**Literature Review Notes (from Evernote)**

***Introduction***

* What the idea behind this document is
* How it is structured

***Sections***

*Interaction*

*Engagement*

*Sequential Data Mining (or other data mining-type techniques?)*

***Longitudinal analysis of low-level web interaction through micro behaviour [interaction - done]***

* A remote capture solution is developed, using a well-established solution as a foundation, to record data on a high-traffic website continuously for two years.
* The data that is collected is low-level; mouse movements (in the form of coordinates), mouse clicks, keystrokes, etc.
* In their analysis, they group the large amounts of interaction data into a temporal perspective.
* They find that rather than users interacting the website quicker as they become more familiar that users actually have increased periods of mouse inactively. Furthermore, the users also spend more time on the website as they become more familiar.
* In this paper, the capture solution presented in [UsaProxy] is modified to allow deployment by adding JavaScript code to the web pages rather than requiring users to set their browser to re-route all connections through a proxy server.

***Understanding users in the wild [interaction - done]***

* They find that there is no need to collect specific information about users, such as their disabilities, as their problems can be identified through emerging behaviours in the experiments.
* *Group these two together.*

The capture solution presented in [usa-proxy] is modified in [cite] to allow deployment by adding JavaScript code to the web pages rather than requiring users to set their browser to re-route all connections through a proxy server. Data; low-level mouse movements, clicks, and keystrokes, in this experiment is recorded from a high-traffic website continuously for two years. They find that users, rather than interacting with the website quicker as they become more familiar, have increased periods of mouse inactivity. Continually, the users also spend more time on the website as they become more familiar. And finally, they find that there is no need to collect specific information about users, such as any disabilities that they might have, as their problems can be identified through emerging behaviours in the experiments [cite].

***Combining mouse and keyboard events with higher level desktop actions to detect mild cognitive impairment [interaction - done]***

* A fully-fledged, desktop application with the aim of trying to detect mild cognitive impairment in older computer users through their interactions with the computer is presented in this paper.
* They monitor operating system events (deleting files through dragging to the recycling bin), web browser extension, and application monitoring (internet explorer, outlook, and word).
* They do collect mouse interactions but the complexity is reduced to mouse drags and phases (time periods between clicks).

***Combining data mining and text mining for detection of early stage dementia: the SAMs framework [interaction - done]***

* More detail is provided in this paper as to the type of data/events that are collected. The desktop logger collects events in; word, outlook, file system, mouse, keyboard, and user interface systems. Potentially providing some insight into the type of events that can, or even should – in this context, be collected. Various web-based events are also captured.
* *Group these two.*

In [cite1, cite2], a fully-fledged, desktop application with aim of trying to detect mild cognitive impairment in older computer users through their interactions with the computer is presented. The monitoring system collects data on operating system events, web browsers, and applications. Furthermore, mouse movements are collected but the complexity is reduced by only recording drags and the time periods between clicks. They have early evidence that this is a promising method to detect cognitive impairments.

***Engagement: What is it, and how can we measure it? [Engagement - done]***

* They present a broad review of measuring and defining user engagement in a range of scenarios.
* The focus is on the reaction of users and to try and understand what engagement means with regards to initial reactions to media-based content.
* They find that that if the audience is emotionally invested in the content then their levels of engagement with the content is subsequently high.

A broad review of measuring and defining user engagement in a range of scenarios is presented in [cite]. The focus is on the understanding of initial reactions to media-based content and what engagement means in this context. They find that if the audience is emotionally invested in the content then their level of engagement is subsequently higher.

***Measuring and defining the experience of immersion in games (might be confusing to write this one up) [Engagement - done]***

* The paper investigates if immersion can be defined quantitatively through three specific experiments. Participants switching from an immersive to non-immersive task, changes in participants eye movements during an immersive task, and the effect of an externally imposed pace of interaction on immersion and affective measures.
* They find that immersion can be measured subjectively, through questionnaires, and objectively, through task completion and eye movements.
* They define three features of immersion; lack of awareness of time, loss of awareness of the real world, and involvement and a sense of being in the task environment.
* They apply Spearman’s Rank-Order correlation on the mouse click data (the mean number of mouse clicks vs the mean number of fixations in the non-immersive condition).

In [cite], the authors perform an investigation to test if immersion can be defined quantitatively through experiments. They devise three experiments; switching from an immersive to a non-immersive task, changes in eye movements during an immersive task, and measuring the effect of an externally imposed pace of interaction which alters the flow of the participant. Immersion is well defined here and has three features; lack of awareness of time, loos of awareness of the real world, and a sense of being in the task environment (a video game). They find that immersion can be measured subjectively, through questionnaires given to the participants before and after the event, and objectively, through task completion and eye movement. The authors apply Spearman’s Rank-Order correlation on the mouse click data (the mean number of mouse clicks vs the mean number of fixations on the non-immersive condition), relying heavily on this in their analysis.

***Knowing the user’s every move – user activity tracking for website usability evaluation and implicit interaction [interaction - done]***

* This paper defines the UsaProxy – a monitoring system for web-based interactions. It modifies HTML pages by adding JavaScript tracking code before they are delivered to the client, this is done through requesting the users to re-route all of their connections through a proxy server. The code collects data on mouse movements, keyboard input, along with other common interaction metrics.
* This is the system that Aitor’s work builds from – *so it should go before it*.

In [ref] a monitoring system for web-based interactions is defined – called UsaProxy. By requesting the users of the system to re-route all of their connections through a proxy server, HTML pages are modified with JavaScript tracking code before they are delivered to the user. The code collects data on mouse movements, keyboard input, along with other, fine-grained interaction metrics.

***Evaluating accessibility-in-use* & *Coping tactics employed by visually disabled users on the web [interaction] Markel and Harper’s work - done.***

* In this paper [evaluating], they isolate the problematic situations faced by users with visual impairments and the tactics they use to attempt to overcome issues while browsing the web. Once these tactics are identified they are considered as behavioural markers to indicate problematic situations with the assumption that these markers infer an issue the user is having. They develop several algorithms and package them together into a web-usage monitoring tool to automatically detect issues while users are using the web.
* This paper [coping] goes into more detail about the tactics discussed in [evaluating]. It also expands on previous analysis by going deeper into the coping tactics the users use; how they react to problems, do they give up or carry on.

Problematic situations encountered by users with visual impairments and the tactics they employ to overcome them are explored in [evaluating]. Through developing several algorithms, and packaging them together into a web-usage monitoring tool, the employed tactics are identified and isolated automatically, in mouse and keyboard data, and treated as markers to infer the user is having an issue. In [coping] more detail is presented about the particular type of tactics and an expansion on the analysis process by going deeper into the tactics the users employed and how they react to problems (do they give up or carry on).

***WevQuery: Testing hypotheses about web interaction patterns [interaction]***

* A scalable system to query user interaction logs that allows designers to test their hypotheses about user behaviour is presented in this paper.
* It’s the combination of the work from Markel and Aitor.
* A graphical user interface allows the users to graphically define queries to run on interaction data (mouse clicks, etc.) stored in a MongoDB instance.

WevQuery, a scalable system to query user interaction logs collected in [aitor, markel], is presented in [cite]. It is a hypothesis testing interface with the aim of better understanding user behaviour stored in the data collected (mouse clicks, movement, etc.).

***Using machine learning to infer reasoning provenance from user interaction log data: based on the data/frame theory of sense making [interaction]***

* In this paper, the authors successfully attempt to infer the reasoning behind analysts’ decision making from low-level user interaction logs.
* They detail some information about their data processing techniques; converting features to numeric expressions and a set of categorical labels to a set of integers (worked for Random Forest and Hidden Markov Models (scale-invariant models) but not SVM, so integers were standardised to have a mean of zero and a variance of one).
* They tested their models against a control model (no information classifier) and provided the baseline for their models to beat.

The authors in [cite] are successfully able to infer the reasoning behind analysts’ decision making from low-level user interaction logs. They detail some information about their data processing techniques; converting features to numeric expressions – a fairly common practise – and a set of categorical labels to a set of integers which worked for Random Forest and Hidden Markov models (scale-invariant models) but not for an SVM, so the integers were standardised to have a mean of zero and a variance of one. To test their models, they used a control model (no information classifier) which provided the baseline for their models to beat.

***Your mouse reveals your next activity: towards predicting user intention from mouse interaction [interaction]***

* The authors of this paper aim to predict user intention from mouse movements, clicks, and positioning (as well as other features). They focus on understanding what the user intended when they triggered an interaction and predicting the type of the next action.
* They build two classification models; a probability model and a machine learning (SVM) model.
  + The probability model uses previous *k* activities of a user to predict the next activity and a second (probability) model considers the time duration of the previous *k* activities.
  + The SVM model is trained on the *k* most recent user activities, the time duration of those activities, and descriptive statistics about the mouse interactions within a set window *W*.
* In their results, they find that both model types have similar performance in their experiments and to achieve the maximum accuracy – for predicting future interactions from historical actions – the combination of both models, in to a multimodal approach, achieves the best results (an average accuracy of 69%).
* In terms of variable values, they test a range of different values for both *k* and *W* and conclude that neither can be too large, the best value for *k* is 3 and when increasing the value of *W,* they find that it introduces a lot of noise.

In [cite] the aim is to predict user intention from mouse movements, clicks, and positioning – among other features – with the focus on understanding what the user intended when they triggered an interaction and then, from this data, attempting to predict the next action. Two classification models are built; a probability model and a machine learning (SVM) model. The probability model uses the previous *k* activities of a user to predict the next activity and a second (probability) model considers the time duration of the previous *k* activities. The SVM model is trained on the *k* most recent user activities, the time duration of those activities, and descriptive statistics about the mouse interactions within a set window *W*. In their results, they find that both model types have similar performance in their experiments and to achieve the maximum accuracy – for predicting future interactions from historical actions – the combination of both models, into a multimodal approach, achieves the best results (an average accuracy of 69%). With regards of variable results, they test a range of different values for both *k* and *W* and conclude that neither can be too large, the best value for *k* is 3 and when increasing the value of *W*, they find it introduces a large amount of noise.

***A grammar-based approach for modelling user interactions and generating suggestions during the data exploration process [interaction]***

* The authors build a system that can assist a user in an answer discovery task.
* Various options are presented to the user in the form of a graph and they are then able to traverse the graph in order to find an answer to their question.
* Through the process they are assisted by a tool that provides suggestions on additional variables (the number of bedrooms, for example) that they could add to their search.
* They provide a general point about determining a *k* value in the *k*-means algorithm when there is some uncertainty about the optimal value to use: test a range of values between 1 and sqrt(n/2), where *n* is the size of the data set. This upper bound prevents over-fitting of the value of *k*.

A system that can assist a user in an answer discovery task is presented in [cite]. Various options are presented to the user in the form of a graph, they are then able to traverse the graph to find an answer to their question. Throughout the process, an assistive tool provides suggestions on additional variables that could help their search (the number of bedrooms when searching for a house value, for example) and suggest a new path in their search. While not much can be taken from this paper, they make a general point about determining a *k* value in the *k*-means algorithm when there is some uncertainty about the optimal value to use; test a range of values between one and the square root of the size of the data divided in half. Adding this upper bound prevents over-fitting the value of *k*.

***An empirical study of user interaction behaviour during visual analysis [interaction]***

* The authors attempt to understand a user’s analytical goal and reasoning through high-level interaction metrics collected from a controlled study.
* They had two main aims from the study; determine common structures of visual analysis behaviour and characterise those structures in terms of what they represent and how they impact a user’s performance.
* In their results and evaluation, they were able to find two behaviour structures of visual analysis activity: patterns (short, ad-hoc sequences of visual actions performed iteratively through the task) and trails (chains of user visual interaction activity leading to insight).

In [cite] the authors attempt to understand a users’ analytical goal and reasoning through high-level interaction metrics collected from a controlled study. They had two main aims in the study; determine common structures of visual analysis behaviour and characterise those structures with regards to what they represent and how they impact a users’ performance. In their results and evaluation, they were able to find two behaviour structures of visual analysis activity: patterns (short, ad-hoc sequences of visual actions performed iteratively through the task) and trails (chains of user visual interaction activity leading to insight).

***Machine learning for sequential data: a review [Sequential data mining]***

* From a sequential data analysis context, the paper discusses difference machine learning approaches.
* There is a primary focus on text-based solutions and presents algorithms such as; sliding window methods, hidden markov models, and graph transformer networks amongst others.
* These methods are not exclusive to text-based solutions and have a place in other types of sequential data analysis.

A review of machine learning approaches, specific to a sequential data analysis context, is presented in [cite]. There is a primary focus on text- and natural language-based solutions and the paper presents algorithms such as; sliding window methods, Hidden Markov models, and graph transformer networks. The point is made that these approaches are not exclusive to text or natural language solutions and have a place in other types of sequential data analysis. An example of these approaches being used in other domains is shown in [using machine learning to infer…] where Hidden Markov models are used in an interaction data context.

***Incremental learning of daily routines as workflows in a smart home environment [Sequential data mining]***

* From annotated sensor data, the aim here is to describe how process mining can be used to learn users’ daily routines.
* Using a First-Order learning approach they propose a solution that can automatically learn daily routines from examples of people’s behaviour.
* They define three terms in this paper:
  + Activity recognition; aims at identifying the occurrence of an activity based on suitable combinations of events.
  + Task modelling; aims at identifying which combination of events determine the occurrence of an activity.
  + Process mining; starts from the activities performed by some agent to carry out a given process, and aims at identifying the valid patterns of activities that support that process.
* Further to this, they also make the point that Hidden Markov model-based approaches are more suitable for representing sequential activities from a stream of sensor events – this backs up the point made previously about the various applications the models can have.

Using annotated sensor data, collected in smart home environments, in [cite] the authors aimed to describe how process mining can be used to learn users’ daily routines. Through utilizing a First-Order learning approach (symbolized reasoning in which each sentence, or statement, is broken down into a subject and predicate), they propose a solution that can automatically learn from daily routines from examples of other people’s behaviour. They find, through experimentation with toy and publicly available datasets, that their approach is effectively able to learn and model the daily routines of people within smart home environments. In addition to their work, the authors provide definitions for three terms: activity recognition; aims at identifying the occurrence of an activity based on suitable combinations of events, task modelling; aims at identifying which combinations of events determine the occurrence of an activity, and process mining; starts from the activities performed by some agent to carry out a given process and identifies the valid patterns of activities that support that process. Finally, they also make the point that Hidden Markov model-based approaches are more suitable for representing sequential activities from a stream of sensor events – this backs up the point made previously about the various applications the models can have.

***Beyond the touchscreen: an exploration of extending interactions on commodity smartphones [interaction]***

* A new family of interactions to extend and enrich the input experience for users of smartphones are introduced in this paper. Various additional inputs for a smartphone are presented, these include the ability to detect in-pocket and on-table touches.
* The authors use both machine learning and rule-based approaches to analyse characteristics of the interactions. The rule-based approach required no specific training for the user of the smartphone, however, the machine learning method (KNN) did.
* The make the point about the difference between session- and user-dependent learning models.
  + Session dependent models require recollecting training data before using the model each time
  + User dependent models demand collecting training data from each user.

A new family of interactions to extend and enrich the input experience for users of smartphones are introduced in [cite]. Various additional inputs for a smartphone are presented, these include the ability to detect in-pocket and on-table touches through detection different physical interactions with the smartphone. The authors use both machine learning (classification) and rule-based (segmentation) approaches to analyse characteristics of the interactions and determine the action the user is attempting to perform. The rule-based approach required no specific training for the users of the smartphone due to looking for particular patterns and spikes in the microphone recording, accelerometer and gyroscope on the phone, however, the machine learning method (KNN) did. They adopted a user-independent model for the implementation of the KNN, which applies a pre-collected training data set and does not require any data collection from each user. In their evaluation, they found that the machine learning approach achieved results between 70% and 90% accuracy for the recognition of interactions and the rule-based interactions can achieve over 92% accuracy.