

## User Information Needs

The interest in Lovecraft's fiction is not purely academic. It is an aesthetic resource and convincing alternate reality that has spawned a host of similar fiction, pop art and role playing games. Just as Lovecraft used the universe he created as a setting for his creative works, others continue to use his world as a setting for their own imaginative expression. As a result, user information needs are not limited to traditional accuracy, access and findability concerns, but also the need to feel that the index is part of this group which views Lovecraft's work in the same inspirational light.

Flatly stating how wonderful Lovecraft is would not be convincing. Instead, the website includes design details that reveal the fact that the index/web designer is a fan as well. First, the handwriting font was selected because of Lovecraft's signature use of deteriorating handwriting as a symptom of mental deterioration. This element is incredibly important to Lovecraft fans. So much so that legibility can be sacrificed to some extent. There are even blog posts which describe the creation of fonts that resemble insane handwriting<sup>1</sup> and collections of insane/archaic handwriting fonts you can buy/download are available in several places online.<sup>2</sup>

In addition to using fonts which recall the handwriting of a man driven mad by the truth of the universe, the presentation of the web page includes uses the format of a set of moldy old file folders. Lovecraft's stories frequently involve someone examining the notes and research of a professor or ancestor who dies mysteriously. The mold and dirt make a connection to the visceral textures of decay that appear throughout the stories as well.

Other elements provide subtle access to supporting information. For example, the photo and text "H.P. Lovecraft?" will connect to biographical information when clicked on. This information is not essential to the index, so it can be presented in a more subtle fashion. Including semi-secret information in the website is another way of

Finally, the Cthulhu Approved graphic, which as created by Wes Jones, recognizes the pop culture status of the Cthulhu, the humorous and, ironically, anthropomorphized versions of the creature, which is the most widely recognized of his creations.

These drawings are just a mockup. If the the project moves on to becoming a real web resource, photos of the actual materials would be used, with different areas of the image mapped as

---

1 Font of All Madness. Dean Adelaide. June 2, 2013. <http://cthulhureborn.wordpress.com/2013/06/02/font-of-all-madness/>

2 <http://www.cthulhulives.org/toybox/propdocs/propfonts.html>

links. This would make the illusion much more meaningful. So with the user's need for the appropriate atmosphere taken care of, they will be more receptive to the more concrete content which the index has to offer.

The concrete half of the user need can said to exist anywhere on a certain spectrum. On one end there are people who are already very familiar with Lovecraft's work and are looking for specific details regarding certain aspects of his oeuvre. Anticipating these questions and providing information specific and rare enough for these users is one concern when designing the index. At the other end of the spectrum are people who may have read a story or two by Lovecraft and want to learn more about his work, or perhaps find another story with a character they liked.

The index should be simple enough that users unfamiliar with his work can find what they are looking for, but transparent enough that they can discover new ways of thinking about Lovecraft's stories. For example, after reading "The Dreamquest of Unknown Kadath" a reader may like to see what other stories the character Randolph Carter appears in. In locating the earlier story, "The Statement of Randolph Carter," the user might also discover that Lovecraft's works are divided up into three phases or wonder why the textures were used in indexing, and be able to find the answers to these questions.

So three main concerns are part of this aspect of user information and will be referred to again and again:

1. Specific and thorough indexing for the advanced user
2. Simple searching for the novice user
3. Transparent and education index structure to facilitate serendipitous learning

### **Subject Scope**

The subject scope of this index are the thematic elements and content of the short stories of H.P. Lovecraft. This index will be part concordance, in that a user can locate all the appearances of a character such as Richard Pickman. However, it will also function as a literary index in that the various dramatic and literary devices and philosophies that Lovecraft makes use of will also be presented for exploration and comparison.

### **Collection Scope and Domain**

Lovecraft's fiction presents one significant difficulty when creating an index. Namely, which of the endless collections and editions of the stories should be used as a foundation for this index.

The scope of the collection is four volumes of short stories by H.P. Lovecraft

published by Arkham House.

Lovecraft, H. P., Derleth, A., & Joshi, S. T. (1985). *At the mountains of madness and other novels (Corrected 6<sup>th</sup> Printing)*. Sauk City, Wis: Arkham House Publishers.

Lovecraft, H. P., Derleth, A., & Joshi, S. T. (1987). *Dagon and other macabre tales (Corrected 5<sup>th</sup> Printing)*. Sauk City, Wis: Arkham House Publishers.

Lovecraft, H. P., Derleth, A. W., & Joshi, S. T. (1989). *The horror in the museum: And other revisions (Corrected 4<sup>th</sup> Printing)*. Sauk City: Arkham House Publishers.

Lovecraft, H. P., Derleth, A., & Joshi, S. T. (1984). *The Dunwich horror and others (Corrected 4<sup>th</sup> Printing)*. Sauk City, Wis: Arkham House.

These collections are corrected printings from the 1980s with comprehensive editing and a new introduction by S.T. Joshi. These volumes are used here because they are recognized as definitive based on the editing and research of Joshi.

All the documentary units in the collection will be stories, pages, and/or paragraph numbers as well for shorter passages. The stories were all published in serialized form previously, but again, for reasons of academic motivation and simplicity this index will reference the above collection only. The stories are all in the horror/fantasy genre, with some sub-genres, written and published in English.

The decision to limit the applicability of the index to these four collections of stories makes the construction of the index feasible, as an attempt to reference even a limited number of different publications would clutter the index as to make it unusable, not to mention the costs to the sanity of the user. It is a wonder Lovecraft did not write a story about a man losing his mind creating an index.

The drawbacks to this approach is the limited use the index will have to those not in possession of the corresponding collections. In addition, the collection indexed here is not an exhaustive collection of Lovecraft's writing, and those students of his work will undoubtedly be interested in important stories that are not in this collection. This situation could be mitigated by including the full text as part of the index, especially because Lovecraft's work is either clearly in the public domain, or treated as such. The problems with this are the significant effort required to assemble this body of text, even using OCR. And to be quite honest, I feel like people should use books to read, not the web, especially when it comes to fiction.

His poems will not be included here because of the difficulty of locating and organizing his poetry, and the relative lack of a significant user base interested in those poems.

Future initiatives may include the addition of his large collection of personal correspondence to the subject scope of the index. Many interesting connections could be made between the letters and the

fiction if queries were able to access both.

### **Classified/Relational Arrangement or Facets**

The index will use top level categories with pre-coordinate index terms to facilitate transparency and learning about the existing body of Lovecraft criticism. Facets could provide an interesting approach to finding a story that contains examples of thematic concepts, supernatural forces, certain recurring characters, etc. However, these stories are not the documentary units, actual pages and paragraphs are the documentary units, and a faceted search will probably have an extremely high precision, frequently zero returns, when using complex faceted queries to look for individual pages. It may be interesting to include a single faceted search page in the index. For instance, I'm looking for a murder mystery, set in Providence, that contains themes of the agency of man in an incomprehensible cosmos, would return *The Shunned House* from these selections in the Genre, Setting and Theme Facets.

Each of the top categories will come with some description of what it includes and why it is a subheading. For example, textures might seem to be an interesting choice for a top level category, but the unorthodox use of textures in the horror genre is an important critical element when studying Lovecraft.

Many of the documentary units will include combination entries, essentially ignoring mutual exclusivity, because a pseudomythological entity, like Nyarlathotep, functions as an element of the setting, a representation of cosmicism, the air element, and will use certain textures, depending upon the story. (Please note: elemental significances are a post-mortem overlay of ideas by one of Lovecraft's contemporaries, and are not included in this index.)

It seems like a surrogate display that includes a 'See Also' section based on a combination of the query and the particular documentary unit would be a great idea, and an incredibly complicated algorithm to develop...

### **Vocabulary Management**

Establishing even top level categories and subsidiary terms for this index is a difficult task. First, Lovecraft's common themes not only upon the work of Lovecraft, but on the work of dozens of literary critics, whose vocabularies, sometimes created specifically to describe aspects of Lovecraft, are not universal. Many terms are extremely complex in meaning, which leads to significant overlap in terminology, but not simple synonymy.

For example, the index term 'questionable parentage' could be synonymous with the theme of 'inherited guilt', but there are other examples of 'questionable parentage' that are more physical than psychological. In addition, it 'questionable parentage,' a theme? A character trait? Actually both, so it will appear under both headings, and sometimes with the subsidiary term 'inherited guilt,' which is another layer of relational indexing.

Establishing a controlled vocabulary with synonyms is especially difficult, due to numerous instances like the ones described above. In the example of the top level category theme, 'questionable parentage' could stand alone with no subsidiary terms, but under human traits, it can have subsidiary elements which clarify each particular case. Theme is broad and character traits are specific.

Despite all these obstacles, a vocabulary authority will be simple to establish, after the index is complete. Much like the American Imprints Series 1, the terms will need to be created and expanded based on the content of the collection. As more and more stories are actually indexed, the list of textures, human character traits, etc. will be expanded. As this is a closed collection there will be an end point, which makes the simultaneous development of the vocabulary authority and index itself possible, as long as the designer is aware of the ongoing development and the intention to create a final authority.

In addition some prominent elements of Lovecraft's work are certainly worth discussion, but are too broad and ubiquitous to allow for indexing. Antiquarian language is used throughout his work, and is one of his most recognizable literary calling cards, however it cannot be practically indexed due to the endless and constant use of this device. Instead, a small aside discussing the general character of his writing will include these elements.

### **Analysis and Indexing methods**

The analysis of the documentary units and the assignment of index terms will be completed solely by myself. This fact makes the simultaneous development of a controlled vocabulary and the index itself possible without a host of tracking and management updates that would keep all the indexers on the same page. The dangers of this approach are the same dangers that haunt all solitary creative efforts, a myopic view of the creation. To address this issue a small group of 12 users, selected to represent a cross reference of the actual user population, will be asked to test the index, and report upon their experience.

Algorithmic processes could be used to develop the aforementioned See Also section of displayed records, but this is a project outside my current capabilities. Otherwise all development of the index will be human.

Full text searching is another option that would alleviate other problems that have been previously discussed. However, this method would also be the least instructive to the user, and should be accompanied with admonitions regarding its strengths and weaknesses in terms of broadening the users horizons.

### **Depth of Indexing**

The depth of indexing will constitute a scanning of the full text and depend entirely upon the top level category being discussed. Texture is a very simple concept, but in Lovecraft's work a certain family of visceral textures are relevant to the study of his work. The indexing in this area will be exhaustive in nature while regular textures are not noted in the index. Every oozing crack in the basements of an imaginary New England shall be painstakingly captured and classified, but traditional textures shall be ignored. Human character traits are a similar matter. There are traits which are integral to Lovecraft's work, and traits which are not especially relevant.

Every instance of a relevant attribute or theme will be cataloged, but this will not mean every page or every paragraph will have material for each subject heading. There will be thematic index terms for each story included. The definition of texture, character trait, theme, etc. will be an essential element of the index terms as presented for browsing and the structural information on the index's about page.

There will be no rule of three or other limiting factors applied to any of the documentary units. Regardless of the complexity or density of index terms relating to a specific story, the documentary unit options of story, page and paragraph number will also provide some flexibility in terms of how many terms are applied to each documentary unit.

In terms of mutual exclusivity, the principle is not really being violated by the application of the same index terms to different top level categories, because the terms are not referring to the same thing, as is made clear by the pre-coordination. 'Cosmic Indifferentism, human awareness of' is different than 'cosmic indifferentism, as theme' or 'cosmic indifferentism, as setting.'

### **Syntax for representation of displayed indexes**

Using pre-coordinate terms under specific top-level categories seems to be the only way to represent the significance of index terms and provide the transparency that the user information need necessitates. As discussed certain terms are multi-faceted in nature and can apply to certain characters or to themes. For example, a user should be able to locate a list of characters who have questionable parentage, as well as search for the theme of questionable parentage separate from any individual

character, and they should be able to know which one of the top level categories, or possibly both, will be searched when they click on a term in the displayed index.

Terms like semi-schizoid, which may not be politically correct or psychologically accurate, or just rub you the wrong way, have been included because they exist throughout other critical discussions of Lovecraft's works. This reason for including these terms should be clear. Also, I have included the names and current addresses of the individuals responsible for the creation and popularization of these insensitive terms, so users can leave an Amulet of the Hound under their pillow.

### **Alphanumeric arrangement**

The closed nature of this collection and the single language and consistent format of the documents contained therein permits a simple alphabetical arrangement following the English language alphabet.

However, there are a number of invented words from the stories that include hyphens and apostrophes to indicate the rhythm of pronunciation. It should be noted in the supporting documents that these marks will not be recognized by the alphabetizing system. They will be ignored and the next letter consulted for the purposes of alphabetization. These marks will not count as spaces either.

### **Locators and links to documents**

There are no links to the full text documents as this index will only provide locators. Locators will be presented in the different forms depending upon the level of the documentary unit. A query for ooze, would have returns with locators that include a piece of the text, and locators down to the paragraph. The passage option displayed here would be ideal for the

**Passage:** “anything more definite than rock and **ooze** and weed was seen”  
**Story:** *The Call of Cthulhu*  
**Page(s):** 151  
**Paragraph:** 2  
**Collection:** *The Dunwich horror and others.*

A query for Randolph Carter, would contain returns with locators such as:

**Story:** *The Dream-Quest of Unknown Kadath*  
**Page(s):** 140-188  
**Collection:** *At the Mountains of Madness and other novels*

Because there are varying sizes of documentary unit and a note specifying the type of documentary unit returned could be helpful, but it should be obvious based on the specificity of the metadata.

## **Surrogate displays**

Documentary units will vary. When searching for textures it would be beneficial to receive a page and paragraph number when locating the referenced text, but when looking for a character of questionable parentage who is a historian, the volume title and story may be sufficient, as a character will appear throughout that story, and in other stories the same character will be mentioned briefly, for example Richard Upton Pickman, who is the main character in “Pickman's Model” and makes a few cameos in “The Dreamquest of Unknown Kadath.”

In order to facilitate quicker access for users who are looking for information regarding a certain story, there is a Browse by Story section of the index. This provides an alphabetical list of the story titles, and each surrogate record provides bibliographic data and a list of index terms that are connected to that story. Users can also select a specific title to search in the Advanced Search Screen.

As was discussed in the user information needs, surrogate displays should be instructive regarding the work of H.P. Lovecraft by providing transparency in their working and by exposing the elements of his fiction that the user may not be aware of.

This is easily done in the Browse by Story section of the index, but particular records, such as the one for 'ooze' in The Call of Cthulhu will require an expository section that illuminates the reasons for cataloging this particular term or type of terms. The term 'insanity' for example could provide advice for searching it as a theme, as a character trait, and why this term was an important theme for Lovecraft.

A 'See Also' section might provide links to other information. Textures which appear in the same story, other stories that involve sailors, etc. The algorithm used to construct this mechanism would need to be aware of the way the query was entered, and the relationships between pre-coordinated terms.

Finally, the surrogate display omits a significant amount of information which is attached to each individual record. Author, language, type, and other basic elements are included in the record in the case that the user wishes to separate it from the index, but while using the index, there is only one author, one language, etc.

## **Record structure**

Here is an example of a typical XML record structure for one paragraph. Pre-coordinated terms have been used to characterize topics with multiple applications in the study of Lovecraft's fiction. The traditional XML subject field is not used. While this would provide flexibility in index design by changing the way the software manipulates the index, not in changing the structure of every record, in



order to facilitate understanding by the user when the record structure is viewed, fields

DOCUMENT-can refer to any story, page or paragraph, which is indicated by the presence of fields

COLLECTION-refers to one of the four volumes.

```
<XML>
<INDEX>
  <DOCUMENT>
    <TYPE>Text</TYPE>
    <LANGUAGE>En</LANGUAGE>
    <PAGE>151</PAGE>
    <PARAGRAPH>2</PARAGRAPH>
    <COLLECTION>The Dunwich Horror and Others</COLLECTION>
    <TITLE>The Call of Cthulhu</TITLE>
    <AUTHOR>H.P. LOVECRAFT</AUTHOR>
    <DATE WRITTEN>1926</DATE WRITTEN>
    <DATE PUBLISHED>1928-02</DATE PUBLISHED>
    <ORIGINALLY PUBLISHED>Weird Tales</ORIGINALLY PUBLISHED>
    <TEXTURE>ooze</TEXTURE>
    <HUMAN TRAIT>cosmic indifferentism, human awareness of</HUMAN TRAIT>
    <PHASE>2</PHASE>
    <PHASE>3</PHASE>
    <RIGHTS>Public Domain</RIGHTS>
  <DOCUMENT>
```

## Metadata schema

Because this is a closed collection, and links to full text would be done by adding the text to the index itself, the metadata schema does not require any interoperability. Field names have been customized in order to provide some organization to the many index terms associated with some documents, and alleviate the increased complexity of pre-coordinated terms that also use the field name. The first example seems to become too convoluted, with the word human used twice, etc. While the second example appears to fit the purposes of this index and its overlapping terms more completely.

```
<SUBJECT>human trait, cosmic indifferentism, human awareness of</SUBJECT>
```

```
<HUMAN TRAIT>cosmic indifferentism, human awareness of</HUMAN TRAIT>
```

The classified index display will be easier to maintain and use if pre-coordinate terms are smaller, while a complete alphabetical list of all terms will be less intelligible. However, viewing the list of index terms without classification, as a whole, will be a function presented for advanced users, and not easily stumbled upon.

Some element of Dublin Core have been added to provide a platform for inclusion in existing library indexes if such a situation arises. However, the fields of type, language, etc. is not the core of this index's functionality, and interfaces specifically designed to exploit the meaningful fields of

texture, theme, human traits, etc. will have to be transferred with the index.

This list of displayed index terms is incomplete. As discussed before, as the analysis of the documents is underway, index terms will be combined/created/eliminated in keeping with the user information needs and literary criticism surrounding Lovecraft's work. This is a predictive/partial list based on the indexers existing knowledge and completed work.

#### Cosmic/Supernatural Entities

- Azathoth
- Cthulhu
- Dagon
- Hastur the Unspeakable
- Nyarlathotep
- Shub-Niggurath
- Yog-Sothoth
- Elder Ones
- Pth'thya-l'yi

#### Cosmic Geography

##### Imaginary New England

- Maine
  - Potowonket
- Massachusetts
  - Arkham
  - Dunwich
  - Innsmouth
  - Miskatonic University
  - Newburyport
- Rhode Island
  - Providence

#### Fictional Texts

- Necronomicon
- Green Meadow Book
- Whispers

#### Human Traits

- Cosmic Indifferentism, human awareness of
- Helplessness, human awareness of
- Hopelessness, human awareness of
- Insanity
- Detachment
- Questionable Parentage
  - Inherited Guilt
- Professions
  - Anthropologist
  - Artist
  - Biologist
  - Engineer
  - Fisherman

Geologist  
Historian  
Investigator  
Linguist  
Musician  
Physicist  
Policeman  
Professor  
Sailor  
Scientist  
Student  
Writer

Semi-schizoid

Phase

- 1, Macabre stories, (approximately 1905–1920)
- 2, Dream-Cycle stories, (approximately 1920–1927)
- 3, Cthulhu Mythos stories, (approximately 1925–1935)

Real People

Poe, Edgar Allan

Sub-Genre/Co-Genre

Mystery  
Crime  
Detective  
Fantasy  
Science Fiction

Themes

Anti-anthropocentrism  
Classicism  
Cosmic Indifferentism  
Cosmicism  
Detachment  
Decline of Civilization  
Distorted Space  
Ethnicity  
Esoteric Knowledge  
Fate  
Forbidden Knowledge  
Helplessness  
Hopelessness  
Inherited Guilt  
Insanity  
Loneliness  
Promethean  
Questionable Parentage  
Race  
Religion  
Science, Dangers of  
Technology, Dangers of

Visceral Textures

Dust  
Gelatinous  
Goo  
Miasma  
Ooze  
Slime  
Sticky