A6.1 Establish Baseline Visitor Intelligence – Library V4
Understanding Visitor Experience at Cultural Heritage Sites using Text Analytics

Work Package: T4
Period: July 2020
Authors: Dr. Pikakshi Manchanda, Valérie Duthoit-Saint-Georges
Principal Investigator: Prof. Andi Smart
1. Recap: Concept Hierarchy and its evolution

1.1. Concept Hierarchy

A comprehensive schema in the form of a concept hierarchy has been proposed to analyse Visitor Experience Interactions at different types of CH sites (in English and France) by capturing aspects which are of interest for the CH site managers and visitors. The concept hierarchy is comprised of:

- **Visitor Experience Interactions**: Root of the schema – Level 1. This node is meant to capture a holistic view of visitor experience based on visitor interactions with site resources at the heritage site.
  - **Visitor Experience**: An individual’s immediate or ongoing, subjective and personal response to an activity, setting or event outside of their usual environment (Packer & Ballantyne, 2016). It is affected by (reaction to) a number of touchpoints or interactions at a CH site (LaSalle & Britton, 2002).
  - **Visitor Interactions**: Studied as Direct, Indirect or Independent Interactions of a visitor with the resources at a CH site (Ponsignon, Smart, Williams, & Hall, 2015; Sampson, 2012; Verhoef et al., 2009)

- The root encompasses key categories/core concept groups – Level 2 – of visitor experience interactions (as informed by research), which are further elaborated with specific constructs at Level 3.
  - **Core Concept Groups**: These key categories/groups are aimed at capturing the core aspects of visitor experience feedback (evidenced from textual feedback collected from different sources) in terms of visitor interactions that are of interest for CH site managers.
  - **Concept Types**: Constructs/Sub-categories of concept groups.
  - **Evaluation Attributes**: Attributes that can be used for aspect/topic level sentiment analysis for visitors’ feedback at the concept type level.

The current version of the concept hierarchy is shown in Figure 1, section 1.2.

1.2. Concept Hierarchy - updates

The concept hierarchy – a collaborative effort of University of Exeter and NEOMA Business School -- is based on evolution of the Dashboard as seen in Deliverable v3 – Baseline Visitor Intelligence for the period September 2019. The hierarchy focuses on Visitor Interactions as a means to study the link between Service Design and Customer Experience. There have been a few updates to the hierarchy since its last version. The updated sections are highlighted in yellow in Figure 1.

Site managers can chose to evaluate sentiment at any level of detail in the concept hierarchy, i.e., a document level sentiment analysis at the concept group level or an in-depth aspect/topic level sentiment analysis at the concept type level based on the evaluation attributes.
In this work, the experience interactions are studied as Direct, Indirect and Independent Interactions that a visitor exchanges with (or responds to) animate or inanimate service offerings within a service encounter. (Consistent with work of Sampson 2012, Ponsignon et al. 2015)

- A **direct interaction** pertains to a person-to-person communication between a visitor and a service offering at a CH site, such as communication between a visitor and staff personnel at the site (Concept group 4 of the hierarchy).

- An **indirect interaction** pertains to a person-to-inanimate resource communication at a CH site. These resources broadly comprise of facilities and equipment and interaction with these resources may range from the use of amenities, availability of discounts/incentives, to ticket purchase experience and availability of digital technology at CH sites (Concept groups 1, 2, 3 and 5 of the concept hierarchy).

- An **independent factor** pertains to elements encompassing situational context which are not part of the site offering, aren’t always within the control of the site and effect/influence visitor
interactions across different phases of the visitor journey (Concept group 6 of the concept hierarchy).

1.2.1. Concept Groups: Updated Definitions

a) CH Site Offering: This concept group refers to the overall site space and the site premises, in association with site offerings in the form of site content and context. It comprises of the following clusters:

• **Site Space**: includes key concepts referring to the CH site’s area on land as well as site premises/surroundings.

• **Site Content**: includes physical artefacts, paintings, sculptures and associated signage (inanimate objects) at the heritage site.

• **Site Context**: includes abstract details at the site which usually lead to a ‘feel-good’ experience at the heritage site.

The three clusters under concept group 1 have been renamed as Site Space, Site Content and Site Context.

• **Changes to Site Space**: Removed ‘Site Upkeep’ (retained under concept group 6 (independent of site offering, yet might effect VX)).

• **Changes to Site Content**: switched ‘Cultural Signage’ from Supporting Services to Site Offering (Site Content) to capture concepts associated with artefacts etc. Removed ‘Site Presentation’ (Site Choir etc) since it has been seen to captures very specific concepts. Furthermore, choir related activities are captured by Interpretation Tools  Non Digital Interpretations  Site Activities (as initially proposed by NEOMA).

b) **Interpretation Offerings**  remains same as defined before (see here).

c) **Supporting Services**: This concept group refers to services/facilities provided and provisions made by the site to aid in understanding/appropriation of site. It comprises of the following concept types:

• Special Needs Adaptability
• Amenities, Facilities
• Information Provision/Promotion (This concept type has been renamed from ‘Website/Social Media/Advertising’)
• Generic Signage, Dedicated Interaction Points
• Food, Beverage

d) **Staff**: This concept group refers to mediation and non-mediation site personnel and corresponding evaluation attributes as observed by visitors in the form of Competency, and Attitude & Behaviour towards visitor needs. It comprises of the following concept types:

• Mediation Staff
• Non Mediation Staff
• **Ambiguous Staff** – This is a newly added concept type and refers to concepts associated with staff members present at a CH site that are difficult/ambiguous to classify in the afore-mentioned categories.
e) **Price**: remains same as defined before (see here).

f) **Situational Context**: This concept group refers to factors/elements outside the boundaries of the site and independent of the site offering, however, which have an effect on a visitor's experience at the site. It comprises of the following concept types:
   - External Climate/Atmosphere/Weather/Nature
   - Site Construction/Upkeep
   - Ease of reach / Geographical Proximity
   - External Facilities

**Motivation for including Situational Context**: (renamed from External Elements) in the concept hierarchy
If visitor experiences are defined as being designed or staged by a provider, the visitor experience is not a purely psychological phenomenon, but includes “the activities, physical surroundings, service providers, other customers, companions, and other elements they engage with” (Chang & Horng, 2010).

### 2. Evaluating VX using the proposed framework

Two approaches have been under evaluation of the proposed hierarchy for performing Aspect-based Sentiment Analysis on a given dataset of visitor feedback:

- **Use of IBM SPSS Modeler**
- **A Linguistics Rule based approach** based on state-of-the-art NLP libraries in Python (developed by UNAI in close co-ordination with UoE and NEOMA).

⇒ An introductory working research paper presenting a research framework for analyzing visitor feedback using Text Analytics has been presented at CAUTHE 2020 (Smart, Manchanda, Bonnin, & Georges, 2020). This paper includes a high-level description of the concept hierarchy, supplemented with sentiment analysis results based on the Modeler approach.

#### 2.1. Comparison of approaches: Preliminary Results (Exeter Cathedral)

The Modeler approach can be considered as a pattern-based approach in that syntactic text link rules and patterns were extracted from the dataset in order to perform sentiment analysis. On the other hand, the Linguistics Rule based approach makes use of state-of-the-art NLP Libraries (i.e., spaCy and Textblob) to perform sentiment analysis. The following sections present a comparative analysis of both these approaches in terms of a.) concept identification and classification as per the concept groups in the concept hierarchy (section 2.1.1) and b.) sentiment analysis for said concept types (section 2.1.2).

##### 2.1.1. Frequency distribution of concept groups

Given the changes to the concept hierarchy, the frequency distribution of the core concept groups has evolved since last reported in deliverable V3. The bars colored in blue reflect the current distribution of concept groups obtained using the (manually) created vocabs in SPSS Modeler for the dataset of Exeter Cathedral. The bars colored in orange reflect the frequency distribution of the concept groups obtained using word vectors utilized in the rule-based approach. As seen in Figure 2, word vectors often identify more occurrences of experience interactions than the manually created vocabs.
2.1.2. Sentiment Analysis

Figure 3 presents the aspect-based sentiment analysis results obtained using both the approaches for the core concept groups found in the Exeter Cathedral dataset. As evident, the performance of the rule-based approach is quite poor than the modeler, in particular for the negative sentiment. This could be due to the fact that either Modeler’s predictions comprise of many false positives which results in false precision, or the rule-based approach has failed to identify and classify many key concepts, thus, performing poor at the sentiment analysis phase. Such poor results for sentiment analysis results called for improvement in the current sentiment analysis algorithm. Additionally, it also established the need for evaluating the performance of both the approaches against a manually annotated baseline.
2.2. Updating the rule based approach

Following the identification of false positives in concept identification and classification phase (Figure 2) and poor sentiment analysis scores (Figure 3), the rule-based approach was updated by making use of Conditional Random Fields (CRF) for the purpose of Concept Extraction from visitor feedback in addition to a Fuzzy Matching approach for the purpose of Concept Identification from visitor feedback and vocabs. This work has been conducted with close coordination between UoE and UNAI. A diagrammatic representation of the updated approach is provided in Figure 4. The concept identification and classification process is conducted as follows:

- Both Fuzzy Matching and CRF Model are used to identify/extract concepts.
- When a concept is identified using Fuzzy Matching, it is automatically classified into a concept type based on Vocab Lookup.
- When a concept is identified using the CRF Model and a suitable concept type match isn’t found in the vocabs, Cosine Similarity scoring is used to classify the said concept. All such imputed concepts and corresponding concept types are saved to imputed vocabs. The imputed vocabs file, therefore, contains the concept, concept type, and similarity scores for how well the concept matched to other concept types. A manual editor can add the concept from the imputed vocab to the relevant concept vocab, making corrections if necessary. In this way, the next time this concept is encountered it will be identified and assigned the correct concept type.

Note that, as seen in deliverable v3, UoE and NEOMA had manually created concept type vocabs by analysing datasets of two heritage sites: a cathedral (Exeter Cathedral) and a castle (Fougères Castle). In order to analyse visitor feedback from different types of sites (castles, museums, gardens, etc.) by making use of the concept hierarchy, UoE, NEOMA and UNAI decided to partition the afore-mentioned concept type vocabs into three levels:

- **Level 1** – Generic vocab comprising of concepts which can be found be in all types of sites
- **Level 2** – Site type specific vocabs with concepts pertaining to specific types of sites, e.g., cathedrals
- **Level 3** – Site vocabs with concepts pertaining to specific sites, e.g., site-specific artefacts, events etc.

Figure 4 illustrates the use of these multi-level vocabularies to perform concept type classification. Additionally, UoE and NEOMA populated these vocabs (in particular, level 2 vocabs) by manually looking at concepts extracted from various other types of sites such as museums, gardens, industrial heritage sites, outdoor heritage sites etc. so that the CRF+Fuzzy Matching approach (hereafter referred to as CFM approach) can now be automated to analyse visitor feedback from different types of sites.

2.3. Performance Evaluation

2.3.1. Manual Annotation framework

In order to evaluate the performance of both the approaches (i.e., the Modeler approach and the CFM approach) in terms of their accuracy, precision and recall, there is a need for establishing the Ground Truth. UoE, NEOMA and UNAI worked together to prepare an annotation template which was used to prepare the ground truth. This ground truth is a manually annotated corpus of a sample of 100 English reviews from the Exeter Cathedral dataset (performed by UoE) and a similar sample of 87 French reviews from the Fougères Castle dataset (performed by NEOMA). The
Annotations were done at five different levels, which would, therefore, help in analysing the performance of both the approaches at those levels (see Figure 5).

The five level annotation and thus the five levels of evaluations that are currently underway based on the annotation template are:

a) **Concept identification** – can the model correctly identify concept phrases?
b) **Concept-type mapping** – can the model correctly map identified concept phrases to concept types and, thus, concept groups?
c) **Evaluation-attribute identification** – can the model correctly identify evaluation attribute phrases?
d) **Evaluation-attribute-group mapping** – can the model correctly map evaluation attributes to groups?
e) **Sentiment analysis** – can the model correctly map evaluation attributes to sentiments?

The sections below provide the evaluation results for levels (a), (b) and (c) obtained using the CFM approach and compared with the results from the manually annotated baseline. Note that analysis of the sampled dataset for these levels using the Modeler is under progress. Evaluation for levels (d) and (e) is also under progress.
2.3.2. Performance Evaluation: Concept Identification and Classification

A total of 696 concepts were identified in the manually annotated baseline for the English dataset of Exeter Cathedral consisting of a sample of 100 visitor reviews. This serves as the ground truth. The rule-based approach, as discussed in Section 2.1, identified 1132 concepts – thus showing a huge portion of false positives (as was also evident for the results of the entire dataset shown in figure 2). On the other hand, a total of 777 concepts were identified using the CFM approach, as discussed in Section 2.2. This shows that the CFM model was able to lower the scale of concept mis-identification and mis-classification.

Figure 6 shows the concept classification results of concepts into concept groups for both these approaches as compared against the concepts found in the manually annotated baseline. A similar process was conducted for the French dataset of Fougères Castle as well.

![Figure 6: Performance Evaluation - Concept Identification and Classification](image)

Additionally, Table 1 shows the precision, recall and accuracy scores for both the approaches evaluated against the respective baselines in terms of concept identification and classification for the English and the French datasets.

<table>
<thead>
<tr>
<th>Concept Identification</th>
<th>English Dataset (ExCath)</th>
<th>French Dataset (FGC)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Rule-based Approach</td>
<td>CFM Approach</td>
</tr>
<tr>
<td>Precision</td>
<td>0.38</td>
<td>0.71</td>
</tr>
<tr>
<td>Recall</td>
<td>0.61</td>
<td>0.86</td>
</tr>
<tr>
<td>Concept Classification</td>
<td>0.45</td>
<td>0.93</td>
</tr>
</tbody>
</table>

*Table 1: Concept Identification and Classification: Performance Metrics*

Here precision is defined as a score of number of correctly identified concepts by a given model divided by the total number of concepts identified by the model. Recall is a measure of number of correctly identified concepts by a given model divided by the total number of true concepts that exist in the ground truth. Finally, in this case, accuracy represents the score of correctly
classified concepts into concept groups by a given model divided by the total number of correctly identified concepts.

It can be clearly seen in Table 1 that the CFM approach improved the concept identification and classification performance by a huge margin, in that precision for concept identification and accuracy for concept classification almost doubled for both the datasets. The CFM approach, therefore, will be used for experimentation with future datasets as well.

### 2.3.3. Performance Evaluation: Evaluation Attribute Identification

This section shows the performance scores for evaluation attribute phrase identification using both the approaches against the respective baselines for the English and the French datasets. A total of 378 phrases were found in the ground truth of the English dataset. As evident from the scores listed in Table 2 below, an advanced rule-based approach has been used to improve the evaluation attribute identification by almost double. Evaluation Attribute type identification and sentiment analysis are currently under progress.

<table>
<thead>
<tr>
<th></th>
<th>English Dataset (ExCath)</th>
<th>French Dataset (FGC)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Simple Rules</td>
<td>Advanced Rules</td>
</tr>
<tr>
<td>Precision</td>
<td>0.32</td>
<td>0.59</td>
</tr>
<tr>
<td>Recall</td>
<td>0.26</td>
<td>0.42</td>
</tr>
</tbody>
</table>

*Table 2: Evaluation Attribute Identification: Performance Metrics*

### 3. Review Capture Form

UoE and UNAI have also collaborated to design a platform for capturing visitors’ feedback in addition to capturing them from TripAdvisor. This platform – called as the Review Capture Form – is tailored to capture descriptive reviews and visitor rating with regards to different aspects of their experience at heritage sites as well as a cumulative experience score, while adhering to the standards of visitor anonymity. Such information, when fed into the Visitor Intelligence Dashboard, would be instrumental in understanding visitor profiles, behaviour and motivations for visit.

The platform is currently in the trial phase where in partners of the project have been asked to leave a review (real or imaginary) for one or all of the five key sites. Any feedback about the usability of the form is also welcome. Please click on the site name to follow the link: Exeter Cathedral, Fougères Castle, East Pool Mine, Botallack Mine and Slapton Sands. Any feedback can be sent to vista-ar@exeter.ac.uk. A snapshot of the platform’s appearance is shown in Figure 7.
Figure 7: A VISTA AR Platform for capturing visitors’ feedback
4. References


