Assignment 4 / Project Specification Template required to include the Dynamic Taxonomy Interface as a class project in CS 401 (Computing Sciences)

**DTSC development: *first prototype of the interface and widgets for future user testing***

Dr. Stroulia (Head of Computing Sciences) manifested her interest in supporting my research project by including the development of the DTSC interface as a class project for one of the senior courses in Computing Sciences at U of A (CMPUT401). As a requirement to formally offer this project as a choice for a CMPUT401 group assignment, a Project Specification Template needs to be completed with information about the project background, targeted audiences, specific components, technical aspects, and future stages in the development (see annex 1 at the end of this document). In order to provide further details and following up CMPUT401 requirements, U of A system analyst from Arts Resource Centre (ARC) Omar Rodrigues-Arenas provided additional recommendations listed in a separate document (annex 2).

A Rich Prospect experience from the planning stages

From the initial planning stages, the DTSC was conceived to cope with the rigor that taxonomic information requires (de Hoog, G. S., 1981) in combination with the flexibility, interactivity and visual dynamism of a Rich Prospect Browsing interface (Ruecker et al, 2011). As such, the DTSC displays visual representations of items from given collections, combined with sophisticated tools for manipulating such complexity. My experience designing such interfaces in the past gives me the confidence to take on the challenge of directing the development of a first DTSC prototype with the advise and support of experts in the involved matters (Computing Sciences, Biology, Library and Information Studies).

The importance of expert collaboration to the DTSC project

Design practice and research implies interacting with other disciplines and areas of expertise outside the design realms. Both for collecting information and for executing solutions, designers cannot work in isolation when solving design problems. My research involves designing a system for potential biomimetic innovation and scientific exploration, which also implies positioning my work in the context of biomimicry: an emerging multidisciplinary field. For these reasons, my advisory committee consists of a mix of design and science researchers, who provide essential input to inform my work. Additionally, the DTSC interface demands a high level of technical complexity at the execution level –programming, prototyping, deploying. For this aspect, the assistance of computing science experts is essential too. Dr. Eleni Stroulia (CS, UofA) and system analyst Omar Rodriguez (ARC, UofA) provide input and support in this matter, especially to make the right decisions in terms of the choices available for programing the digital tools. Without a consistent computing science backup along this project, the possibilities to test and improve the DTSC idea may be very limited. The experts in computing sciences may provide support in the following points:

**Programing the interface**

The DTSC interface needs to be programmed as a website page, compatible with usual devices, operative systems, and browsing programs. The more compatible the environment the more effective the assessment will be for the user experience studies. This interface needs to be also compatible with additional components such as a search engine, protected log-in areas, image galleries, and other dynamic features.

**Programing the widgets**

Among the additional dynamic features, the five ‘widgets’ or data visualizations are at the core of the DTSC: Phylogeny, Homology-Analogy, Evolutive Disruptions, Maps (Research, Geography, Ecosystems) and Author’s Wordcloud. The advise of computing science experts, and the support of students and developers will be essential to create these data visualizations. Some of these will be based on open source templates such as D3 wiki contributions (<https://github.com/d3/d3/wiki/Gallery>), (e.g. wordcloud, phylogeny tree, gallery of images, or ‘sunword’ idea in the Homology), but some other may need to be entirely programmed from scratch (e.g. disruptive timeline, or the hyperspectral imaging layers in the Homology).

**Building the data-base**

Another key component of the DTSC is the data-base needed to collect and store the data, which will feed the results and visualizations obtained by the interface users. This data-base is an implicit taxonomy that runs at the background of the DTSC system. Administrating this data-base also requires a back-end interface, accessible only to administrators and authorized reviewers.

**Designing the access to data and metadata**

The collection of the data to feed the data-base implies identifying and implementing tools for accessing and data-mining available digital repositories (such as digital libraries systems or the internet). The effectiveness of DTSC in the process of finding and collecting data relevant to a Structural Colour taxonomy depends on using the appropriate tools for the task. This may involve considering specific areas such as metadata[[1]](#footnote-1) and granularity[[2]](#footnote-2) For this crucial background component, more input from experts in computing science, digital humanities, and library and information studies will be of vital importance.

Summary of the four main components of DTSC

The DTSC is a system for collecting, organizing, visualizing and sharing scientific information in interactive and reciprocal ways. For this, the complexity of the system can be understood by identifying four main components:

1. Database: back-end structure where information is collected and organized to be accessed in the front-end interface and widgets
2. Interface: front-end environment where information can be searched, found and visualized
3. Widgets: front-end wiki modules, complementing with additional data visualizations
4. Workflow Admin: back-end structure capable for administrate contributions, accounts, access privileges, peer-review process, etc.

Future of the tool

Beyond the completion of a first prototype of the DTSC interface for user testing purposes and doctoral dissertation materials, future design versions of the DTSC interface (beta versions and final versions to be deployed for public access) may involve issues of security, privacy and scalability and extendibility of the system, as well as dissemination and commercialization of accessing information. These issues will be considered at later stages of my research and rigorously analyzed in final materials of my dissertation. Yet, the goal of my doctoral dissertation is to provide a template for developing this and other tools similar to DTSC, and may not include any of these versions completed.

References

de Hoog, G. S. (1981). “Methodology of Taxonomy.” *Taxon* *30* (4): 779. doi:10.2307/1220079.

Ruecker, S.a, M.b Radzikowska, and S.c Sinclair. (2011). *Visual Interface Design for Digital Cultural Heritage: A Guide to Rich-Prospect Browsing*. Surrey, UK: Ashgate Publishing, Ltd. doi:10.1108/02640471211204150.

Tufte, E. R. (1983). *The visual display of quantitative information*. Cheshire, CT: Graphics press.

Annex 1

401 Project Specification

*If you cannot fill in all the details for each part of the template, don’t worry. If your project is selected your student team will work with you to further elaborate and specify the requirements.*

# Project Background:

In this section, please describe the motivation for this project and the general context in which the envisioned software system will be deployed. In general, we are interested in the following questions:

1. Does the project involve the development of a new system?
2. If this is not a green-fields engineering project, how will the stability of the requirements be ensured?
3. What is community envisioned to use the system to be developed?
4. What is the computer-usage experience of the envisioned users?

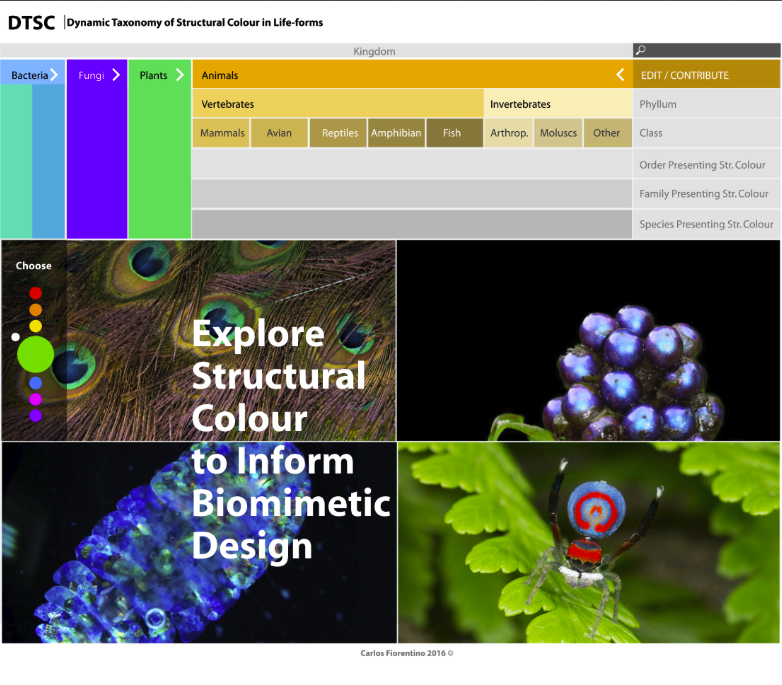
More generally, please outline existing background knowledge about your organization that the students should be aware of in order to complete the project

**A Dynamic Taxonomy Visual Interface on Structural Colour to Bridge Science and Design Innovation**

Carlos Fiorentino (Human Ecology, Art & Design)

Supervised by Dr. Megan Strickfaden (Human Ecology, U of A) and Dr. Tomislav Terzin (Biological Sc. U of A, Augustana)

The emerging discipline of Biomimicry studies “nature’s best ideas” with the purpose of emulating them to address human problems by design. One of this ideas is Structural Colour –in physics “light interference”– which is a way of achieving colour without relying on pigmentation or chemical coloration processes, but rather by adding “information” to material surfaces at the nano-scale. In nature Structural Colour is observed in an abundant number of species across animals, plants, fungi and bacteria. While Structural Colour has been initially explored and tested for application in new materials and technologies, it is still at an early stage of development, and no substantial progress has been observed in last years. Part of this situation is due to limitations in communication when crossing disciplinary boundaries. Existing information on structural colour is presented in a language rooted in hard sciences –biology and physics– and may limit access and understanding for design inspiration and implementation to those working outside science. My doctoral quest is focused on addressing such limitations, and is articulated in one main research question: *How can available scientific information/ knowledge on structural colour be more accessible to biomimetic practitioners?* As a result of my research, I aim to provide dynamic digital tools to facilitate the access to available knowledge on structural colour in nature, and may fill communication gaps between scientists and designers involved in current biomimetic projects, as well as inspire new ones.  My proposal is called “Dynamic Taxonomy of Structural Colour –DTSC” (fig.1). It is an interface with the capacity of visualize and arrange scientific information based on a Rich Prospect Browsing experience, where users customize their search. In order to create such dynamic tool, I need to move from mock up design versions done to first prototype versions for future user testing. This involves choosing the appropriate formats, language and other technical aspects of the interface, and visualization widgets included. It also involves setting up an initial data-base capable of accessing external repositories (e.g. digital libraries systems) and to be fed by Data and Text analysis tools. The DTSC interface is also an opportunity for scientist to contribute with feedback, comments and new entries about Structural Colour. For this, an important part of the interface is the “contribution layer” which collects contributions to be peer-reviewed.



*Fig.1: mockup homepage of DTSC interface*

In general, we are interested in the following questions:

1. Does the project involve the development of a new system? YES Is there code to be reused? POSSIBLY, THE WIDGETS MAY BE BASED ON EXISTING VIZ, FROM D3, ETC. (e.g. wordcloud, phylogeny tree, ‘sunword’ idea) Is there data to be ingested in the system? YES, INITIALLY MANUALLY (SPREAD SHEETS), THEN IN FUTURE STAGES BY DATA-MINING, ACCESSING DIGITAL REPOSITORIES
2. If this is not a green-fields engineering project, how will the stability of the requirements be ensured?
3. What is community envisioned to use the system to be developed? BIOMIMETIC DESIGNERS INTERESTED IN COLOUR STUDIES, SCIENTISTS (EVOLUTIONARY BIOLOGISTS, PHYCISISTS IN BIOPHOTONICS), AND BROADER AUDIENCES INTERESTED IN STR.C FROM ACADEMIA AND INDUSTRY
4. What is the computer-usage experience of the envisioned users? BASIC TO INTERMEDIATE

More generally, please outline existing background knowledge about your organization that the students should be aware of in order to complete the project

THIS IS PART OF MY DOCTORAL RESEARCH. THE PROTOTYPE RESULTED FROM THIS EXPERIENCE WILL BE USED FOR USER TESTING (TARGET: DESIGNERS AND SCIENTISTS). THE FINDING OF THIS STUDY WILL BE USED AS PART OF MY DISSERTATION MATERIALS, PUBLICATIONS AND PRESENTATIONS IN SYMPOSIA.

# Project Sponsor and Stakeholders:

Please name the person(s) from your organization that will be involved in the project and their roles.

1. Who will sign off? SUPERVISOR: MEGAN STRICKFADEN; EXT.ADVISOR: TOMISLAV TERZIN
2. Who will be available to meet with the students? ME. EVENTUALLY MY ADVISORS TOO. We anticipate that at least four meetings will take place between the development team and the client team.

# Key Functional Requirements of the System:

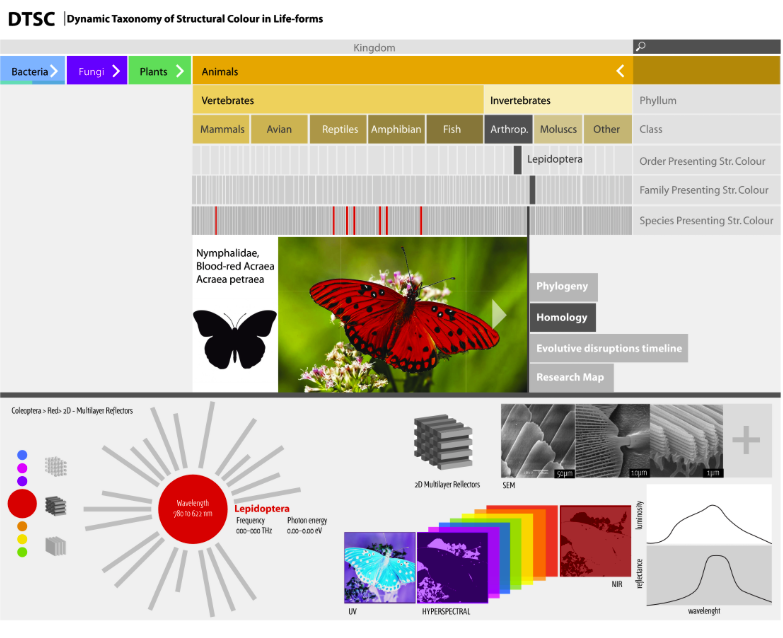
Provide a list of at most ten functions that the system should have. Consider organizing these functions according to the types of users who will exercise them.

DESIGNERS:

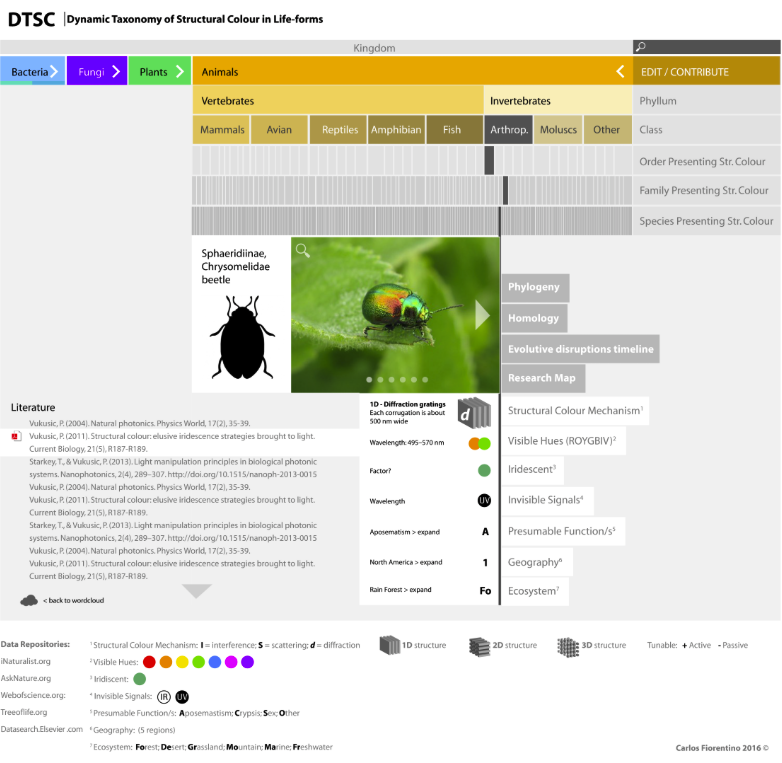
* SEARCH FOR A SPECIFIC HUE (COLOR) AND FIND SPECIES ASSOCIATED TO THAT HUE (Fig.2)
* GET RICH DESCRIPTIONS, IMAGERY AND SPECS ABOUT THOSE SPECIES (Fig.2)
* FIND LITERATURE ASSOCIATED (Fig.3)
* FIND TOOLS TO UNDERSTAND PATTERNS, CONNECTIONS AND INTERRELATIONS WITHIN THOSE SPECIES (WIDGETS, GALLERIES) (Fig.4 and 5)
* FIND CASE-STUDIES OF IMPLEMENTATION ASSOCIATED
* GENERAL DB SEARCH BY KEYWORD

SCIENTISTS:

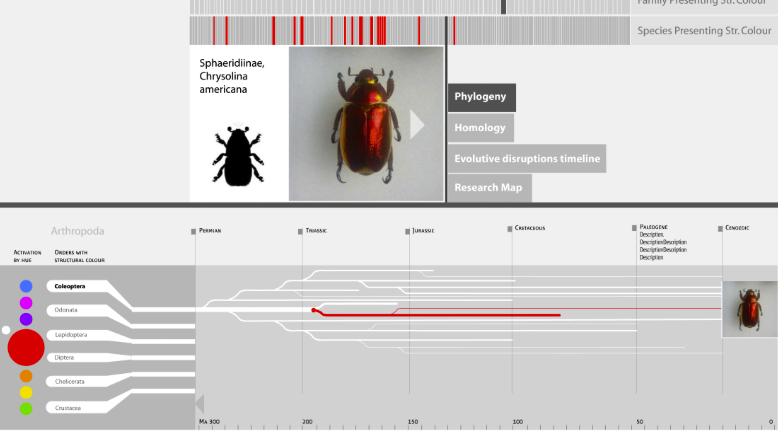
* FIND SPECIFIC TAXA (SPECIES, FAMILIES) AND AVAILABLE INFO ABOUT STR.C. (Fig.6 and 7)
* FIND TOOLS TO ANALYZE PATTERNS, CONNECTIONS AND INTERRELATIONS WITHIN THOSE SPECIES THOUGH STR.C.
* FIND LITERATURE ASSOCIATED
* CONTRIBUTE TO THE DB WITH FEEDBACK, NOTES, QUESTIONS, NEW FINDINGS, ETC. (Fig.8)
* USE DTSC AND ITS WIDGETS FOR EDUCATIONAL PURPOSES
* GENERAL DB SEARCH BY KEYWORD



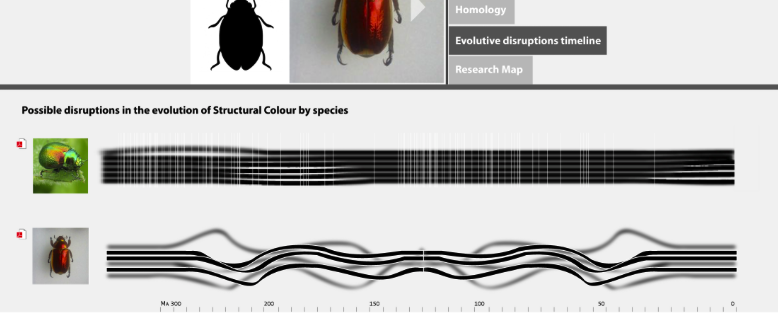
*Fig.2: the top of the screen shows one identified species found under ‘red’ structural colour; the bottom shows the widget “Homology” activated, which gives access to scientific details of that particular species.*

**

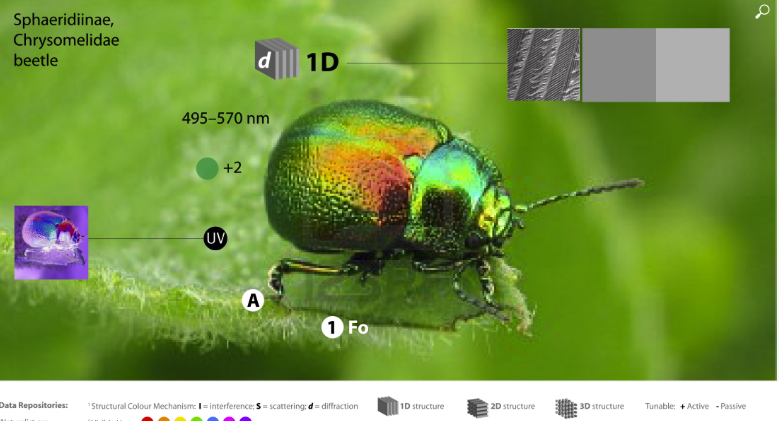
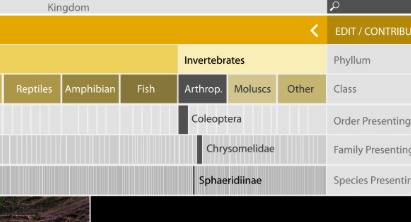
*Fig.3: Left corner of the screen shows a list of publications associated to a particular species selected.*

**

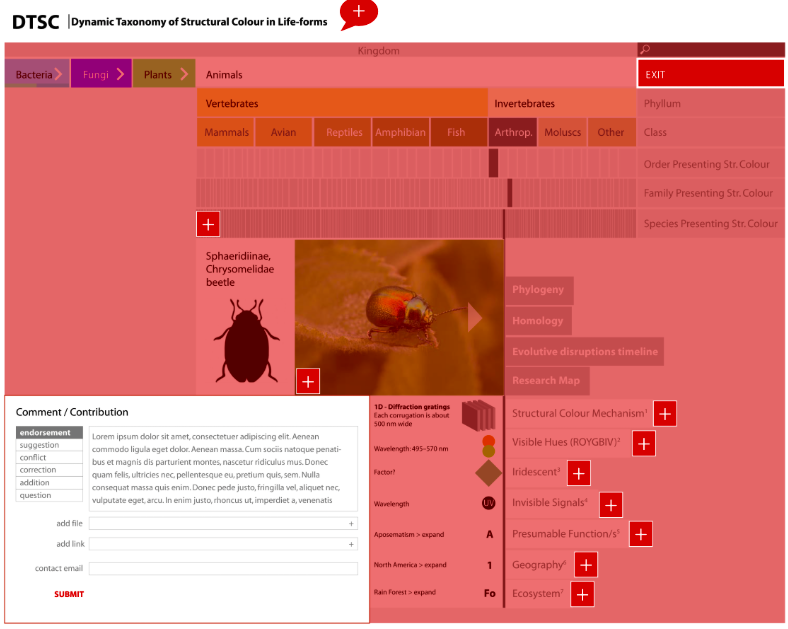
*Fig.4: the bottom of the screen shows the widget “Phylogeny” activated, which provides an evolutionary branched timeline of a particular species containing a particular structural colour.*

**

*Fig.5: a different widget “Evolutive Disruptions” shows another, this time comparing two species and visualizing the number and intensity of disruptions associated with structural coloration change.*

**

*Fig.6 and 7 show detail of accessing one species from the main taxonomy, and the results showing main aspects supported by photographic material*



*Fig.8: the Edit / Contributions ‘layer’ of the interface, would allow registered users to submit comments, suggest changes to the information, propose new entries, etc. for peer-reviewing and eventually update the DTSC data-base.*

The development team will elaborate upon these functional requirements as part of their project.

Attach (or point to) any other documents that already exist that may be relevant: (a) procedure manuals, (b) competitor/related products, …

MOCKUP VERSION: <https://xd.adobe.com/view/db3f8da5-fc16-4a54-b75a-68ff8eba2f9e/>

FLOWCHART OF THE PROJECT:

<https://carlosfiorentino.files.wordpress.com/2016/11/flow-diagram-copy.jpg>

(password protected: carlos)

FLOWCHART OF DTSC INTERFACE CONTENT: <https://carlosfiorentino.files.wordpress.com/2016/11/dtsc_flowchart.pdf>

(password protected: carlos)

# Technical Requirements:

What programming language and/or development tools must be used? SUGGESTED (TO BE DISCUSSED):

**Backend:** Linux or whatever works for CMPUT401 students

**Data-base:** Mongo DB or any other that supports changes in the scheme, including search modes based on geographic location, tagging, etc.

**Language:** RUBY or PYTON

**Framework:** RAILS (if Ruby)

**Format of Data:** JSON or XML

**Front-end (interface) and widgets:** JAVA Script (D3 wikis use Java)

**Workflow:** System capable for admin, contributions, sign up, accounts, peer-review, privileges, etc.

What type of operating systems does it need to run on?

ANY. IT NEEDS TO BE A VERY COMPATIBLE TOOL RUNNING IN ANY DEVICE, OS AND BROWSER

Is the system to be web-enabled (i.e., all interactions will be through a web browser)? YES. It should work in any updated browser

Will your organization be providing the development platform?

STILL NOT SURE. Possibly, first prototype could be hosted on campus by ARC or IST as a virtual server.

Will your organization be providing the deployment platform?

TBC. OPTIONS SUGGESTED: WestGrid (Compute Canada), ARC, IST or Docker

Maintenance: ARC or IST

**Required deliverable from CS 401:**

* Prototype running on line for testing purposes,
* Users’ info/access,
* Documentation,
* Source Code,
* Copies and Backups

# Quality Requirements:

List the three most important quality requirements (i.e., user-response time, throughput, usability, security, privacy, extendibility) for the envisioned system.

THIS IS A PROTOTYPE TO BE DEVELOPED FOR TESTING THE INTERFACE WITHIN A SMALL AUDIENCE OF DESIGNERS AND SCIENTIST, WITHIN THE DOCTORAL STUDIES’ PURPOSE ONLY. THE PROTOTYPE SHOULD SIMULATE REAL TIME INTERFACE USE. A FUTURE PUBLIC VERSION DEVELOPED WILL CONSIDER ISSUES OF SECURITY, PRIVACY AND EXTENDIBILITY.

Annex 2

Additional input on technical specifications for DTSC

*Consultant: Omar Rodriguez-Arenas, Systems Analyst, Arts Resource Centre,*

*University of Alberta*

Recommendations for DTSC interface development:

**OS:** any

**Backend:** Linux or whatever works for CMPUT401 students

**Web:** any updated browser

**Data-base:** Mongo DB or any other that supports changes in the scheme, including search modes based on geographic location, etc.

**Language:** RUBY or PYTON; Framework RAILS or …

**Format of Data:** JSON or XML

**Front-end (interface) and widgets:** JAVA Script (D3 wikis use Java)

**Workflow:** System capable for admin, contributions, sign up, accounts, peer-review, privileges, etc.

**Temporary host for prototype:** ARC or IST, virtual server

**Permanent host, deployment options:** Westgrid by Compute Canada, ARC, IST or Docker

**Maintenance:** ARC or IST

**Required deliverable from CS 401:**

* Prototype running on line for testing purposes,
* Users’ info/access,
* Documentation,
* Source Code,
* Copies and Backups

1. **Metadata:** data that describes other data, summarizes basic information about data, which can make finding and working with particular instances of data easier. [↑](#footnote-ref-1)
2. **Granularity:** The *granularity* of data refers to the size in which data fields are sub-divided. For example, *coarse granularity* as a single field, *fine granularity* as multiple fields, or even *finer granularity* at the level of details. Finer granularity has overheads for data input and storage. This manifests itself in a higher number of objects and methods in the object-oriented programming paradigm or more subroutine calls for procedural programming and parallel computing environments. It does however offer benefits in flexibility of data processing in treating each data field in isolation if required. A performance problem caused by excessive granularity may not reveal itself until scalability becomes an issue. [↑](#footnote-ref-2)