**Forming a Corpus – An Invitation to the Digital Humanities**

**Abstract:** The Digital Humanities is a field that continues to grow and change rapidly as the digital age progresses. It represents the overlap between the humanities and digital technology, as well as the implications of this overlap. It renders itself distinct from the Traditional Humanities not only in its use of technology, but also in that it is a field that exists with the built-in notion that virtual media are now the primary means by which people communicate, write, and produce art. Consider for a moment how computerization and technological advancement have allowed for such an increase in the production of artistic work that the means by which the average individual creates such work have changed entirely from advent of digital technology, and then compare that to the production of artistic work prior.

In this course students will create a small text corpus comprised of a collection of works from a variety of media, contexts, and background in order to establish a theme for the purpose of developing a fuller understanding of text analysis, interdisciplinary study, and intertextuality. Over the semester, this corpus will grow and take form: as this occurs, students will use their corpuses to explore ideas and answer questions about that theme. This corpus is a “living” body of work; students are encouraged to add and remove texts from it as the semester progresses, and even after. Additionally, students will put works from their corpuses and their corpuses as a whole in conversation with assigned works, and use digital tools and platforms to perform analysis on assigned texts as well as texts of their choice.

A secondary component of this course is the Class Omnibus – a collaborative, less-formal corpus that all students, as well as the instructor will contribute to over the semester. During any 3 weeks of the semester, each student is expected to contribute at least one digital or physical work that relates to the current class topic to the Class Omnibus. This contribution will be done via a short, informal written justification (roughly 100 words), in which students will explain their reasoning for the contribution, how the work relates to the topic, and if they wish, if and how the work connects back to their personal corpus. Each week through week 12, one student will also lead a brief presentation on that week’s topic, followed by a class discussion. Students will retain access to the completed Class Omnibus after the end of the semester.

Note: This syllabus contains information that will remain relevant throughout the semester. It will be referenced regularly, and a digital copy is available via the course website.

**Content Disclaimer:** This course will not frequently address disturbing content. On the occasions that it does, there will be a clear warning about it. However, students may encounter disturbing content while exploring tools, websites, and self-selected works. Use caution and discretion; students should know and respect their own personal limits. Students are encouraged to contact the instructor in the case that any course content is unacceptably upsetting to them.

**Course Requirements and Grading:** Below is an explanation of how one can be expected to be graded in this course, as well as some general expectations for conduct:

**Attendance and Participation (10%)**All students are expected to come to class prepared, ready and willing to participate in discussion, and on time. Additionally, students are expected to remain undistracted during class and to show respect towards one another as well as to the instructor. If a student is going to be absent, that student is expected to present a doctor’s note or other appropriate documentation excusing that absence. More than two unexcused absences during the duration of the semester will result in significant decrease in overall course grade.

**In-Class Presentation (5%)**At some point of the semester, students are asked to present on a topic. This presentation was covered in the abstract, but more information will be given in class.

**Miscellaneous Assignments and Weekly Upkeep (20%)**Each week, students are expected to Consider, Explore, Read, and complete Assignments (CERA, for those in favor of mnemonics). In the “course schedule” portion of this syllabus, each week’s material is organized in this “CERA” format. This portion of the grade also covers contributions to the Class Omnibus. Students should also be aware that it is generally the case that each week’s readings and explorative content will be very helpful, if not outright integral to understanding material present in the following week’s readings and explorative content.

While the “CRA” of “CERA” is fairly self-explanatory, the “E” may be somewhat unfamiliar. To properly “Explore” assigned content, students are expected to browse and/or experiment with the content. Students are absolutely not expected to achieve expertise or mastery with explorative content, but they should familiarize themselves with it to the point of being prepared to discuss it critically in class.

The last three weeks of the term only have “CE,” as that will be the time that students will be finalizing their corpuses and final projects.

**Projects (25%)**Several larger-scale projects, usually spanning multiple weeks, are part of this course’s curriculum. Students will be provided with detailed handouts covering the specific expectations for each project when it is assigned.

**Final Project and Completed Corpus (40%)**At the end of this course, students will complete a three-piece final assignment. Students will be provided with detailed handouts covering the specific expectations for each project when it is assigned.

**Course Material:** In the interest of accessibility, a significant amount of course material will be made available through a course packet or handout. Works on this list can be added to your corpus, if appropriate.

* Ariella Rotramel, Rebecca Parmer, and Rose Oliveira’s *Engaging Women’s History through Collaborative Archival Wikipedia Projects* (link available on course website)
* **Henry James’ *Professions of Taste* (.pdf available on course website)**
* Luigi Serafini’s *Codex Seraphinianus* (.pdf available on course website)
* Maggie Nelson’s *Bluets* (.pdf available on course website)
* Meredith Broussard’s Artificial Unintelligence: How Computers Misunderstand the World
* R. Watson Todd’s Building and Using Your Own Corpus and Concordance (included in course packet)
* Roland Barthes’ “The Death of the Author” (included in course packet)
* Walter Benjamin’s “The Work of Art in the Age of Mechanical Reproduction” (handout)
* Stephen Ramsay’s *Reading Machines: Towards an Algorithmic Criticism* (.pdf available on course website)
* Excerpts from Susan Sontag’s *Against Interpretation* (.pdf available on course website)
* Excerpts of James Joyce’s *Ulysses* (.pdf available on course website)
* Excerpts of George Kingsley Zipf’s *Human Behavior and the Principle of Least Effort* (included in course packet)
* Excerpts of Gilles Deleuze and Félix Guattari’s *A Thousand Plateaus* and *Anti-Oedipus* (included in course packet)

**Recommended Texts:** These are suggested, mainly for the purposes of gaining background knowledge, but not required reading. Feel free to use them in your corpus, too. They will not necessarily be included in the course packet, however.

* Dante Alighieri’s *Divine Comedy* (week 6)
* Miguel de Cervantes’s *Don Quixote* (week 6)
* Wayne Koestenbaum’s *Humiliation* (week 12)
* William Ian Miller’s *Anatomy of Disgust* (week 9)

**Course Schedule:** This course is divided into three loose units, each comprised of four weeks. The first four focus on digital text analysis, the second four focus on critical appraisal and examination of material that exists within and around the sphere of the Digital Humanities, and the third four focus on combining aspects of traditional literary analysis and critical theory with the Digital Humanities. The final three weeks of this course are dedicated to the final project.

* **Week 1: Introduction**
  + Consider: What is the purpose of a corpus? What can one attain by analyzing a corpus as a whole? Is a corpus a “body without organs?”
  + Explore: [JSTOR Labs Text Analyzer](https://www.jstor.org/analyze/), [Voyant](https://voyant-tools.org/)
  + Read: *Human Behavior and the Principle of Least Effort* (pages 313-347)
  + Assignments: Use Voyant to analyze any 5 digital pieces you have and write a brief, informal (roughly one page) reflection on your experience.
* **Week 2: The Formation of a Corpus**
  + Consider: What is a corpus? What forms a corpus? What are the base components and scaffolding of a corpus? Additionally, is the collection of texts assigned by this class a corpus itself?
  + Explore: [Lexos](http://lexos.wheatoncollege.edu/upload), [On-Line Encyclopedia of Integer Sequences](https://oeis.org) (optional), [SkELL](https://www.sketchengine.eu/skell/), [Sketch Engine](https://www.sketchengine.eu) (has free trial, can be bothersome to download, optional, but recommended)
  + Read: Building and Using Your Own Corpus and Concordance, *Capitalism and Schizophrenia* excerpt set 1 (selected passages concerning the “body without organs.”)
  + Assignments: Formally begin intro project (see “Projects”).
* **Week 3: “Web Corpus”** 
  + Consider: Is the creation of large-scale corpuses necessary, and if so, why? How do large-scale corpuses differ from small-scale corpuses?
  + Explore: [GitHub](https://github.com/), [Google NGram](https://books.google.com/ngrams), [Pastebin](https://pastebin.com/)
  + Read: Artificial Unintelligence: How Computers Misunderstand the World (Introduction), *Reading Machines: Towards an Algorithmic Criticism*
  + Assignments: Select a word and set the lower bound of time range on Google NGram to 1500. Then, run that word through Google NGram. No formal writing is required this week, but be prepared to discuss your findings in class: note taking is encouraged.
* **Week 4: Digital Integration**
  + Consider: What effects does the digitalization of a work have on that work and related pieces? Does this affect analysis of that text? How do a physical text and a file containing that text differ?
  + Explore: [Grammarly](https://www.grammarly.com/), [Hypothes.is](https://web.hypothes.is/) (optional), [UsingEnglish Text Statistics Analyzer](https://www.usingenglish.com/resources/text-statistics/)
  + Read: “The Work of Art in the Age of Mechanical Reproduction”,
  + Assignments: Intro project due this week. Use a tool we have explored (I.E., Voyant, Lexos) to perform a cursory comparative text analysis of any work of your choice and the digital edition of Ernest Vincent Wright’s *Gadsby* uploaded to the course website. Note that *Gadsby* was written with the intent of possessing virtually no instances of the letter “e." Write a one-page response on your findings: including screenshots is highly encouraged.
* **Week 5: Limitations and Technochauvinism**
  + Consider: Is there any meaningful truth to comparisons between the human brain and a computer? How can the beliefs and biases of a creator be reflected in a creation? What role does restriction play in an analysis? How would one describe “limitations” of text analysis? How does, for instance, MonkeyLearn compare to other tools?
  + Explore: [MonkeyLearn](https://monkeylearn.com/word-cloud/) (using .txt files strongly recommended), [Wikipedia](https://www.wikipedia.org/)
  + Read: Artificial Unintelligence: How Computers Misunderstand the World (sections 1 and 2), *Engaging Women’s History through Collaborative Archival Wikipedia Projects*
  + Assignments: Find an idea that speaks to you in the readings and write a two-page response reflecting on it. Using texts we’ve examined in class is encouraged, as is using texts in your corpus. Also, just for fun – choose a Wikipedia article, then click “random article.” From the random article, try to get back to the original article in as few clicks as possible.
* **Week 6: Hypertexts**
  + Consider: Can a physical work be a true hypertext? What are the identifiable aspects of a hypertext, and how do they define a hypertext? Can a hypertext be in and of itself a corpus? Furthermore, could last week’s “just for fun” assignment be a means of turning Wikipedia into a hypertext in a sense?
  + Explore: [SCP Foundation](http://www.scpwiki.com) (a collection of recommended articles is available via the course website). Come to class with at least two articles you’re prepared to talk about.
  + Read: *Ulysses* excerpts
  + Assignments: This week, reflect on your corpus. Consider if any of the works in there are in any way hypertextual. Schedule a meeting with the instructor for before week 9 for an in-depth conversation with about your corpus.
* **Week 7: Mixed Media and Interdisciplinary Content**
  + Consider: How does examining a work that combines mediums differ from examining a work that only uses one? Can you separate a work’s mediums when you analyze it? Should you?
  + Explore: [Twinery](https://twinery.org/), at least two of the Twine games listed on the course website (if you choose *Birdland*, you don’t need to play a second game if you don’t want to). Feel free to play other Twine games not listed here.
  + Read: Twine documentation (optional, available via course website)
  + Assignments: Begin on mid-semester project (see “Projects”).
* **Week 8: Derivation, Parody, Pastiche, Remix**
  + Consider: At what point does a work based on another work become a work of its own, if ever? Is this point different from different perspectives (i.e. the creator, the readers, the law)? (Additionally, here’s some food for thought: “Zipf’s Law,” attributed to George Kingsley Zipf, covers the fact that many forms of scientific data can be extrapolated via a certain distribution. However, Zipf is not the first to have identified this phenomenon – despite this; the law is named after him.)
  + Explore: [SiivaGunner](https://www.youtube.com/channel/UC9ecwl3FTG66jIKA9JRDtmg/videos) (YouTube channel), selected animutation and related content (available via course website)
  + Read: *Capitalism and Schizophrenia* excerpt set 2 (selected passages concerning rhizomatics.)
  + Task: Find 3-5 instances of authors directly referencing another work in the course material, in any of the explorative content, or in works in your corpus. Compile the instances in a list, and be prepared to discuss them in class next week.
* **Week 9: Context, Nuance, and Taste**
  + Consider: When should a work be taken out of context, and when shouldn’t it? Is there overlap between “remixing,” re-contextualizing, and negative appropriation? Is there a “tasteful” way to re-contextualize? Have you ever enjoyed a work generally considered or considered by those around you (or even considered by you) to be of low-quality or “problematic” in some way or another?
  + Explore: “Curated Collection of Reviews,” part 1 (available via course website), “Questionable Taste Collection” (selection of works, available via course website, optional).
  + Read: ***Professions of Taste* (pages 1-79)**
  + Assignments: Mid-semester project due.
* **Week 10: Death of the Author**
  + Consider: What is the “death of the author?” What does it entail? Is it a concept that, like the rhizome, the body without organs, or deterritorialization, can be applied outside of its original context? Is doing so indeed an instance of the “death of the author” in and of itself? Is there a relationship between excerpting a piece and the “death of the author?”
  + Explore: [*Petscop*](https://www.youtube.com/watch?v=6e6RK8o1fcs&list=PLpf9TFcVakcZzagnbfbQ4ZQMaEW7diVEo)(optional, contains disturbing content), return to *SCP Foundation* and give [SCP-5500](http://www.scpwiki.com/scp-5500) a try, if you haven’t looked at that article already (optional).
  + Read: Excerpts from *Against Interpretation*, “The Death of the Author”
  + Assignments: Write a one-page critical response on a piece in your corpus in which you take into account the author as little as possible. Then, write the response again, as closely as possible to the first one, but take into account the author as much as possible.
* **Week 11:** **Reflection Through Fan Works – Between the Deterritorialized and the Remixed**
  + Consider: What fan works have you enjoyed? Does your corpus contain any works that are fan works? How do the works in your corpus link to material outside your corpus? When does a work stop being a fan work, and start simply being a work inspired by another work?
  + Explore: “Curated Collection of Reviews,” part 2 (available via course website). Additionally, search the web for a fanwork, or choose one you’re already familiar with or that is in your corpus. Recommendations will be available on the course website, but you are strongly encouraged to find one yourself. Be prepared to discuss it critically in class.
  + Read: *Capitalism and Schizophrenia* excerpt set 3 (selected passages concerning deterritorialization.)
  + Assignments: Reflect on the possible inspirations of works in your corpus and in the course material.
* **Week 12: Esotericism, Breaking Apart, Bringing it All Together**
  + Consider: Think about “phonemes,” “tonemes,” “visemes,” even “memes,” – can there be a “narreme,” distinct unit of narrative, or a “proseme,” a distinct unit of prose? What makes a work what it is? What makes a narrative a narrative? A text a text?
  + Explore: [Library of Babel](https://libraryofbabel.info/). Also, choose a piece of explorative content from a past week and return to it, but try performing self-guided exploration of it.
  + Read: *Bluets, Codex Seraphinianus* (You’re not required or expected to read either of these texts cover-to-cover. Rather, take the time to read of them what you can.)
  + Assignments: Prepare yourself for the roundtable discussion next week. Additionally, schedule an appointment for your panel, and meet with the instructor to discuss your final project.
* **Week 13: Beginning of the End**
  + Consider: Was this course intellectually stimulating? How has what you’ve learned, for instance, during week 1, impacted your understanding of the content of week 12, or vice versa? What loose ends do you want to tie up?
  + Assignments: Roundtable discussion this week. Essentially, we will be running through each of the topics this semester and giving them a once-over, but with the knowledge of each of the other topics. We will also be discussing the final project in-depth as a class: come prepared with questions. Continue work on your final project.
* **Week 14: Middle of the End**
  + Consider: Are you satisfied with your corpus? If not, how could you reach this point?
  + Assignments: Continue to work on your final project. Also, the end-of-semester celebration and final presentations are this week. Sign up to bring in food or supplies.
* **Week 15: End of the End**
  + Consider: What’s next for you? Do you think your corpus will be a tool you use in the future? If so, do you think it will remain mostly the same, or do you think you’ll continue to change and adjust it?
  + Assignments: Final project due.