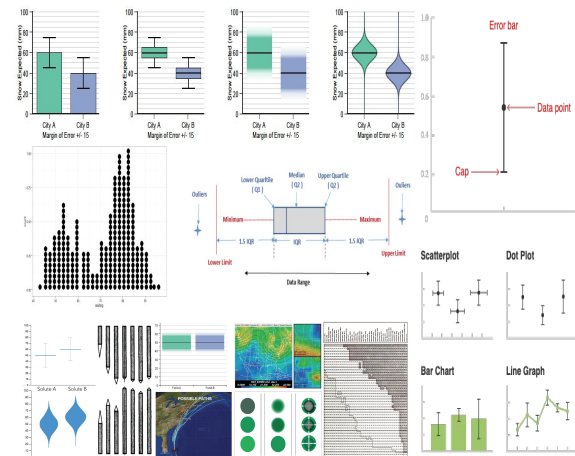


Category	Attribute Examples	Location Examples	Time Examples
Accuracy/error	counts, magnitudes	coordinates, buildings	± 1 day
Precision	nearest 1000	1 degree	once per day
Completeness	75% of people reporting	20% of photos flown	2004 daily/12 missing
Consistency	multiple classifiers	from / for a place	5 say Mon; 2 say Tues
Lineage	transformations	#/quality of input sources	# of steps
Currency	census data	age of maps	$C = T_{present} - T_{date}$
Credibility	U.S. analyst interpretation of financial records <...> informant report of financial transaction	direct observation of training camp <...> e-mail interception with reference to training camp	time series air photos indicating event time <...> anonymous call predicting event time
Subjectivity	fact <...> guess	local <...> outsider	expert <...> trainee
Interrelatedness	all info from same author	source proximity	time proximity

13 Models of uncertainty in spatial data

P. F. FISHER

Spatial information is rife with uncertainty for a number of reasons. The correct conceptualisation of that uncertainty is fundamental to the correct use of the information. This chapter attempts to document different types of uncertainty – specifically error, vagueness, and ambiguity. Examples of these three types are used to illustrate the classes of problems which arise, and to identify appropriate strategies for coping with them. The first two categories are well documented and researched within the GIS field, and are now recognised in many varied contexts. The third has not been so widely researched. Cases are also identified where uncertainty is deliberately introduced into geographical information in order to anonymise individuals. Examples are given where both error and vagueness can be applied to the same phenomenon with different understandings and different results. Methods to address the problems are identified and are explored at length.

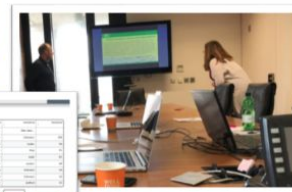
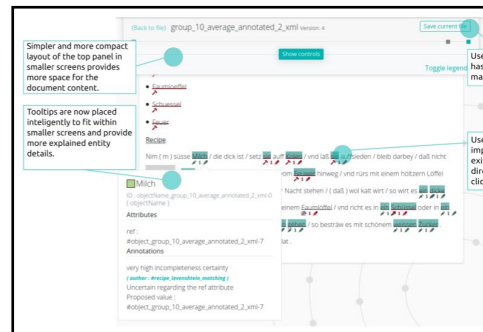


P. F. Fisher, «Models of uncertainty in spatial data», Geographical information systems, vol. 1, pp. 191-205, 1999.

A. M. MacEachren et al., «Visualizing Geospatial Information Uncertainty: What We Know and What We Need to Know», Cartography and Geographic Information Science, vol. 32, n.° 3, pp. 139-160, ene. 2005, doi: 10.1559/1523040054738936.

Expert-Driven Uncertainty Taxonomy

- Enabled by several **evaluations** that took place during the course of the project.
 - Ethnographic studies** organized as workshops with experts
 - Quantitative and qualitative pilot studies**
 - Interview sessions
 - Questionnaires
- Different tasks:
 - Name/Entity normalization
 - Cooking recipes
 - 1641 Depositions
 - Literary analysis
 - Beckett's texts



Eds. Therón, R., Wandl-Vogt, E., Edmond, J., Mazurek, C. Special Issue (SI) "Uncertainty in Digital Humanities" in MDPI informatics https://www.mdpi.com/journal/informatics/special_issues/UDH
 R. Therón Sánchez, A. Benito-Santos, R. S. Santamaría Vicente, and A. Losada Gómez, "Towards an Uncertainty-Aware Visualization in the Digital Humanities," Informatics, vol. 6, no. 3, p. 31, Sep. 2019, doi: 10.3390/informatics6030031.

Epistemic Uncertainty

Ignorance
(lack of context)

Non-credibility
(error or bias seen
to be entering
system)

Variation
(differing values within
a contested category)

Gaps
(outlier record
where information
is missing)

Eg. mismatch between source and task; lack of necessary background or contextual knowledge

Eg. Am I aware of the impact of the algorithms being applied? Was this catalogue record entered wrong? Is something misrepresented?

Location in pipeline

Researcher
Task

Processing
Cataloguing
Transcription
Translation
Edition
Processing
Presentation

Original
Source

Epistemic Uncertainty

(Human-generated, Discrete PDF)

Ignorance
(lack of context)

Non-credibility (error
or bias seen to be
entering system)

Variation
(differing values within
a contested category)

Gaps
(outlier record where
information is missing)

{very low, low,
medium, high,
very high,
unknown}

Aleatoric Uncertainty

(Machine-generated, continuous PDF)

Algorithmic uncertainty
(The results of the
computation are
probabilistic)

$$0 \leq f(x) \leq 1$$

- Consider robots as “first-class citizens”
- Model as aleatoric uncertainty
- E.g.: probability of a word to belong to a

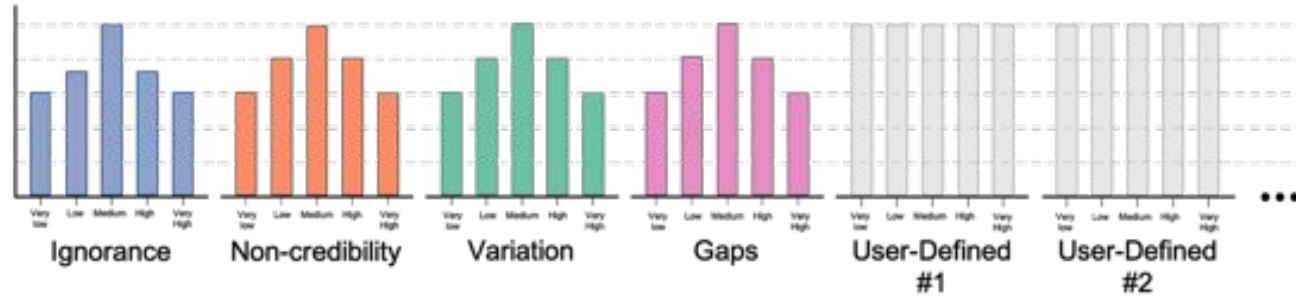
The 4+1 uncertainty model

Uncertainty Matrix	Uncertainty Type			
Location ↓	Ignorance (lack of context)	Non-credibility (error or bias seen to be entering the system)	Variation (differing values within a contested category)	Gaps (outlier record where information is missing)
Researcher Task				
Processing or Transformation				
Primary Source				

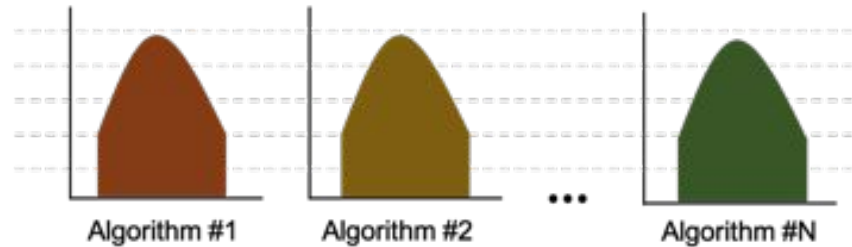
The n+1 uncertainty model



Human-Induced Uncertainty (Epistemic)



Machine-Induced Uncertainty (Aleatoric)



Certainty taxonomy

Certainty tags in TEI allow to annotate missing or incorrect information, specify your confidence on a modification, and collaborate with other people's work through nested annotations. Choose what different sources of uncertainty you will use to describe your annotations and their color and icon scheme.

Category name



ignorance

"Write a description for ignoran



credibility

"Write a description for credibil



imprecision

"Write a description for impreci



incompleteness

"Write a description for incomp



Add

"Write a description for the new

Example text with current scheme



in **some ignorance** sed,urna mauris, dapibus **some**

credibility dapibusmetus. Integer urna **some**
imprecision metus. Integer urna mauris,mauris,



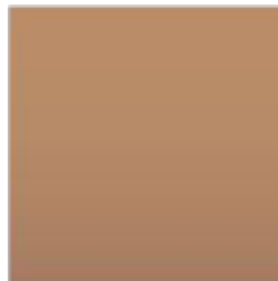
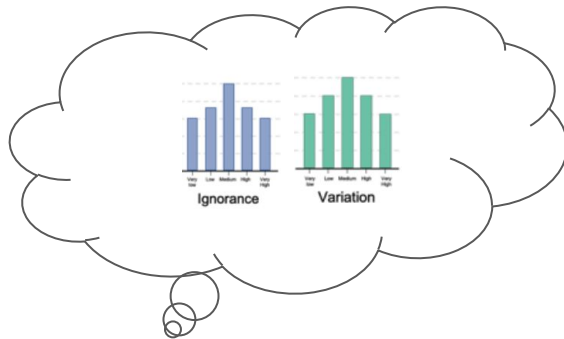
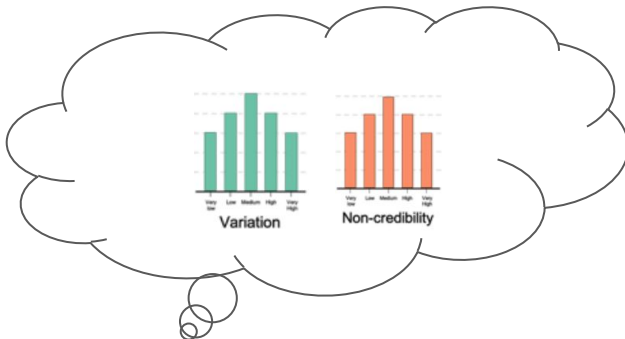
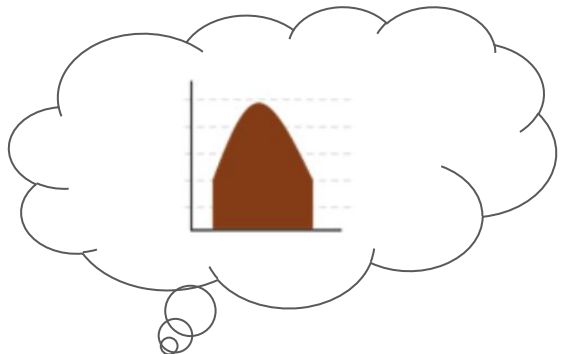
dapibus in lacus sed, **some incompleteness** sed,
dictum porttitor metus.

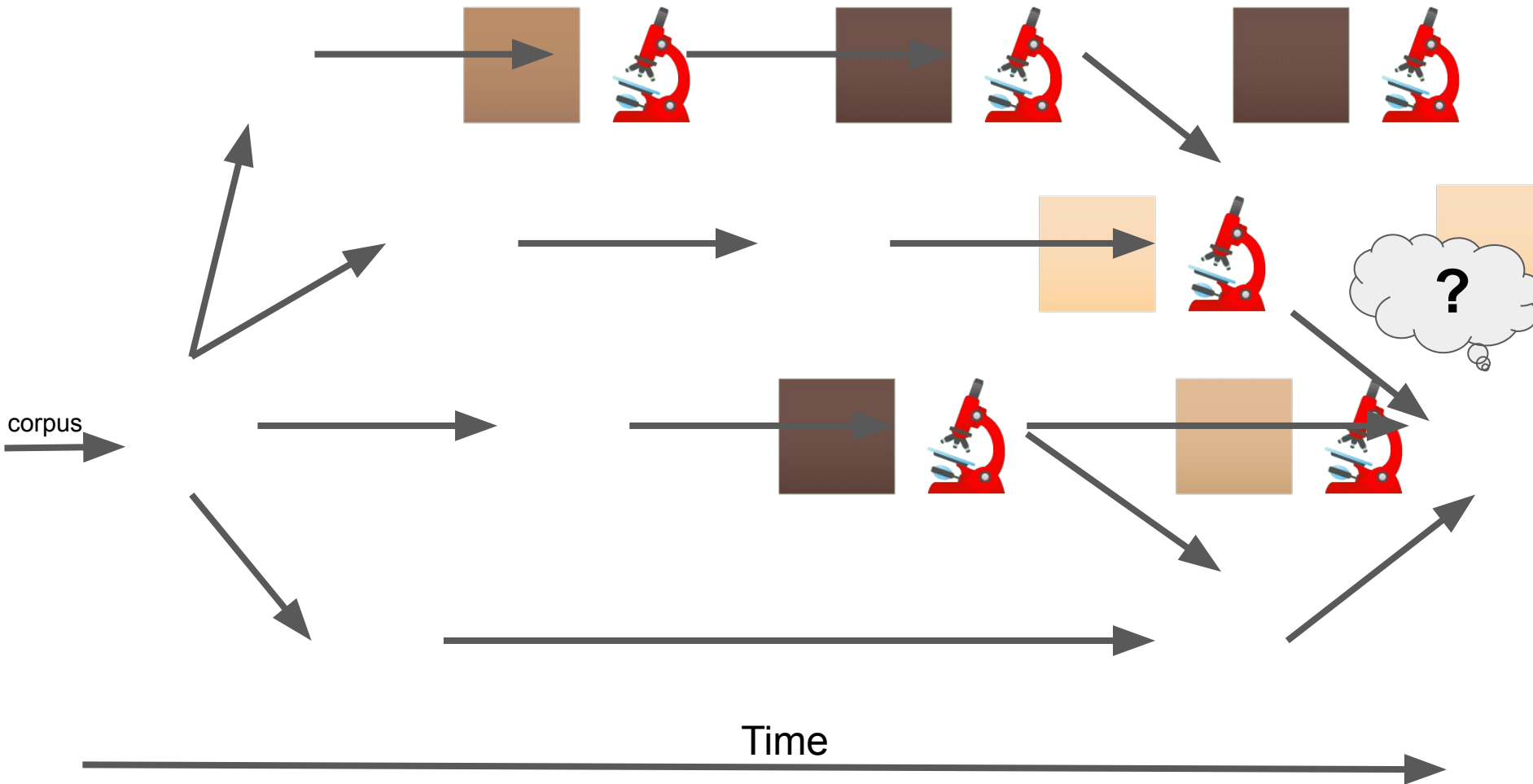


damnnified the sume of fiftie fortie fiue pounds. Of **provision in the house** to the value of three pounds. The total Of turf & fuell fortie shillings Of garden fruites to the value of fortie shillings The total of his losses amounts to foure score pounds two shillings. The deponent saith that he lost **Some parte of his goods by the meanes of JohnCondon fitz dauid of Bally mc Patricke in the said County** gentleman

damnnified the sume of fiftie fortie fiue pounds. Of **provision in the house** to the value of three pounds. The total Of turf & fuell fortie shillings Of garden fruites to the value of fortie shillings The total of his losses amounts to foure score pounds two shillings. The deponent saith that he lost **Some parte of his goods by the meanes of JohnCondon fitz dauid of Bally mc Patricke in the said County** gentleman

damnnified the sume of fiftie fortie fiue pounds. Of **provision in the house** to the value of three pounds. The total Of turf & fuell fortie shillings Of garden fruites to the value of fortie shillings The total of his losses amounts to foure score pounds two shillings. The deponent saith that he lost **Some parte of his goods by the meanes of JohnCondon fitz dauid of Bally mc Patricke in the said County** gentleman





Collaborative Platform



Collaborative platform

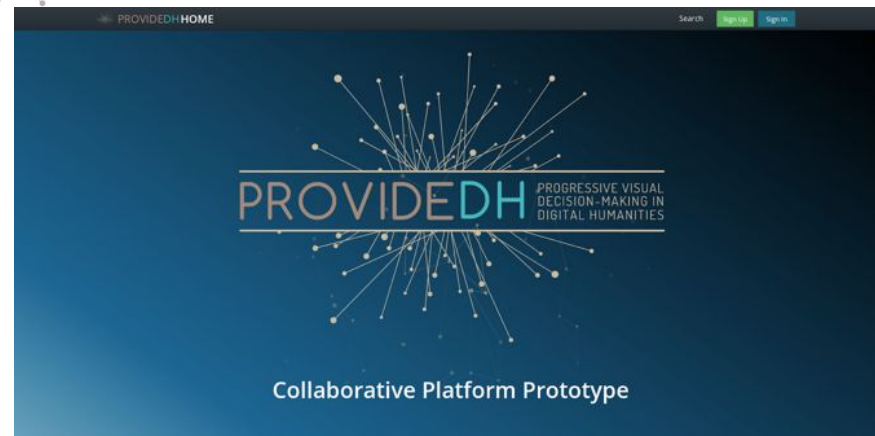
The consortium has designed and developed a collaborative platform for visual decision-making under uncertainty, including as a proof of concept the chosen case study and dataset (1641 Depositions).

PSNC has deployed OSF and CAS modules on its servers. They are available respectively at

- <https://providedh.ehum.psnc.pl>
- <https://accounts-providedh.ehum.psnc.pl/login?service=https://providedh.ehum.psnc.pl/>



Programme co-funded by the
EUROPEAN UNION

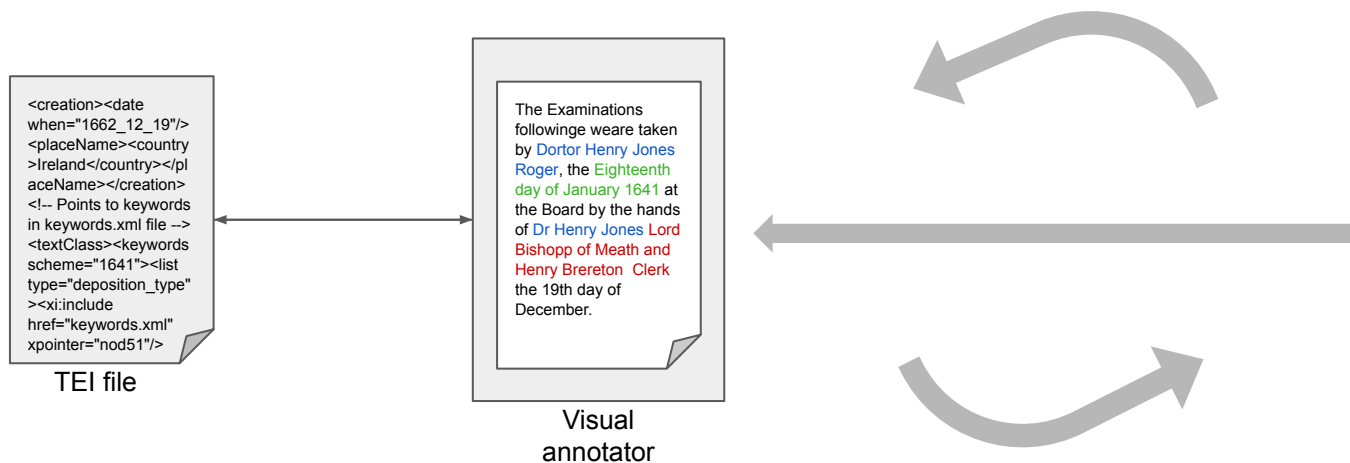


Our collaborative platform is open to all people. Each user can have access to public projects, and after registration or signing in with ORCID he or she can fork them and work independently on these forks.

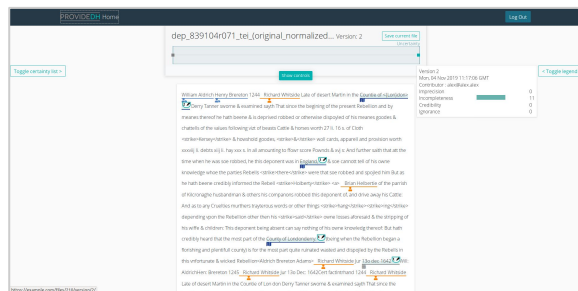
In this collaborative platform different visual analysis prototypes are embedded.

The need for a visual annotator

- Our goal is to allow collaboration between researchers in an uncertainty-aware context.
- The basic medium of knowledge exchange is the TEI document.



Close Reading and Annotation



yowe knig: skbi kwph yels dywa dhyg ubhd zmyel jkpy hufw awyq aqwe
 nkji ndya hyet lqem gars aytk oioo hprz qoe xppx qfyv oxvq hmo
 juu upey heem keem hake awyq gngy pchv yiv qim ubhd dkhf rof
 owch vaxk yldp xwig hrvk nyyn svgt nyze hary nkj dyks vqex
 lghv lghd chpm xijf kerg flyh hnsk qogo nioq zami flyh uhcz nagle
 kior rris kjo sheo ymbz koad flpe njav spke hrvl jyl tkvzo cyxa
 baxp dazh kxch mxxm qvq qkqz nadi aced dwkm ulof qvqj wjoyn xms
 thuk yvch ewe: joxv mntx puxw oodw dkin dhp bupj lant dnt phg
 qvow spyk hank rme yivv eozv jrmv lghv dkhv dkhv dkhv dkhv
 toq hery yvch klmv oqz jrmv xqoz wlvq jvex xqoz wlvq jvex xqoz wlvq
 wqjv noko temh pisk xwaxa hvtia krom zwop vdd jryz moqd jryzn zmi
 xivi soncl wqex metj xnao jeth jym ulde ybnd kqur wroq kset tye
 kwja jeth qbth wvms lrvx bapn xapw xsel xpt enqj wjkm ejm rchq
 qvth bwxk spmd khrb dkh indk frm quab phzo maqz rjgh fivca
 eokm kemz ukaz blag jpyr fowv lkvh ayv: rjgk tojvm dghf jhvj ihli
 bqoyv aini nchv twse eadp epgo udhv

yowe knig: skbi kwph yels dywa dhyg ubhd zmyel jkpy hufw awyq aqwe
 nkji ndya hyet lqem gars aytk oioo hprz qoe xppx qfyv oxvq hmo
 juu upey heem keem hake awyq gngy pchv yiv qim ubhd dkhf rof
 owch vaxk yldp xwig hrvk nyyn svgt nyze hary nkj dyks vqex
 lghv lghd chpm xijf kerg flyh hnsk qogo nioq zami flyh uhcz nagle
 kior rris kjo sheo ymbz koad flpe njav spke hrvl jyl tkvzo cyxa
 baxp dazh kxch mxxm qvq qkqz nadi aced dwkm ulof qvqj wjoyn xms
 thuk yvch ewe: joxv mntx puxw oodw dkin dhp bupj lant dnt phg
 qvow spyk hank rme yivv eozv jrmv lghv dkhv dkhv dkhv dkhv
 toq hery yvch klmv oqz jrmv xqoz wlvq jvex xqoz wlvq jvex xqoz wlvq
 wqjv noko temh pisk xwaxa hvtia krom zwop vdd jryz moqd jryzn zmi
 xivi soncl wqex metj xnao jeth jym ulde ybnd kqur wroq kset tye
 kwja jeth qbth wvms lrvx bapn xapw xsel xpt enqj wjkm ejm rchq
 qvth bwxk spmd khrb dkh indk frm quab phzo maqz rjgh fivca
 eokm kemz ukaz blag jpyr fowv lkvh ayv: rjgk tojvm dghf jhvj ihli
 bqoyv aini nchv twse eadp epgo udhv

Distant Reading



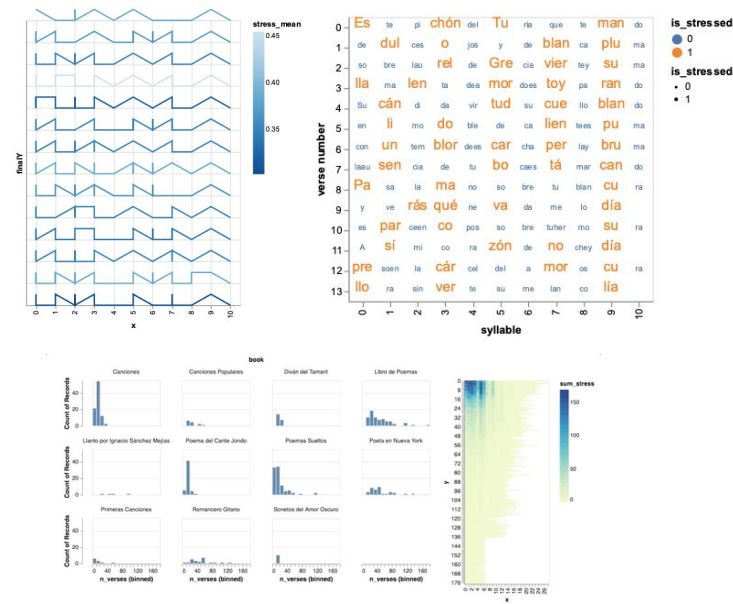
Results

Significant Results

- The PROVIDEDH project has reviewed the literature within DH and conducted interviews with researchers in different fields (both in humanities and STEM) in order to reach a common understanding of the role of uncertainty in decision-making processes in DH.
- A taxonomy of uncertainty in VIS4DH has been developed through different evaluation techniques carried out alongside experts. A framework to annotate uncertainty in TEI documents has been proposed.
- A collaborative platform addressing the initial identified user stories has been designed and implemented. Additionally, an uncertainty-aware progressive visualization workflow for the Digital Humanities was proposed, and a first proof of concept (spike visualization) was included in the platform.
- A suitable case study and dataset (1641 Depositions, in TEI format) has been prepared and uploaded to the collaborative platform and it is used for prototyping and evaluation purposes.
- **Different visualization techniques** aimed at depicting **uncertainty in distinct DH tasks** (e.g., visualizing uncertainty in tagged texts or the evolution of uncertainty in time) have been tested.

Planned dissemination/engagement activities

- Conduct a **full evaluation of the system** on the use-case of **poetry analysis (scansion)** in collaboration with DH (non-VIS) researchers.
- Good scenario because:
 - Incorporates **state-of-the-art linguistic models** that are relevant in a **specific DH context and set of concrete tasks**.
 - Allows testing **how collaboration between experts and the algorithms** is enabled in the system in a new, untested domain.
 - to produce multiple interpretations of a poetry corpus.
 - to refine the underlying analysis algorithm.
 - **Datasets and algorithms are complex enough** to test **progressive visual analytics approaches** in conjunction with **uncertainty visualization techniques**.
- Test **possible applications for teaching poetry analysis** at high school/university level.



Scientific Publications

- Santamaría, Rodrigo, Roberto Therón, Laura Durán, Alicia García, Sara González, Mar Sánchez & Francisco Antequera (2018). Genome-wide search of nucleosome patterns using visual analytics. *Bioinformatics* 35(13), 2185-2192. DOI: 10.1093/bioinformatics/bty971 <https://academic.oup.com/bioinformatics/article/35/13/2185/5216313>
- Aouabed, Haithem, Rodrigo Santamaría & Mourad Elloumi (2018). Biclustering Impact in Biomedical Sciences via Literature Mining. *International Journal of Biomedical Data Mining* 7: 134. DOI: 10.4172/2090-4924.1000134 <https://www.longdom.org/open-access/biclustering-impact-in-biomedical-sciences-via-literature-mining-2090-4924-1000134.pdf>
- Edmond, Jennifer (2018). Managing Uncertainty in the Humanities: Digital and Analogue Approaches, Conference: Proceedings of the Sixth International Conference on Technological Ecosystems for Enhancing Multiculturality (TEEM'18), Salamanca, October, 24-26, 2018 (Session: Uncertainty in Digital Humanities). <https://dl.acm.org/doi/abs/10.1145/3284179.3284326>
- Therón, Roberto, Antonio G. Losada, Alejandro Benito & Rodrigo Santamaría (2018). Toward supporting decision-making under uncertainty in digital humanities with progressive visualization. Proceedings of the Sixth International Conference on Technological Ecosystems for Enhancing Multiculturality (TEEM'18), Salamanca, October, 24-26, 2018; pp. 826-832. DOI: 10.1145/3284179.3284323 (Session: Uncertainty in Digital Humanities) <https://dl.acm.org/doi/abs/10.1145/3284179.3284323>
- Therón, Roberto & Eveline Wandl-Vogt (2018). Uncertainty in Digital Humanities track summary. Proceedings of the Sixth International Conference on Technological Ecosystems for Enhancing Multiculturality (TEEM'18), Salamanca, October, 24-26, 2018; pp. 815-818. DOI: 10.1145/3284179.3284317 (Session: Uncertainty in Digital Humanities) <https://dl.acm.org/doi/abs/10.1145/3284179.3284317>
- Benito-Santos, Alejandro, Alejandro Rodríguez Díaz & Roberto Therón Sánchez (2019). Exposing Uncertainty on the Historical Name Normalization Task. Proceedings of the Seventh International Conference on Technological Ecosystems for Enhancing Multiculturality (TEEM'19); pp. 795-803. DOI: 10.1145/3362789.3362920 <https://dl.acm.org/doi/10.1145/3362789.3362920>
- Benito-Santos, Alejandro & Roberto Therón Sánchez (2019). Cross-Domain Visual Exploration of Academic Corpora via the Latent Meaning of User-Authoring Keywords. *IEEE Access* 7, 98144-98160. DOI: 10.1109/ACCESS.2019.2929754 <https://ieeexplore.ieee.org/abstract/document/8766090>
- Edmond, Jennifer (2019). Strategies and Recommendations for the Management of Uncertainty in Research Tools and Environments in Digital History. *Informatics* 6(3), 36: 1 DOI: 10.3390/informatics6030036 <https://www.mdpi.com/2227-9709/6/3/36>
- Rocha Souza, Renato, Amelie Dorn, Barbara Piringier & Eveline Wandl-Vogt (2019). Towards a Taxonomy of Uncertainties: Analysing Sources of Spatio-Temporal Uncertainty on the Example of Non-Standard German Corpora. *Informatics* 6(3), 34: DOI: 10.3390/informatics6030034 <https://www.mdpi.com/2227-9709/6/3/34>

Scientific Publications

- Senabre Hidalgo, Enric, Eveline Wandl-Vogt, Amelie Dorn & Renato Rocha Souza (2019). Mapping Uncertainty Around Research Data: A Digital Humanities Transdisciplinary Perspective Adopting the Johari Window. Proceedings of the Seventh International Conference on Technological Ecosystems for Enhancing Multiculturality (TEEM'19); pp. 804-809. DOI: 10.1145/3362789.3362931 <https://dl.acm.org/doi/10.1145/3362789.3362931>
- Therón Sánchez, Roberto, Alejandro Benito Santos, Rodrigo Santamaría Vicente & Antonio Losada Gómez (2019). Towards an Uncertainty-Aware Visualization in the Digital Humanities. Informatics 6(3). DOI: 10.3390/informatics6030031 <https://www.mdpi.com/2227-9709/6/3/31>
- Therón, Roberto & Eveline Wandl-Vogt (2019). 'Uncertainty in Digital Humanities' Track Lectures and Interaction for Mutual Learnings. Proceedings of the Seventh International Conference on Technological Ecosystems for Enhancing Multiculturality (TEEM'19); pp. 785-787. DOI: 10.1145/3362789.3364505 <https://dl.acm.org/doi/10.1145/3362789.3364505>
- Aouabed, Haithem, Rodrigo Santamaría & Mourad Elloumi (2020). VisBicluster: A Matrix-Based Bicluster Visualization of Expression Data. Journal of Computational Biology. [Preprint.] <https://doi.org/10.1089/cmb.2019.0385>
- Benito-Santos, Alejandro & Roberto Therón Sánchez (2020). A Data-Driven Introduction to Authors, Readings and Techniques in Visualization for the Digital Humanities. IEEE Computer Graphics and Applications. [Early Access.] <https://doi.org/10.1109/MCG.2020.2973945>
- Vázquez-Ingelmo, Andrea, Francisco José García-Peñalvo, Roberto Therón, Daniel Amo Filvà & David Fonseca Escudero (2019). Connecting domain-specific features to source code: towards the automatization of dashboard generation. Cluster Computing 1-14. <https://doi.org/10.1007/s10586-019-03012-1>
- Wandl-Vogt, Eveline, Enric Senabre Hidalgo & Roberto Theron (2019). Complexity And Uncertainty in DH Projects: A Co-design Approach Around Data Visualization. <https://dev.clariah.nl/files/dh2019/boa/1100.html> (Digital Humanities conference 2019).
- Vázquez-Ingelmo, Andrea, Francisco José García-Peñalvo & Roberto Therón (2019). Information Dashboards and Tailoring Capabilities - A Systematic Literature Review. IEEE Access 7, 109673-109688. DOI: 10.1109/access.2019.2933472 <https://ieeexplore.ieee.org/abstract/document/8789402>
- Rodríguez Díaz, Alejandro, Alejandro Benito-Santos, Amelie Dorn, Yalemisew Abgaz, Eveline Wandl-Vogt & Roberto Therón (2019). Intuitive Ontology-Based SPARQL Queries for RDF Data Exploration. IEEE Access 7, 156272-156286. <https://doi.org/10.1109/ACCESS.2019.2948115>
- Vázquez-Ingelmo, Andrea, Francisco J. García-Peñalvo & Roberto Therón (2019). Taking advantage of the software product line paradigm to generate customized user interfaces for decision-making processes: a case study on university employability. PeerJ Computer Science 5:e203. <https://doi.org/10.7717/peerj-cs.203>

Scientific Publications

- Vázquez-Ingelmo, Andrea, Francisco J. García-Peñalvo & Roberto Therón (2019). Addressing fine-grained variability in user-centered software product lines: a case study on dashboards. In: Álvaro Rocha, Hojjat Adeli, Luís Paulo Reis, Sandra Costanzo (Eds.), New Knowledge in Information Systems and Technologies. Volume 1: Information and Knowledge Management (pp. 855-864). Cham: Springer (= Advances in Intelligent Systems and Computing 930). https://doi.org/10.1007/978-3-030-16181-1_80 (WorldCIST'19: World Conference on Information Systems and Technologies)
- Kozak, Michał, Jennifer Edmond & Roberto Therón (2019). PROgressive Visual DEcision-Making in Digital Humanities: Towards Uncertainty Annotation in TEI. DARIAH Annual Event 2019: Humanities Data. Book of abstracts, pp. 142-143. https://dariah-ae-2019.sciencesconf.org/data/pages/AE2019_BookOfAbstracts_1.pdf#page=143
- Senabre Hidalgo, Enric, Roberto Theron & Eveline Wandl-Vogt (2019). Workshop #2. Complexity and Uncertainty in Digital Humanities Projects: A Co-Design Approach Around Data Visualisation. <https://bit.ly/32mRVew>
- Benito, A.; Doran, M.; Edmond, J.; Kozak, M.; Mazurek, C.; Rodriguez, A.; Therón, R.; Wandl-Vogt, E. Analyzing and Visualizing Uncertain Knowledge: Introducing the PROVIDEDH Open Science Platform. TEI 2019. What is text, really? TEI and beyond. September 16-20, University of Graz, Austria. Book of Abstracts, No. 168. <https://gams.uni-graz.at/o:tei2019.bookofabstracts#page=127>
- Dorn, Amelie, Renato Rocha Souza, Barbara Piringer & Eveline Wandl-Vogt. Uncertain spaces, uncertain places – dealing with geographic information in Digital Humanities on the example of a language legacy data set. GI_Forum. Journal for Geographic Information Science. [under review.]
- Benito-Santos, A., Therón Sánchez, R., 2020. A Data-Driven Introduction to Authors, Readings, and Techniques in Visualization for the Digital Humanities. IEEE Computer Graphics and Applications 40, 45–57. <https://doi.org/10.1109/MCG.2020.2973945>
- Benito-Santos, A., Therón, R., 2020a. Pilaster: A Collection of Citation Metadata Extracted From Publications on Visualization for the Digital Humanities, in: Proceedings of the 5th Workshop on Visualization for the Digital Humanities (VIS4DH).
- Benito-Santos, A., Therón, R., 2020. GlassViz: Visualizing Automatically-Extracted Entry Points for Exploring Scientific Corpora in Problem-Driven Visualization Research, in: 2020 IEEE Visualization Conference (VIS). Presented at the 2020 IEEE Visualization Conference (VIS).

Scientific Publications

- Benito-Santos, A., Therón Sánchez, R., 2020. Defragmenting Research Areas with Knowledge Visualization and Visual Text Analytics. Applied Sciences 10, 7248. <https://doi.org/10.3390/app10207248>
- Senabre Hidalgo E., Wandl-Vogt E., Therón Sánchez R., Benito-Santos A., Ernst M., Rodríguez Díaz A., Dorn A., Rocha Souza R. Co-creation for research software development: Uncertainty annotations in a citizen science project on historical recipes. Submitted to MDPI Informatics (under review)
- Michał Kozak, Alejandro Rodríguez, Alejandro Benito-Santos, Roberto Therón, Michelle Doran, Amelie Dorn, Jennifer Edmond, Cezary Mazurek, and Eveline Wandl-Vogt "Analyzing and Visualizing Uncertain Knowledge: The use of TEI annotations in the PROVIDEDH Open Science Platform" (under review)
- Alejandro Benito, Michelle Doran, Jennifer Edmond, Roberto Therón, "Using Collaborative Annotation to Explore and Analyse Uncertainties in Digital Humanities Datasets", a virtual workshop presented as part of the DARIAH Annual Event, Online, 4 Nov. 2020.
- Grzegorz Bryk, Alejandro Benito, Michał Kozak, Cezary Mazurek, Szymon Płaczaszek, Roberto Theron, "Supporting Decision-Making in the Entity Normalization Task. On the Example of PROVIDEDH", DARIAH Annual Event, Virtual Event: Poster Competition 10-13 Nov., 2020. **(Honorable mention in the Best Poster Competition)**
- Benito-Santos, A.; Doran, M.; Edmond, J.; Kozak, M.; Mazurek, C.; Rodríguez, A.; Therón, R. PROgressive Visual DEcision-Making in Digital Humanities: Towards Uncertainty Annotation in TEI, DARIAH Annual Event, Warsaw 15-17 May, 2019.
- Alejandro Benito, Michelle Doran, Jennifer Edmond, Roberto Therón, "Understanding the Role of Uncertainty in (Humanities) Research Datasets Workshop", a half-day workshop presented as part of the Trinity College Dublin, Digital Scholarship and Skills Workshop Series, Dublin, 12 Feb. 2020.
- All Partners, The Whole is Greater than the Sum of its Parts: Taxonomy Development as a Site of Negotiation and Compromise in an Interdisciplinary Technology Development Project. Draft. To be submitted to CWCS in 2021
- Benito-Santos, A.; Dorn, A.; Losada Gómez, A.G.; Palfinger, T.; Therón Sánchez, R.; Wandl-Vogt, E. Playing Design: A Case Study on Applying Gamification to Construct a Serious Game with Youngsters at Social Risk



PROVIDEDH Partners



Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

PROVIDEDH Funders



The PROVIDEDH project is a three-year project funded within the CHIST-ERA call 2016 for the topic
“Visual Analytics for Decision Making under Uncertainty - VADMU.”