Methods, Tools and Case Studies for User Interaction Modelling

Markel Vigo

Human-Computer Systems Group

Department of Computer Science

markel.vigo@manchester.ac.uk

www.markelvigo.info

@markelvigo

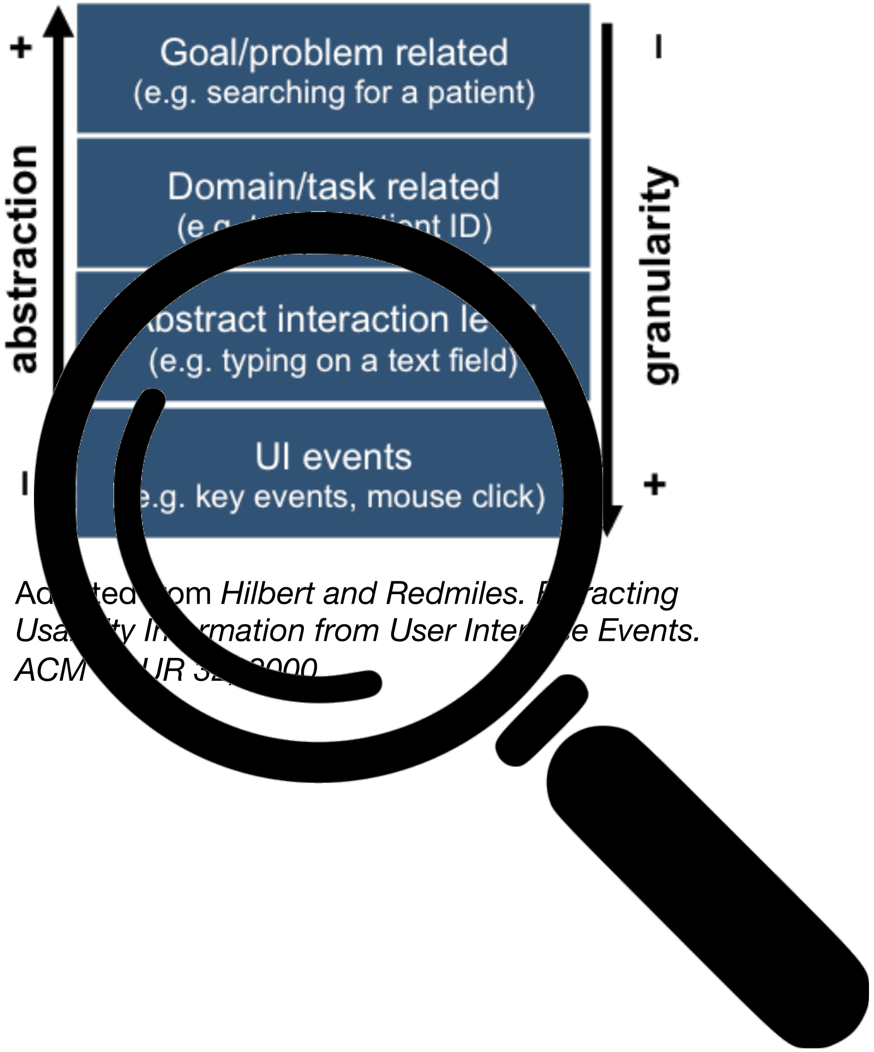
MANCHESTER
1824
The University of Manchester

IAMlab
Interaction Analysis and Modelling Laboratory

Goal

- Understand user behaviour on data intensive systems
- Extract objective and actionable understanding of systems in use
- To inform the design of user interfaces, digital health interventions, workflows and data governance policies
- Design pipelines to harvest, store, transform, query and analyse human-computer interactions
- Build computational user models to predict
  - Interest
  - Engagement
  - Interaction barriers
  - Knowledge acquisition

Scope: low-granularity interaction data



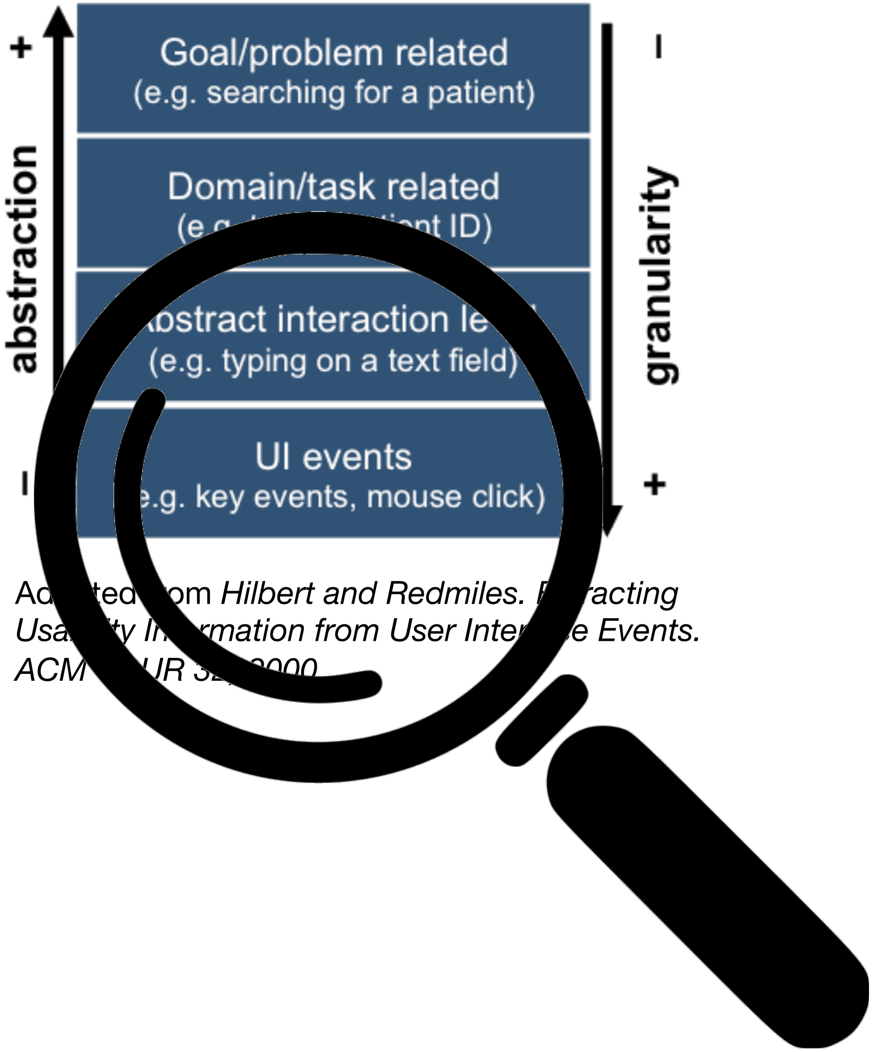
Adapted from Hilbert and Redmiles. Extracting Usability Information from User Interface Events. ACM CSUR 32, 2000

- It contains implicit behavioural markers that are indicators of cognitive process
  - Location of mouse cursor → attention
  - Exploration → engagement
  - Quick scroll down → information overload
- Appropriate for *in the wild* naturalistic studies

✅ High ecological validity, easy recruitment

❌ No control for tasks, objective ground truth or other factors

Scope: low-granularity interaction data

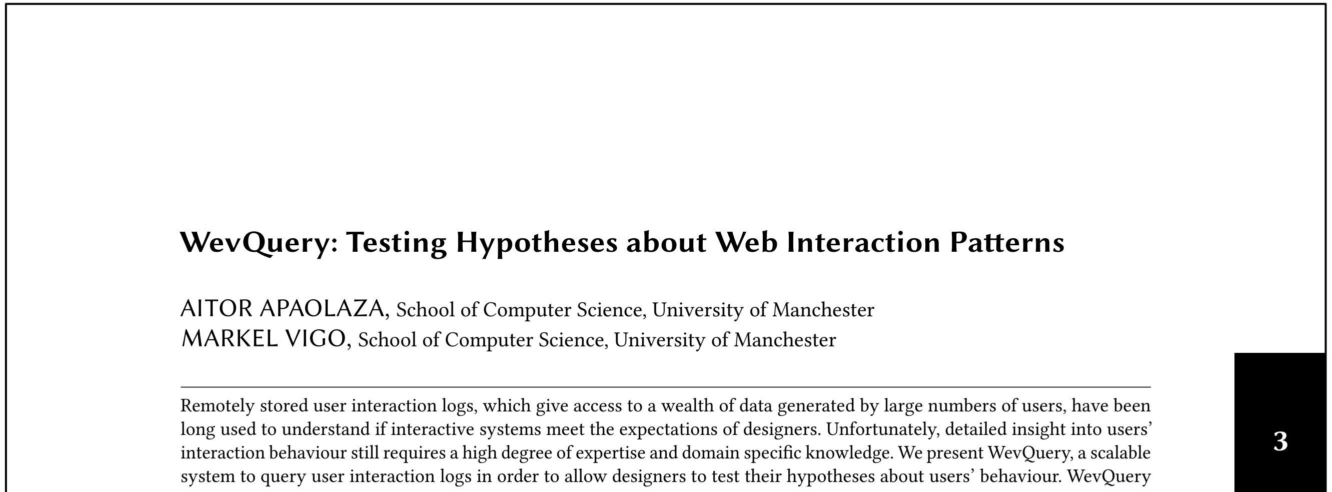


Challenges:

1. Limited semantics
2. High cardinality/noisy outputs
3. Hypothesis formulation
4. Scalability

Tool support for addressing the challenges

ACM SIGCHI Symposium on Engineering Interactive Computing Systems, **EICS 2017**



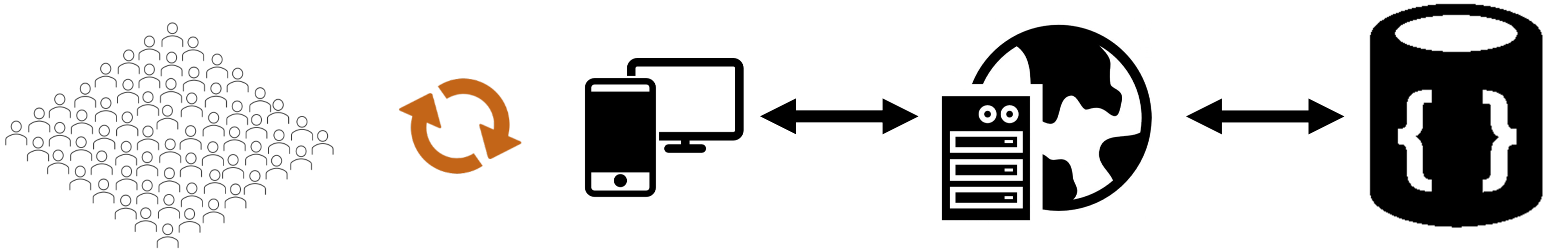
WevQuery: Testing Hypotheses about Web Interaction Patterns

AITOR APAOLAZA, School of Computer Science, University of Manchester
MARKEL VIGO, School of Computer Science, University of Manchester

Remotely stored user interaction logs, which give access to a wealth of data generated by large numbers of users, have been long used to understand if interactive systems meet the expectations of designers. Unfortunately, detailed insight into users' interaction behaviour still requires a high degree of expertise and domain specific knowledge. We present WevQuery, a scalable system to query user interaction logs in order to allow designers to test their hypotheses about users' behaviour. WevQuery

WevQuery: Addressing the challenges

1. Limited semantics: include interaction context



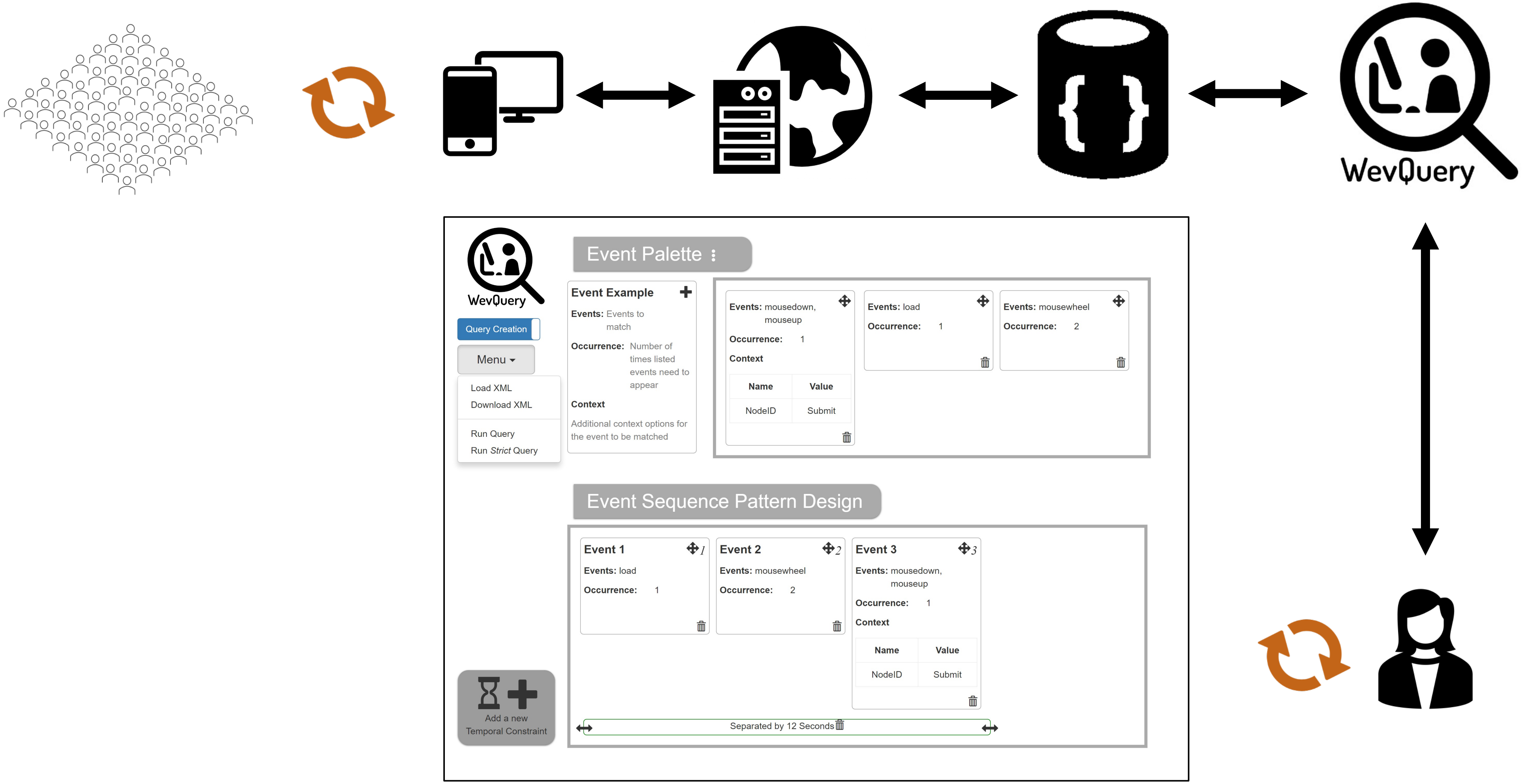

```
{ "_id" : ObjectId("51c98279e4b0978a196f12e1"),
  "mouseCoordinates" : {
    "coordX" : 379,
    "coordY" : 599,
    "offsetX" : 105,
    "offsetY" : 40},

  "nodeInfo" : {
    "nodeDom" : "id(\"content-primary\")/DIV[6]/P[1]",
    "nodeType" : "P",
    "nodeTextContent" : "Each student \nhas a(...)",
    "nodeTextValue" : "undefined"},

  "ip" : "XXX.XXX.XXX.XXX",
  "timestamp" : "2013-06-25,12:43:29:208",
  "sessionstartparsed" : "2013-06-25,12:12:47:918",
  "usertimezoneoffset" : "-60",
  "sd" : "10006",
  "sid" : "h0cBPqMBrP7c",
  "event" : "mousemove",
  "platform" : "win32",
  "browser" : "Chrome28.0.1500.44",
  "url" : "http://www.cs.manchester.ac.uk/
      undergraduate/studentprojects/"}
```

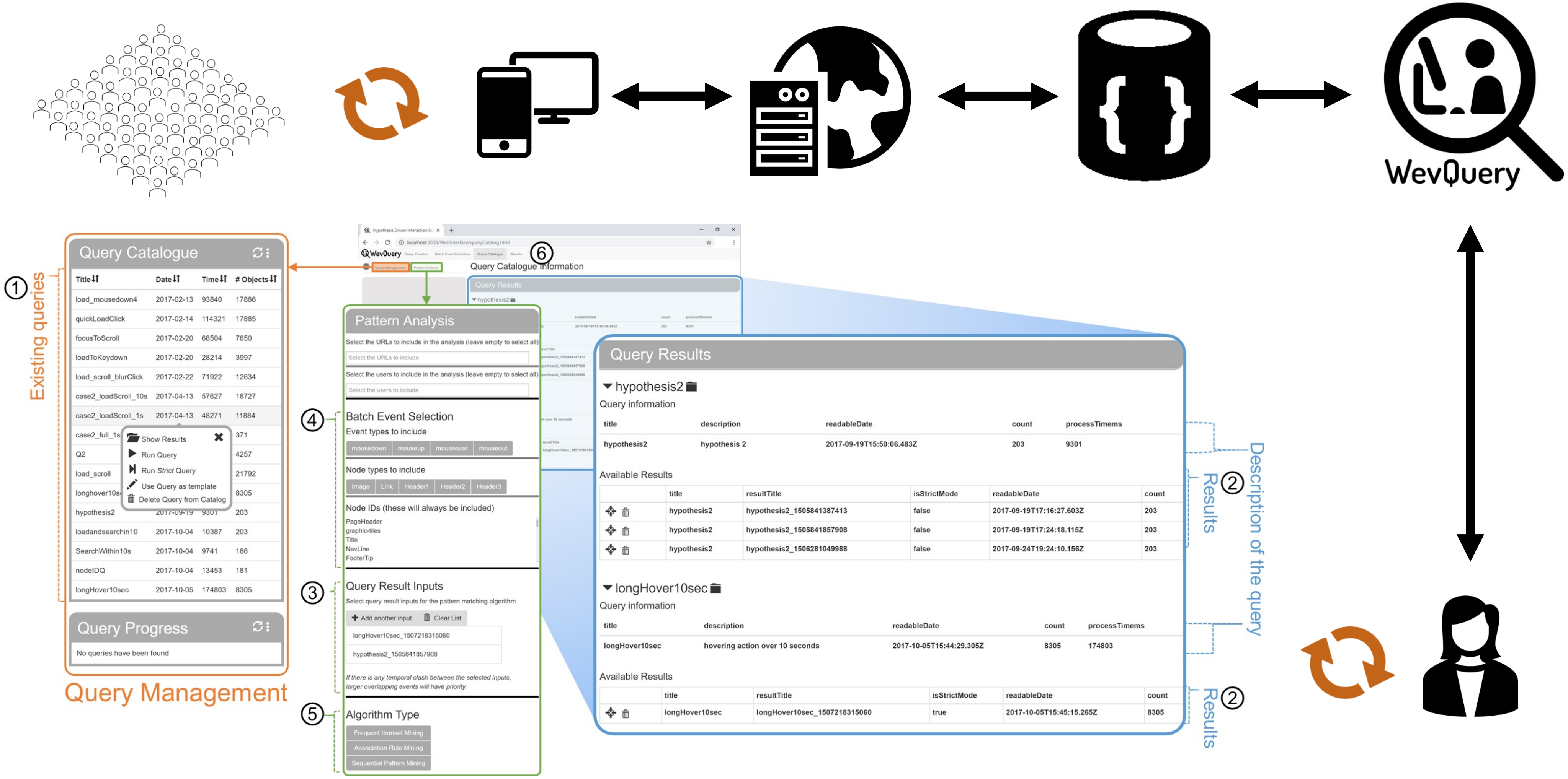
WevQuery: Addressing the challenges

2. High cardinality: transforming, subsetting and filtering



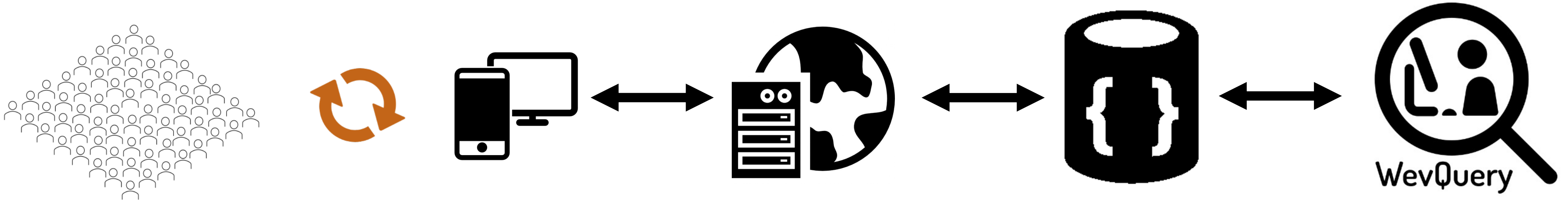
WevQuery: Addressing the challenges

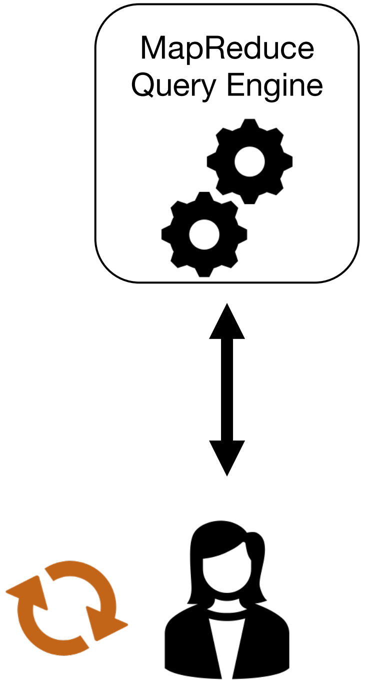
3. Hypothesis formulation

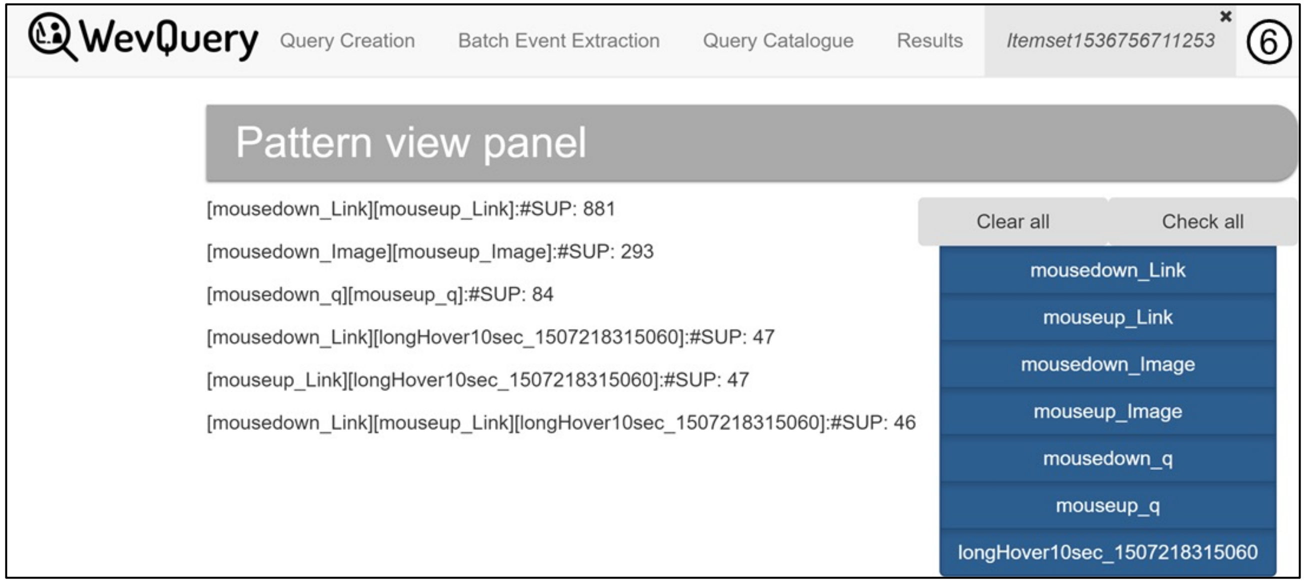


WevQuery: Addressing the challenges

4. Scalability







WevQuery: Testing Hypotheses about Web Interaction Patterns

AITOR APAOLAZA, School of Computer Science, University of Manchester
MARKEL VIGO, School of Computer Science, University of Manchester

International Journal of Human-Computer Studies 130 (2019) 196–208

Contents lists available at ScienceDirect

International Journal of Human-Computer Studies

journal homepage: www.elsevier.com/locate/ijhcs

Assisted pattern mining for discovering interactive behaviours on the web

Aitor Apaolaza, Markel Vigo*

School of Computer Science, University of Manchester, United Kingdom

ARTICLE INFO

Keywords:
Interaction logs
Assisted pattern mining
User interface evaluation

ABSTRACT

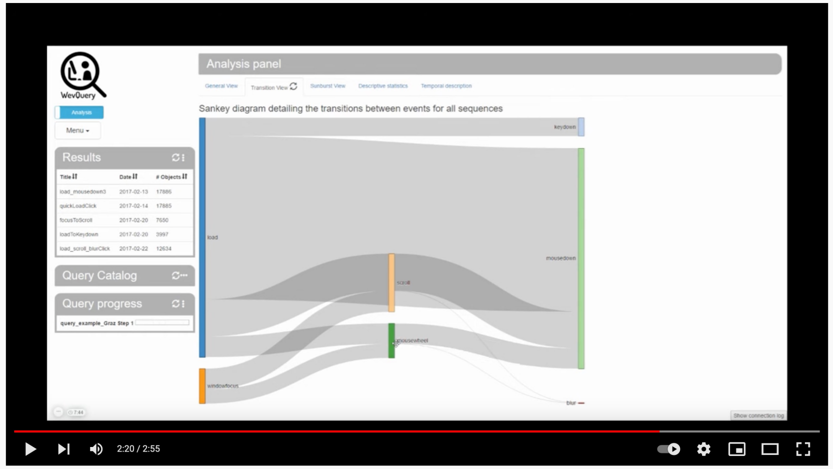
When the hypotheses about users' behaviour on interactive systems are unknown or weak, mining user interaction logs in a data-driven fashion can provide valuable insights. Yet, this process is full of challenges that prevent broader adoption of data-driven methods. We address these pitfalls by assisting user researchers in customising event sets, filtering the noisy outputs of the algorithms and providing tools for analysing such outputs in an exploratory fashion. This tooling facilitates the agile testing and refinement of the formulated hypotheses of use. A user study with twenty participants indicates that compared to the baseline approach, assisted pattern mining is perceived to be more useful and produces more actionable insights, despite being more difficult to learn.

1. Introduction

Understanding users' interaction with complex interactive systems is a challenging endeavour. While task-oriented user evaluations help to optimise the user interface elements involved in the execution of known tasks, user behaviour beyond the established boundaries of the tasks remains unknown. This pragmatism is understandable in that evalu-

surrounding interactions (Nebeling et al., 2013).

Challenge 2: Limited semantics. Raw user interaction events lack a rich context of use from which one can extract meaningful conclusions. To increase this lack of meaning, events should be associated with elements on the website and mapped into the appropriate abstraction levels (Hilbert and Redmiles, 2000; Liu et al., 2017; Perer and Wang, 2014). This would allow, for instance, to transform mouse clicks on a



Wevquery - A scalable system for testing hypotheses about web interaction patterns

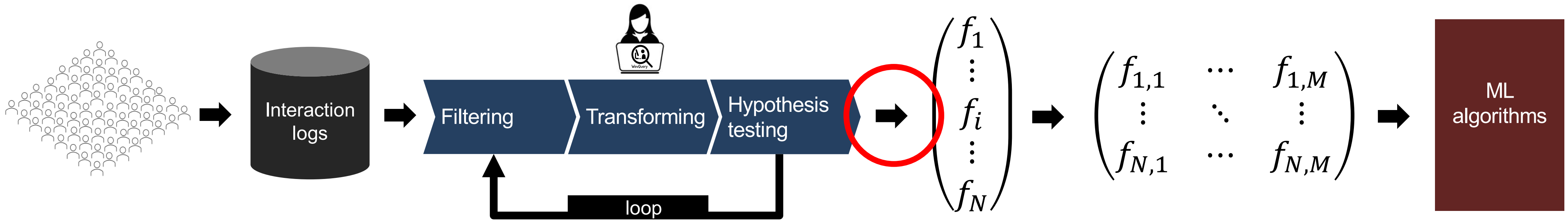
https://youtu.be/0oUD-1RzydI

GitHub

https://github.com/aapaolaza/WevQuery

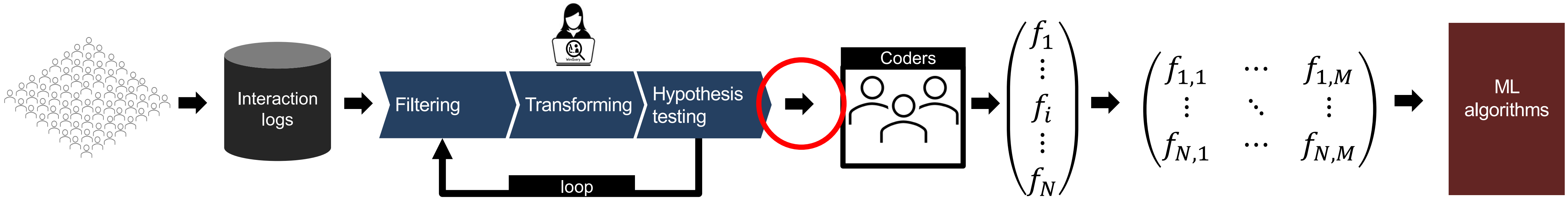
Methods (and pipelines too)

Interactive behaviours as features of user models



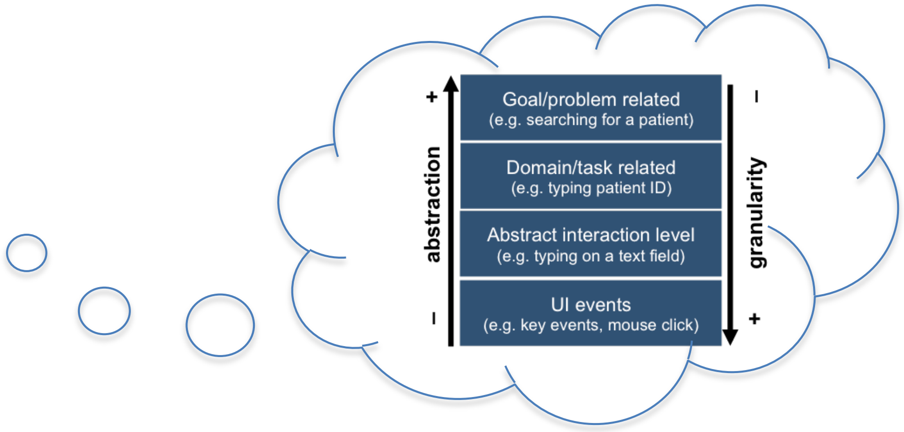
- Behaviours become features
  - Hypothesis driven vs data driven
- Computational representation of users/behaviours over time
- Data is segmented into sessions
- Features are fed into learning algorithms
- Challenges when generating features on a data-driven fashion

Interactive behaviours as features of user models



- Challenges when generating features on a data-driven fashion
  - Noisy outputs: pattern overload due to subpatterns and minor variations
  - Limited expressivity (domain experts needed)

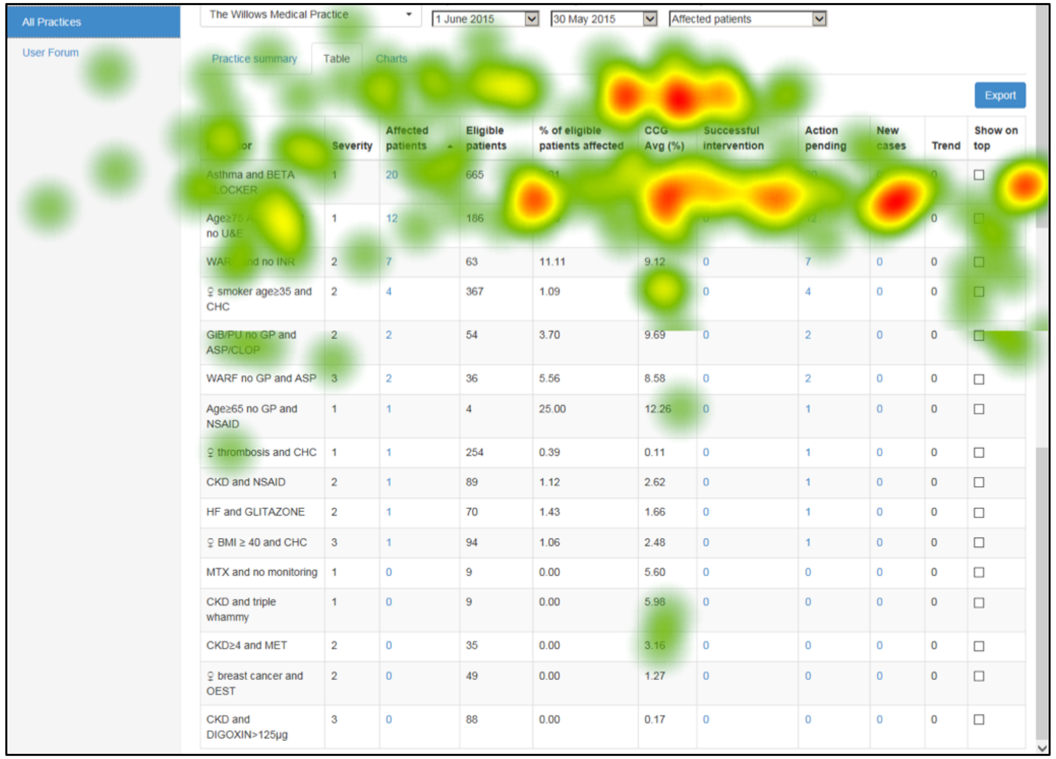
1. Further filtering: maximise expressivity while reducing idiosyncrasy
   - Longer patterns
   - Frequent patterns
2. Further grouping: thematic analysis of patterns
   - Treat patterns as codes
   - Humans in the loop generating themes

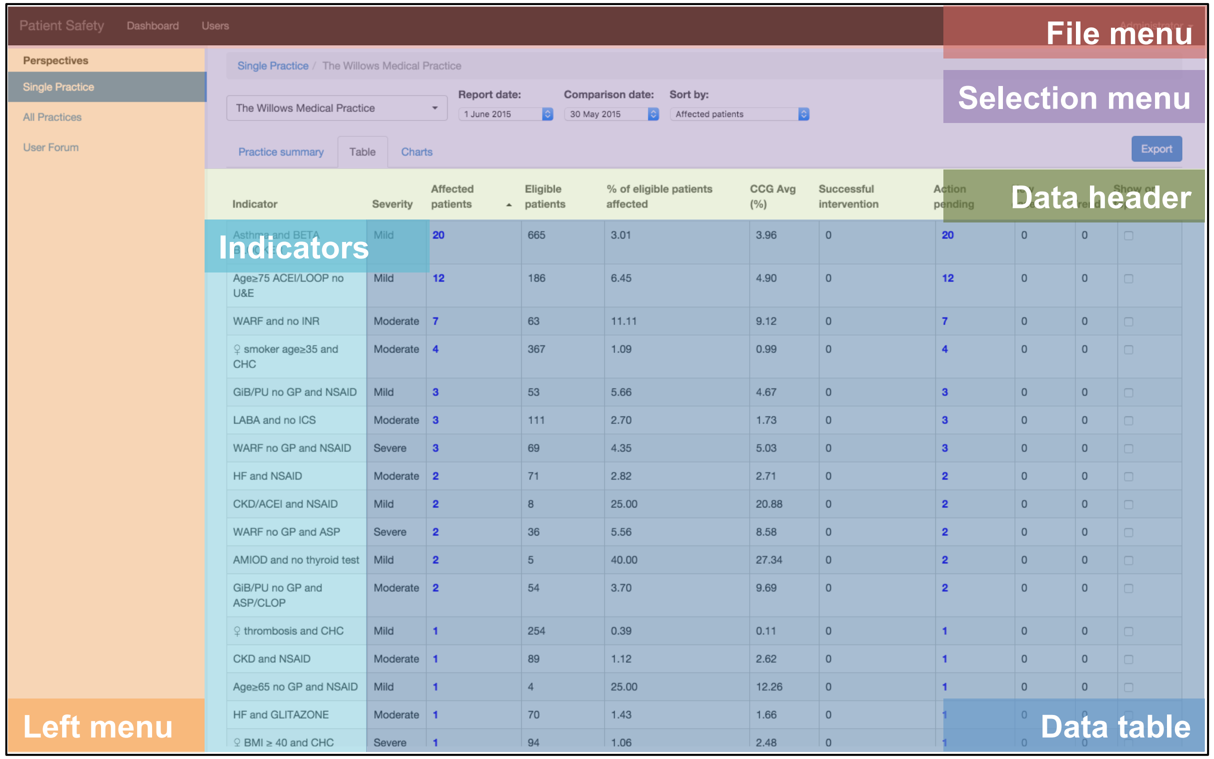


Case studies

Medication safety dashboard

- Goal: characterise the use of clinical pharmacists vs non primary users
- Hypothesis driven features: dwell time, mouse hovers between clicks
- Supervised learning
- N=35, 10-months

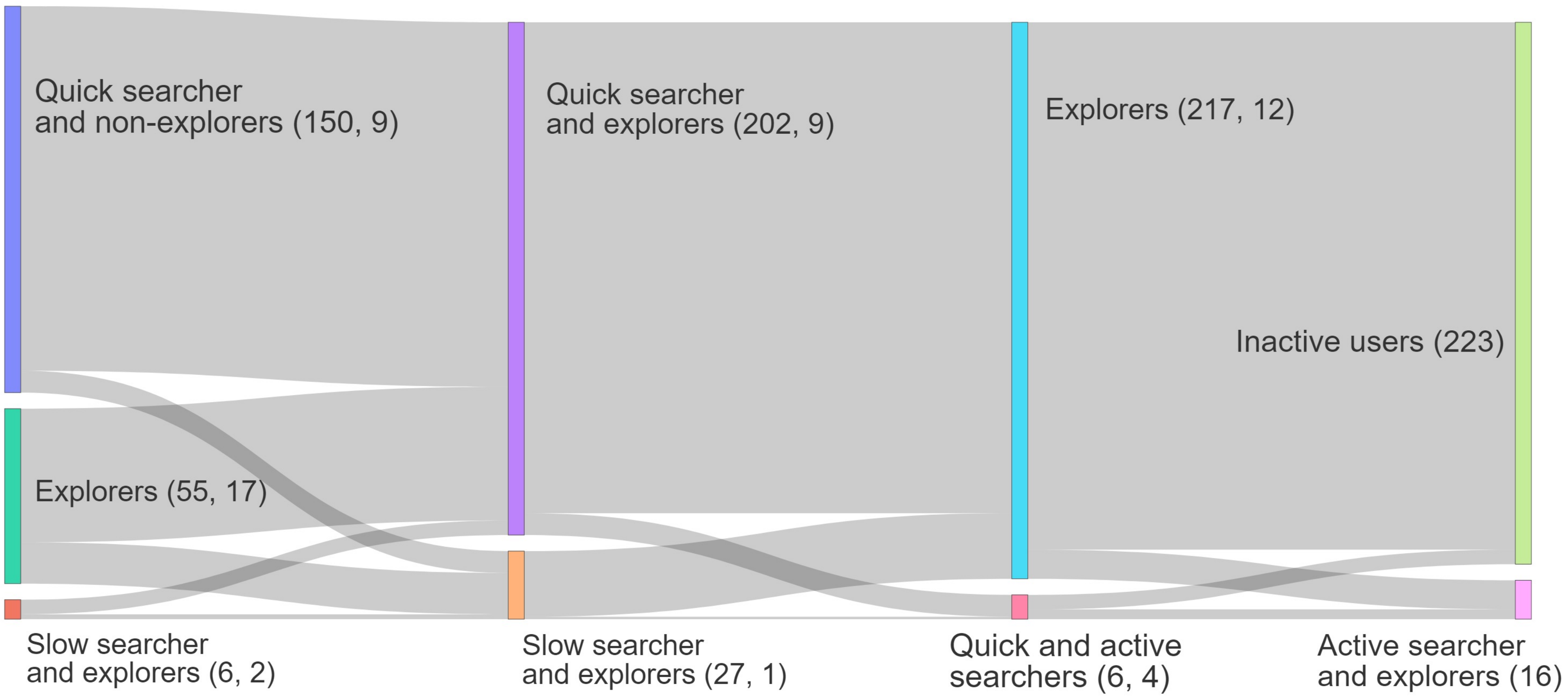




➡️ Yera, Muguerza, Arbelaitz, Perona, Keers, Ashcroft, Williams, Peek, Jay, Vigo (2019) *Modelling the interactive behaviour of users with a medication safety dashboard in a primary care setting*. International Journal of Medical Informatics 129, https://doi.org/10.1016/j.ijmedinf.2019.07.014

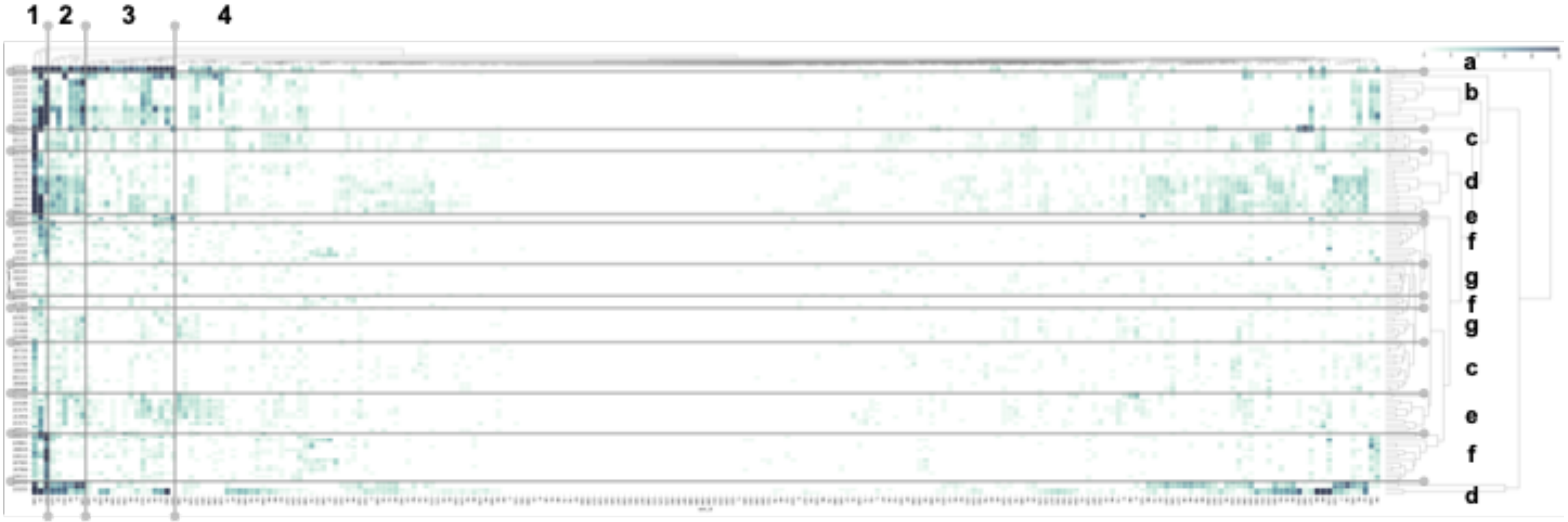
Behaviour evolution on a specialist search engine

- Goal: monitor search and exploration behaviours longitidinally
- Hypothesis driven features: clicks, scroll, dwell time, search terms…
- Unsupervised learning
- N=239, 20-months



Engagement on a cMOOC

- Goal: engagement patterns of early career researchers on a cMOOC
- Data driven features: 130 activity patterns
- Unsupervised learning
- N=224, 4 weeks



➡️ Gledson, Apaolaza, Barthold, Günther, Yu, Vigo (2021) *Characterising Student Engagement Modes through Low-Level Activity Patterns*. ACM Conference on User Modelling, Adaptation and Personalization, UMAP 2021. https://doi.org/10.1145/3450613.3456818

Achievement on a cMOOC

- Goal: associate patterns of use with achievement in online learning
- Data driven features: 23 activity patterns
- Supervised learning
- N=193, 4-weeks

- Assessed the data driven approach against known features
- Increase of 7% in accuracy in detecting students who are not going to achieve a badge in the course

➡️ Yu, Harper, Vigo (2021) *Modeling Micro-Interactions in Self-Regulated Learning: a Data-Driven Methodology*. International Journal of Human-Computer Studies.
https://doi.org/10.1016/j.ijhcs.2021.102625

Conclusion

- Low level interactions contain implicit cognitive markers
- Added value in using low level interactions
- Analysing low level interactions comes at a cost
- Which can be mitigated with tool support
- Specially for not data scientists
- There are opportunities to interactively explore the problem space

Methods, Tools and Case Studies for User Interaction Modelling

Questions?

Markel Vigo

Human-Computer Systems Group

Department of Computer Science

markel.vigo@manchester.ac.uk

www.markelvigo.info

@markelvigo