**MARKET RESEARCH ANALYSIS**

***CSIRO***

**Text Annotation Tool**

Presented By:

*Degala Anurag Yadav*

*Harsha Lakshmipuram*

*Amelia Wigg*

*Malayamarutham Sai Pradyumna*

*Sai Gautam Mandapati*

*Pavani Gujjarlapudi*

*Eswar Sai Krishna Gowra*

*Tharindh*u

Table of Contents

**OUR AIMS......................................................................................................................... 3**

**OUR DELIVERABLES........................................................................................................... 3**

**TARGET MARKET............................................................................................................... 3**

**TARGET NEED.................................................................................................................... 3**

**COMPETITOR ANALYSIS..................................................................................................... 4**

*Bioportal* **........................................................................................................................... 4**

*LightTag***............................................................................................................................ 9**

*TagTog***............................................................................................................................ 13**

*Diigo***................................................................................................................................ 17**

*ZipBoard***.......................................................................................................................... 19**

*Asana***.............................................................................................................................. 22**

**FEASABILITY ANALYSIS..................................................................................................... 24**

**REFERENCES.................................................................................................................... 25**

Market Research Analysis

OUR AIMS

To develop a web service that annotates text with definitions, taken from controlled online vocabulary services. Make it easier, simpler and more accurate to get a definition, which has been provided by a collaboration of online vocabulary services. A quick understanding which is enabled by hyperlinking technical terms to their corresponding definitions.

OUR DELIVERABLES

A tool to:

* Detect words or phrases with in a text that have definitions in one or more source vocabularies
* Allow the user to select which definition applies to the usage of the term in the context
* Allow the user to select an alternate or preferred term to replace the original
* Insert a hyperlink to the definition

TARGET MARKET

The target market of our tool includes anyone who is undertaking online research using the internet. There are no gender specifications of our target market, nor is there a specific age group due to the tool being aimed at anyone undertaking research. However, the target market must be old enough to be undertaking research. The text annotation tool is particularly helpful for reading and understanding research papers and complex term, therefor the market is targeted at those likely to examining research papers for as part of their job, education or personal research.

MARKET NEED

**Description of the tool:**

Quite often technical documents have complex terminology and specific words or phrases can denote precise definitions of some concept. In most of the cases a single term has more than one meaning. By hyperlinking the text with precise and accurate meaning, we can develop clarity and specificity of terms and meanings. This project will develop a tool and web micro service to enable semi-automated annotation of text with hyperlinks to the complex definitions of terminology used in research papers.

**Need for this tool:**

Commonly in research papers there are a set of typical words which can derive multiple meanings, easily causing confusion or misunderstanding. This highlights the need for a tool that provides users with the assistance of hyperlinking text which can provide easy definitions or the term. Additionally, the user can be redirected to the webpage which can assist in finding a particular meaning or search for terms in a particular field.

**Demand of this tool:**

There a small number of existing tools which assist in annotating words and providing links to users. There is high demand for these type of tools, due to the complexity and array of meaning in research papers. There are a lot scientific terms which user may require assistance understanding or comprehending. A tool could have a range of available ontologies and therefore could be useful in a range of different fields.

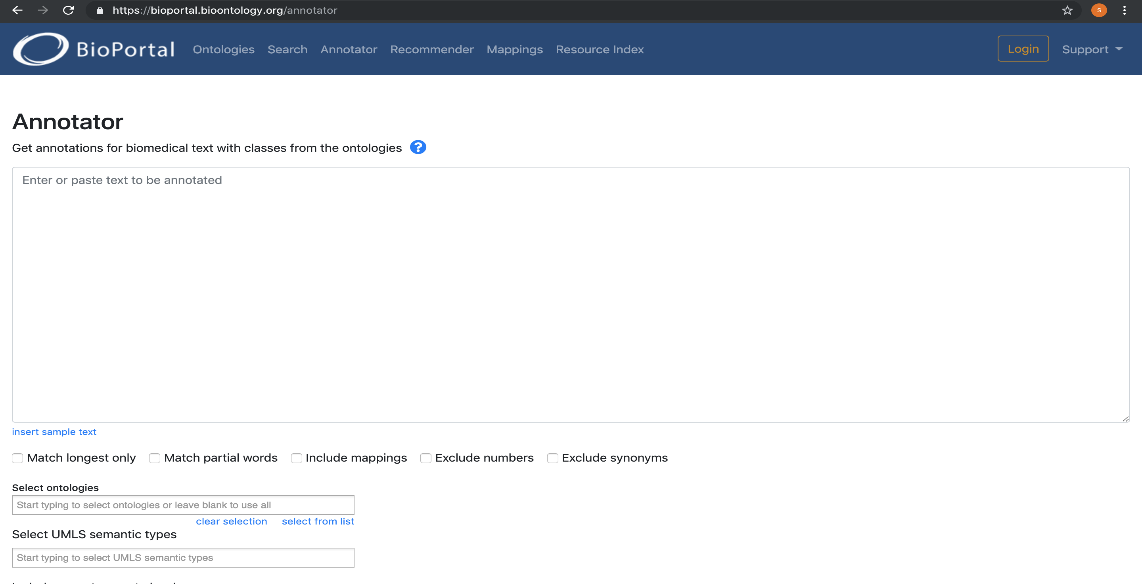
COMPETITOR ANALYSIS

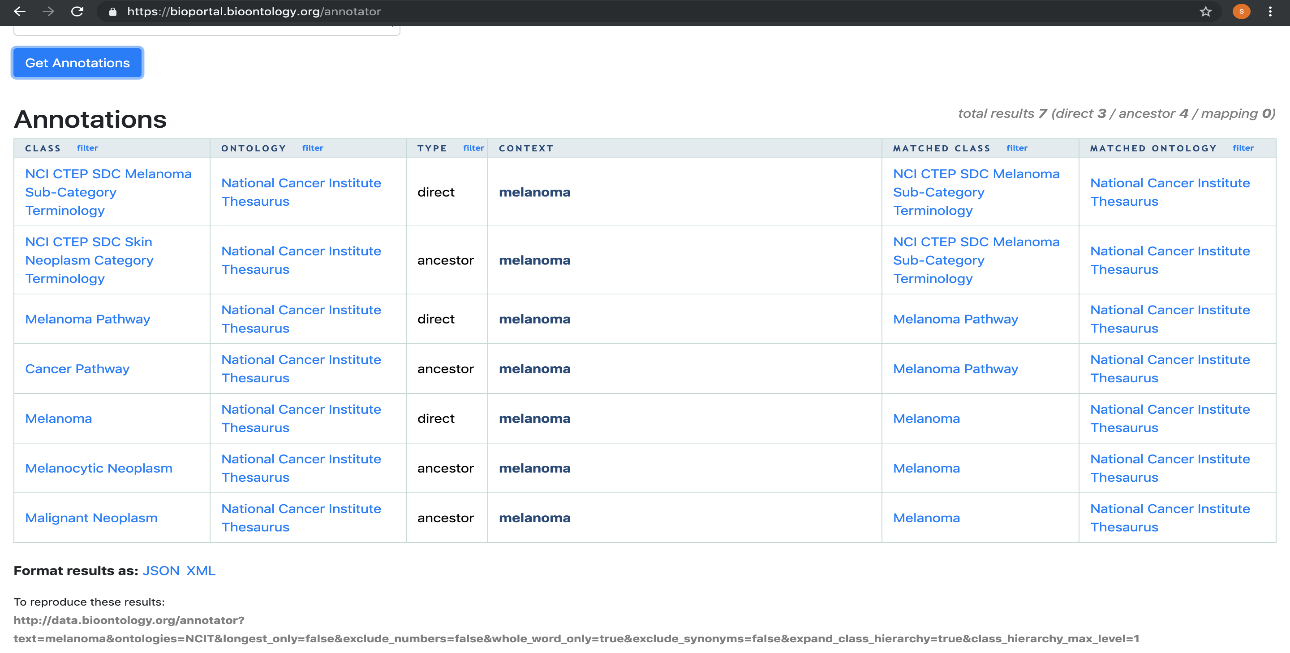
*Bioportal*

***Description of the tool***

Bioportal is a bio ontology website which assists in annotating words by providing links for various different websites, in which the user can obtain meanings and definitions of words. The website includes an ontologies library where the user can explore different libraries dependant on their specific area, examples include National Cancer Institute Thesaurus (NCIT) and other medical libraries.

The below image depicts the process where text can be entered and searched in the NCIT ontology search library. The user needs make certain selections before clicking the ‘get annotations’ tab, in order to derive results.

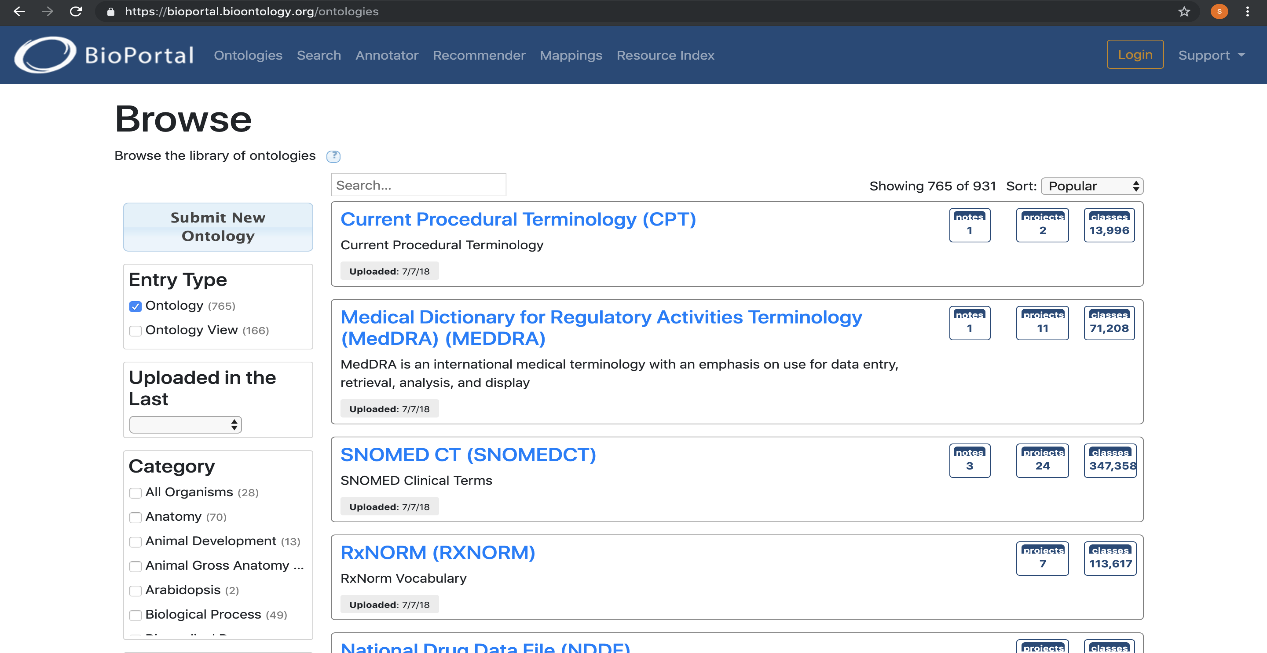




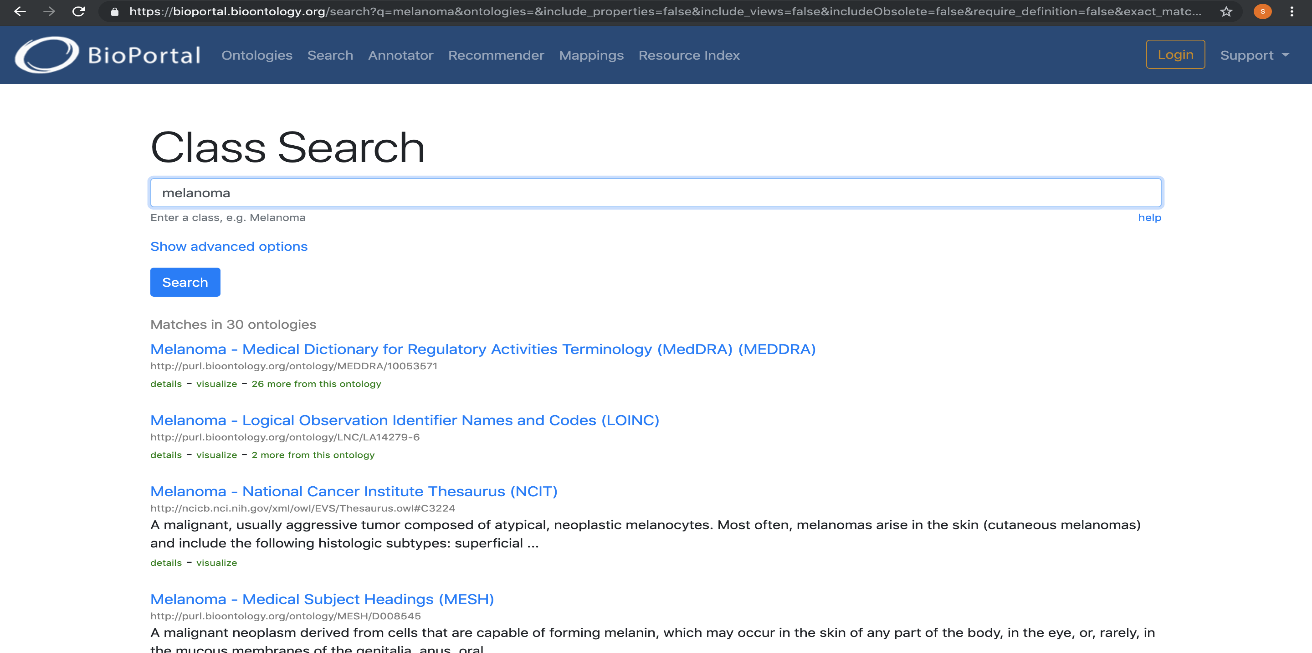
***Features of the tool***

Below six of the main features of the tool are outlined with corresponding images to represent the feature.

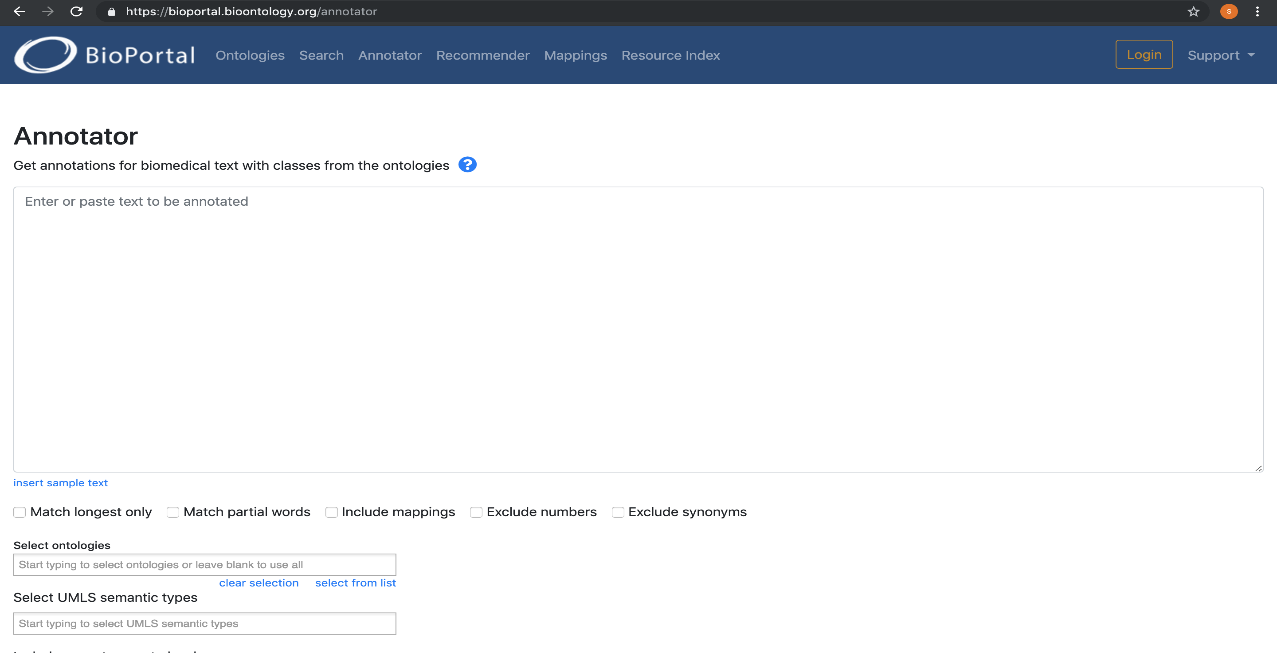
1. The first feature this website provides is the ontologies tab, where the user can select different ontologies. The user can submit a new ontology to add into the library. The user can search based on the entry type eg. category, format, group.



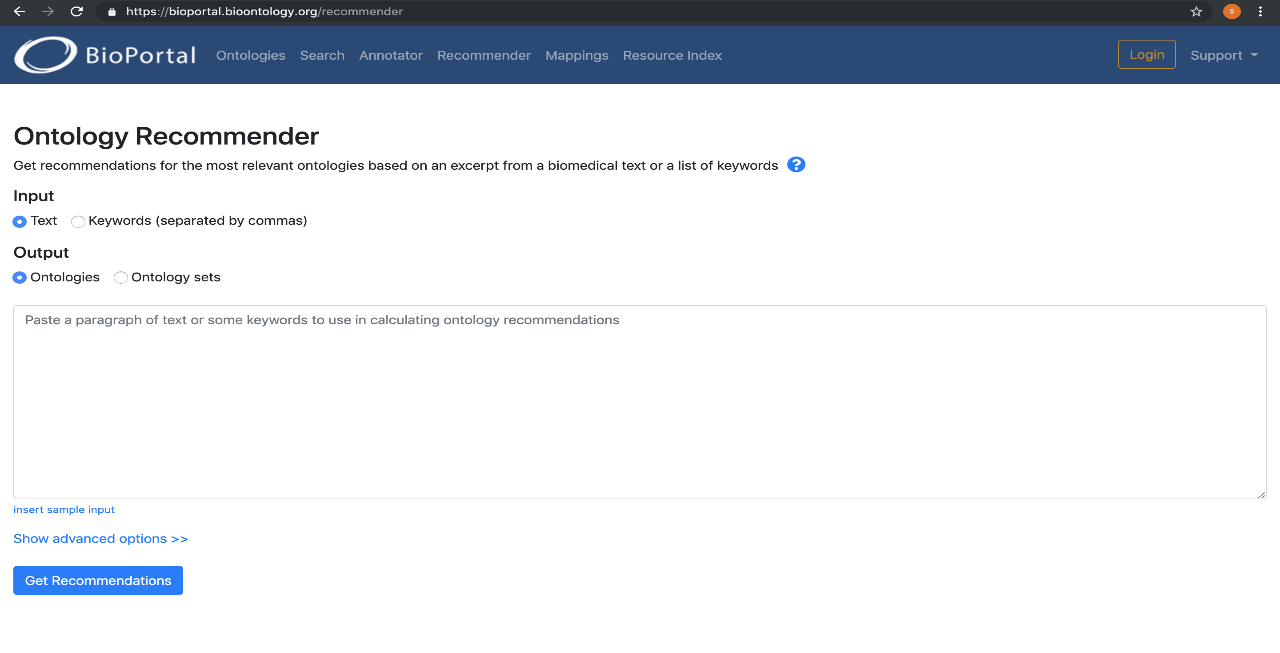
1. The second feature that Bioportal offers is the class search option. When the user enters certain medical terms, the selected class will be matched with corresponding ontology libraries and the results will be provided.



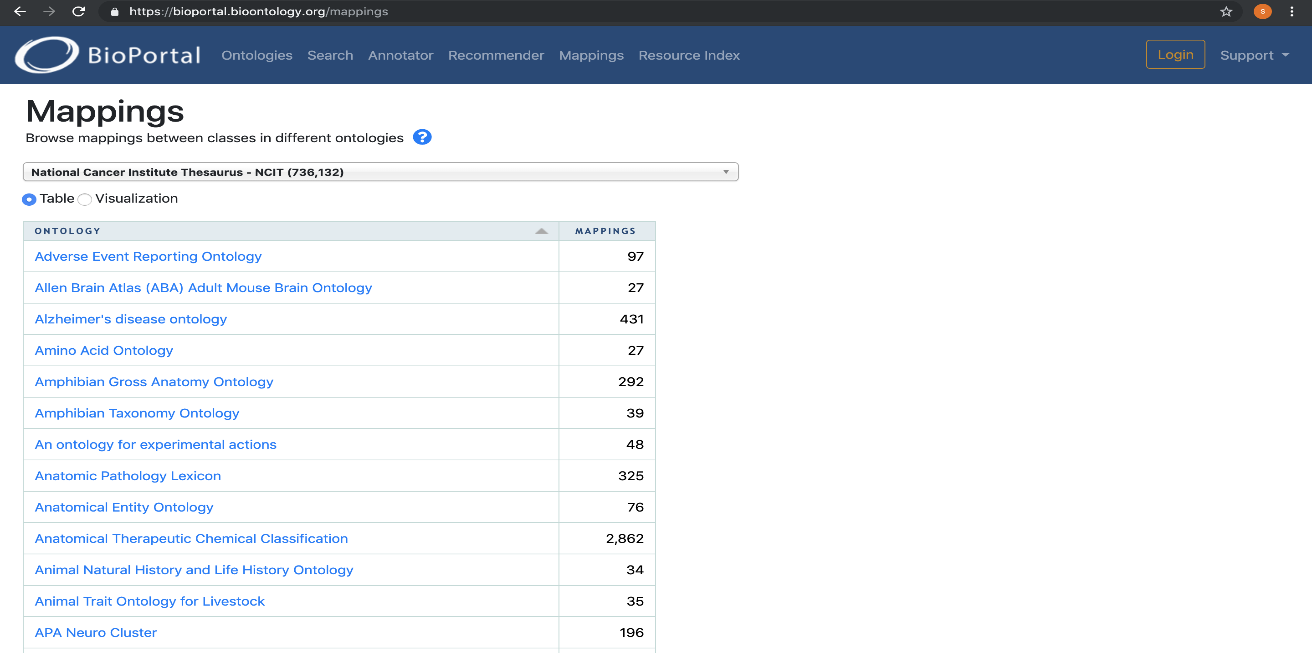
1. Another feature Bioportal offers is the Annotators option. In which the user can get a range of different data from a variety of different annotators.



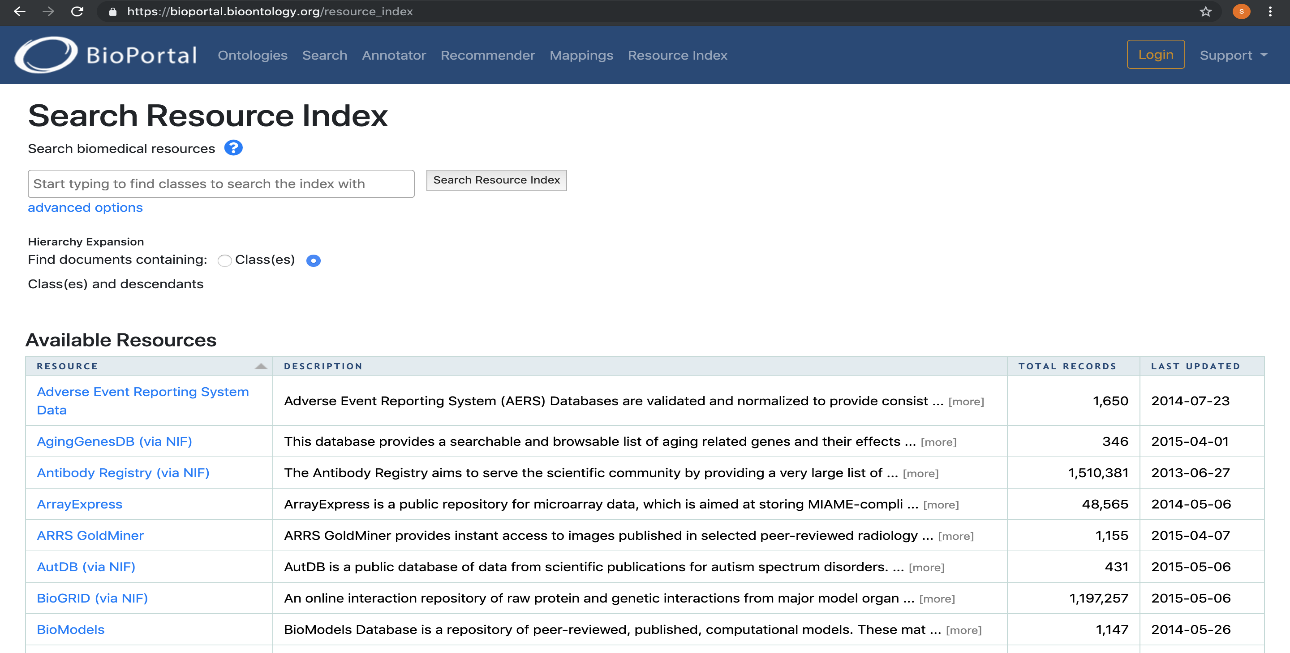
1. An additional feature is the Ontology Recommender. The user can get recommendations for ontologies based on the text entered. There is an additional option of check boxes, where the user can select text or keywords and ontology or ontology sets. The results are provided to the user based on their selections.



1. Bioportal offers a Mapping feature, which allows the user to map between different classes in different ontologies.



1. Another features is the Resource Index, which provides the use the option to select data from various Biomedical resources.



***Features that align with our aim***

Many of the Bioportal features align with our project aim. However, objectively the most important feature that aligns with aim and deliverables is the annotator. As our project is based around the features and uses of text annotation. We will integrate two API’s from the existing ontology libraries and develop an annotator.

***Features that would need to be altered***

The tool is very similar to the tool we are developing and therefor there are not many features that would need to be altered to align with our aims and deliverables. However, certain features could be added to enhance the tool to encompass all of our requirements.

***Benefits of the Tool***

* Free web page annotation tool
* Advance search options available based on databases
* Customisation options according to user queries
* Easy to use and handle
* Get recommendations for the input query we provided

***Limitations of the Tool***

* We can search only ontology & Biomedical bases keywords
* Limited to Biomedical uses not compatible all fields

***User Scenarios***

* The user needs to open the bio-portal browser and just search the annotated word.
* The user doesn’t need to login or signup for annotated words.

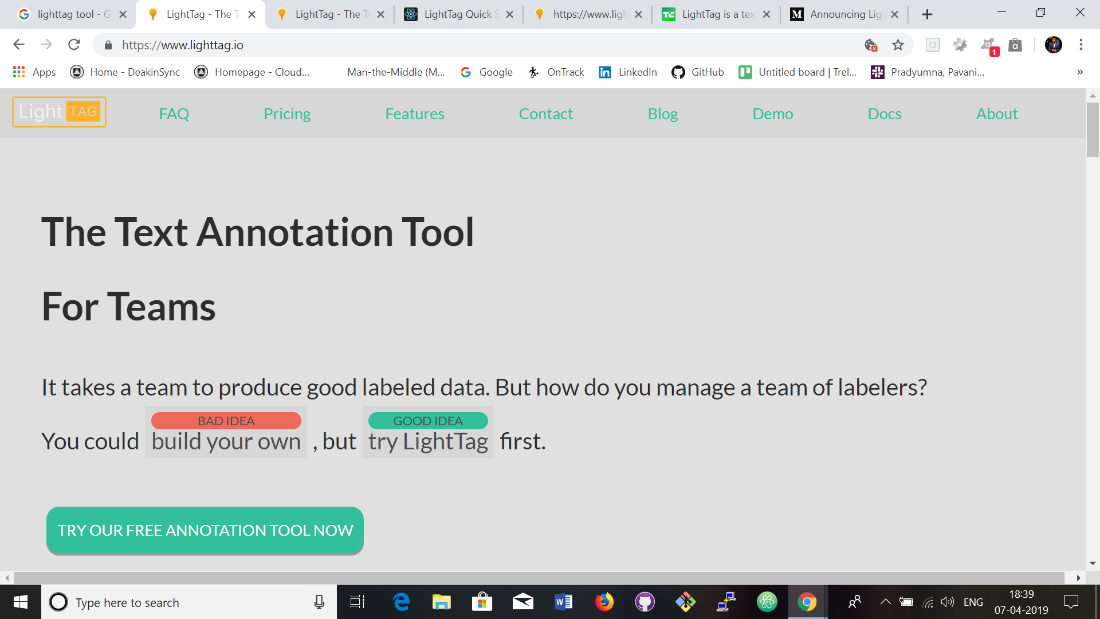
***Price***

Bioportal annotator is an open source tool, free for anyone to use at no cost. There is a login option where user is able to create their account and download the latest versions of ontologies for their work.

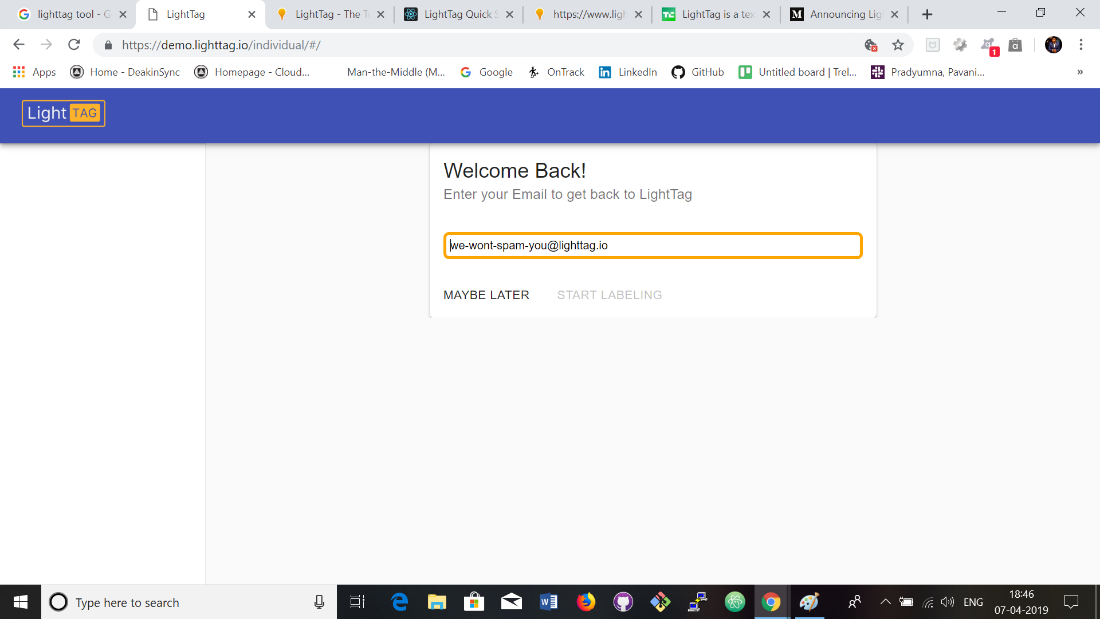
*Lightag*

***Description of the tool***

Lightag is a useful text annotation website/tool. It is predominantly used as text annotation platform for data scientists creating AI training data. LightTag aims to solve one of the main bottlenecks of the deep learning-based AI development.

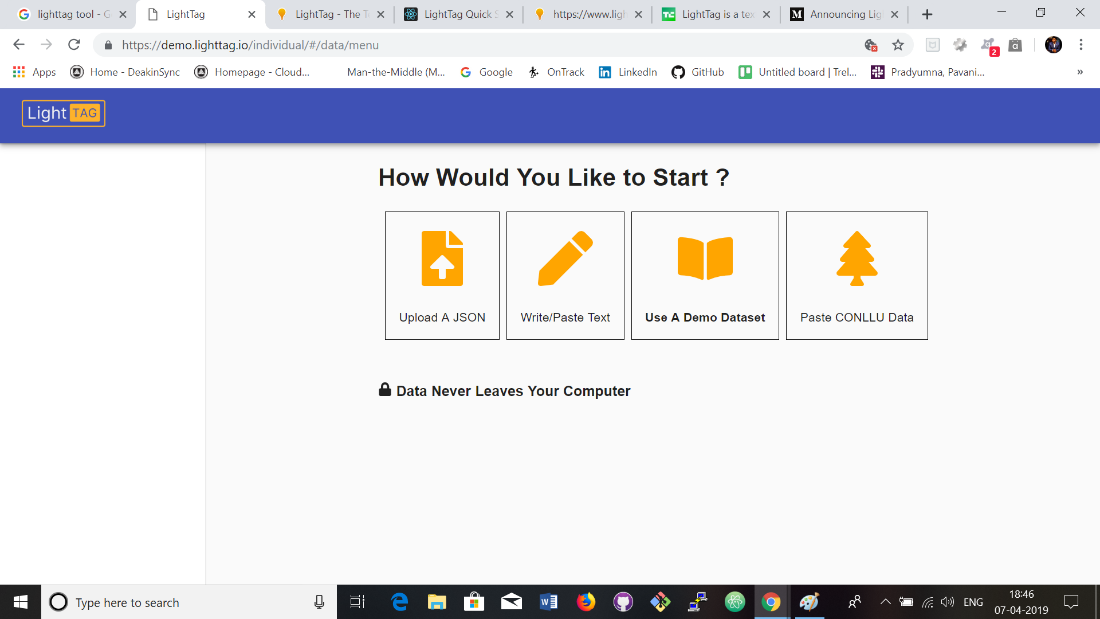


***Features of the tool***

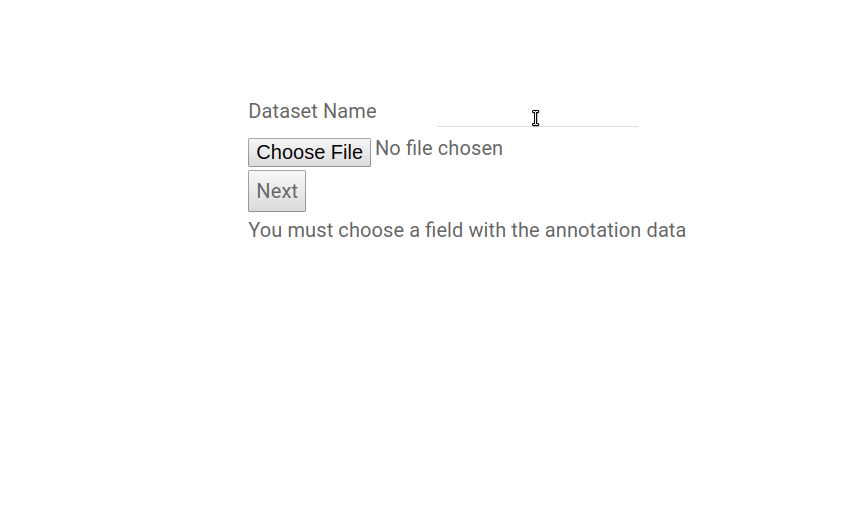
Lightag includes the feature of its own ‘AI’, which could learn from previous labelling and make annotation suggestions. Additionally, the platform also automates the process of managing a project. It assigns tasks to labellers ensuring there is enough overlap and duplication to keep accuracy and consistency high. LightTag’s annotation interface is both easy to navigate and optimised to maximize annotator productivity, while minimizing subtle biases that a UX can introduce. LightTag’s interface allows annotation with a mouse or keyboard. Keyboard shortcuts assist by making annotators work fast and easy. It can annotate any language like English financial jargon, Chinese legalese, Arabic tweets or Hebrew medical records. LightTag allows you to define an annotation task and specify how many annotators you’d like to have annotating each example. LightTag will assign work to annotators as they arrive, making sure that you reach your annotation goals in the most efficient way possible.

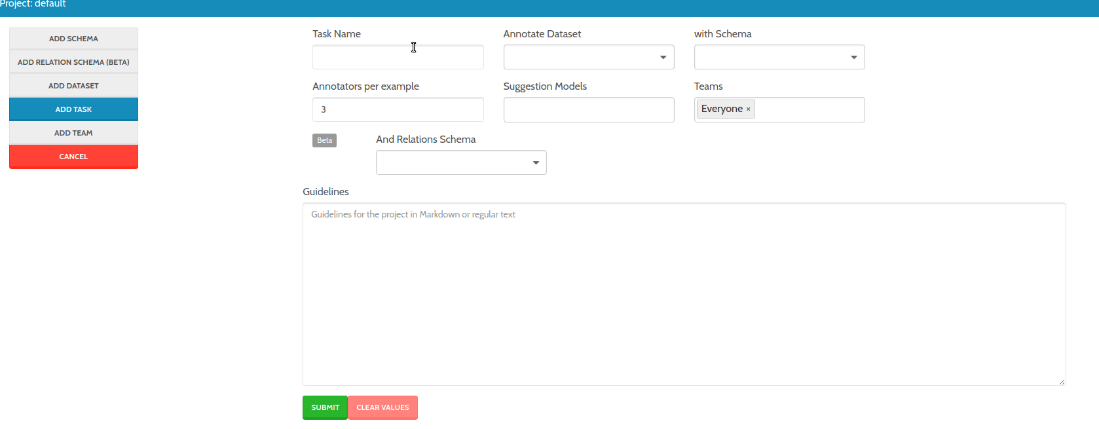
LightTag's annotation interface features:

* Keyboard Shortcuts
* No tokenization assumptions
* Full Unicode Support
* Subword and phrase annotations
* RTL and CJK languages
* Entity, Classification and Relation annotations.



In the above image we can see four options “Upload a JSON”, “Write/ Paste Text”, “Use a Demo Dataset”, “Paste CONLLU Data”. LightTag expect you to upload your data as a JSON array of objects.





***Features that align with our aim***

Whilst Lightag has many features it does not completely align with the tool we aim to develop. It does however annotation capabilities which are similar to our aims. Some of the annotation or user features could be similarly designed in our project, although Lightag is aimed at AI, which our project does not have much focus on.

***Features that would need to be altered***

As discussed above, some of our projects tool features could be based off Lightag’s features although, there would need to be significant changes to many of the features to align with our aim and deliverables.

***Benefits of the Tool***

* Can annotate in any language
* Measures inter annotator agreement to understand the quality of annotation
* Continuously boot straps the projects and provides suggestions
* Can add new meaning from our own library and upload it to LightTag with Rest API

***Limitations of the Tool***

* Provides hourly rate for engineers to build own labelling tools
* Comes with only in on-premise version

***User Scenarios***

* The user needs to sign in first to use the tool.
* The user can see various options like upload JSON, write or paste text, use demo data base and paste conllu data.
* The user can choose any option regarding the project and annotate the words.

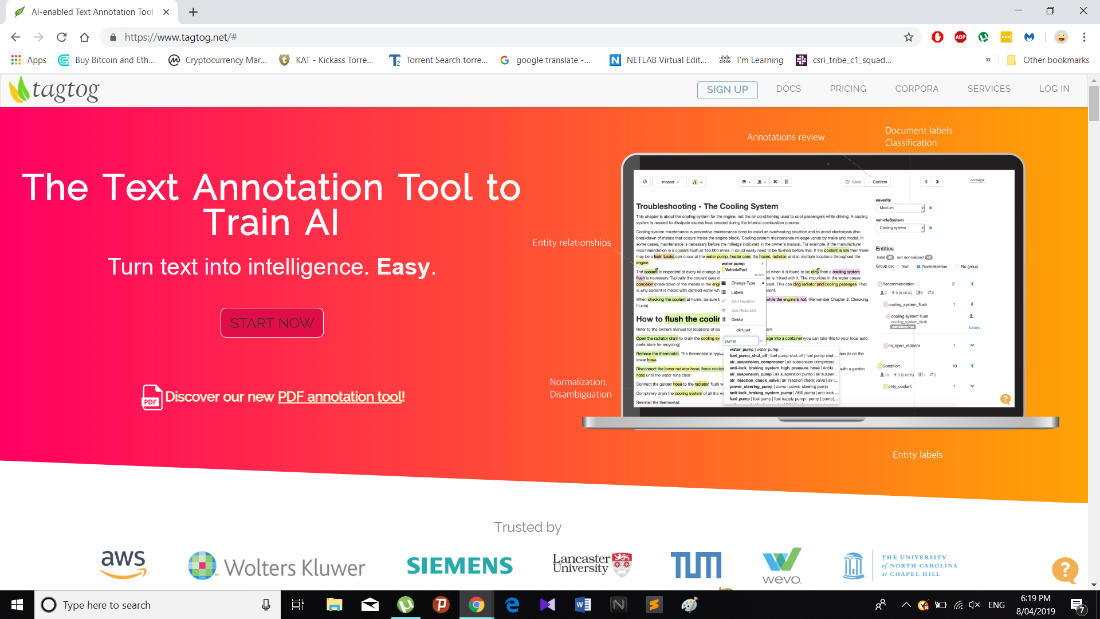
***Price***

The tool is It is free of cost and only requires the user to create an account in order to use the tool. The tool can be downloaded and is especially useful for team projects.

*TagTog*

TagTog is a Text Annotation Tool to Train AI. Mainly TagTog use for train text mining methods. Generally, there are two types of people use this tool,

1. People who want to have annotated and register part of the knowledge with in their organization.
2. People wo need want to train their own material models.



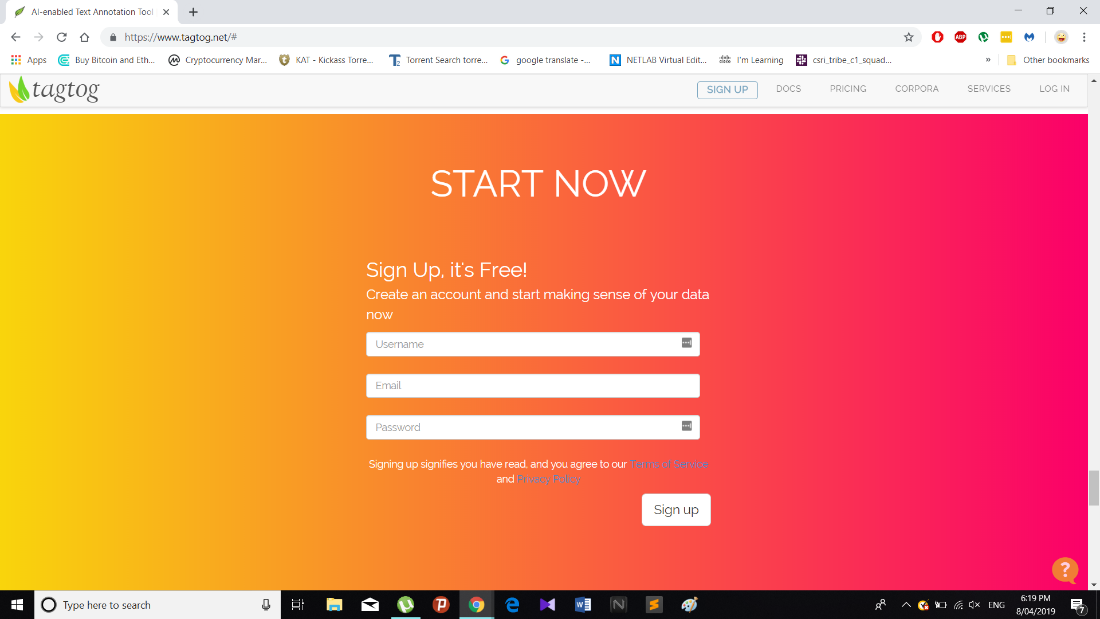
***Description of the tool***

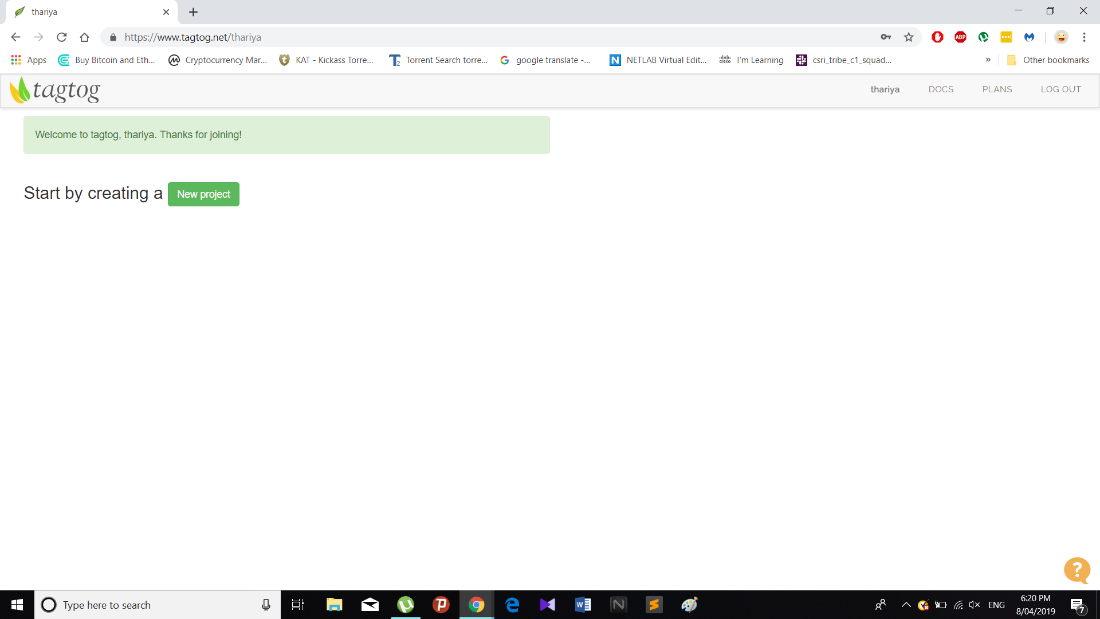
Users can install TagTog for their computer or they can use website to begin their work. Tagtog is mainly used for group projects. The admin of the project has the ability to add others to the project and allow all users of a group to contribute.

The annotation features of this tool have the ability to be used either manually or automatically. The opportunity for collaboration is also a key feature of the tool for projects. Not only that, below options are the major part of TagTog,

* Index data or augment
* Find custom insights inn text automatically
* Classify text
* Train your own AI

The basic process of setting up an account requires the user to sign up and select the option to start a new project. The below images depicts the process of beginning a new project.

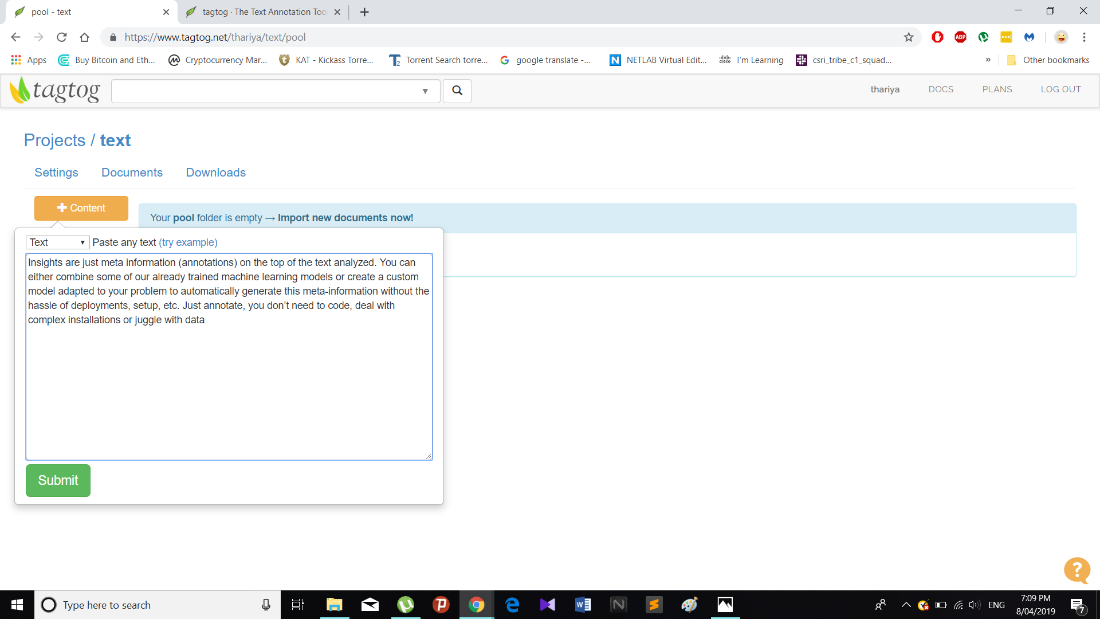




***Features of the tool***

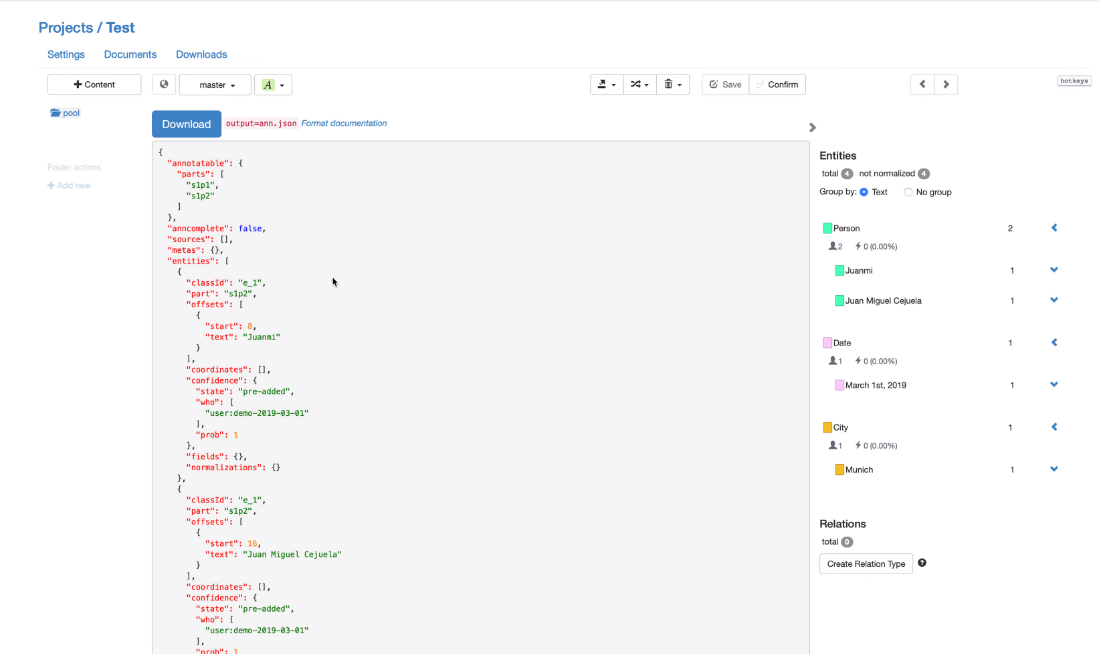
Two of the main features of Tagtog outlined below include, adding a document to the project and manual text annotation.

1. Adding a document to project:



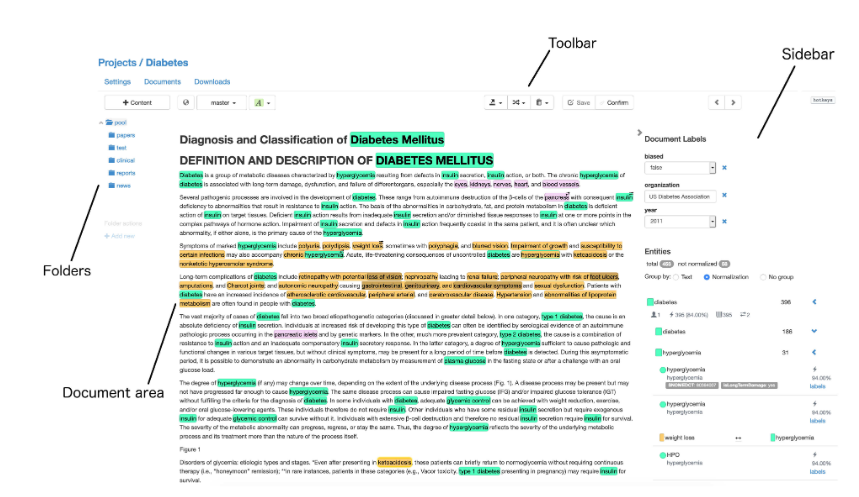
This stage is very critical, as the project owner/admin has to define the requirements and resources for the project. Not only that, TagTog has several functions such as Annotables, Entities, Dictionaries, Relations, Document Levels, Entity Labels, Webhooks, Annotations, Members and admin.

Additionally, only members of the team can access and edit any of the coding, which emphasises the importance of inviting and establishing a group.



1. Manual text annotation:

This feature is used in order to quickly annotate and normalize entities.



***Features that align with our aim***

Some of the Tagtog features that align with the aim of our project include automatic text annotation and the ability to have many members of a team working a project at the same time. Other features involving text annotation also align with our aims, although it executed with a different purpose to our project.

***Features that would need to be altered***

There are many features that would need to be altered in order to align with our aims. Tagtog does has text annotation capabilities that generally have a similar purpose to our project, however it does not have the precise annotation capabilities that our project requires.

***Benefits of the Tool***

* Customization the options.
* Easy to use and handle.
* Can select different members.
* Has a good speed.

***Limitations of the Tool***

* Only leader can add or delete comment
* Need to log in every time.

***User Scenarios***

* The user needs to sign in and create the project first.
* The user can add the document to the project and annotate the text.
* The user can edit the code of the project.

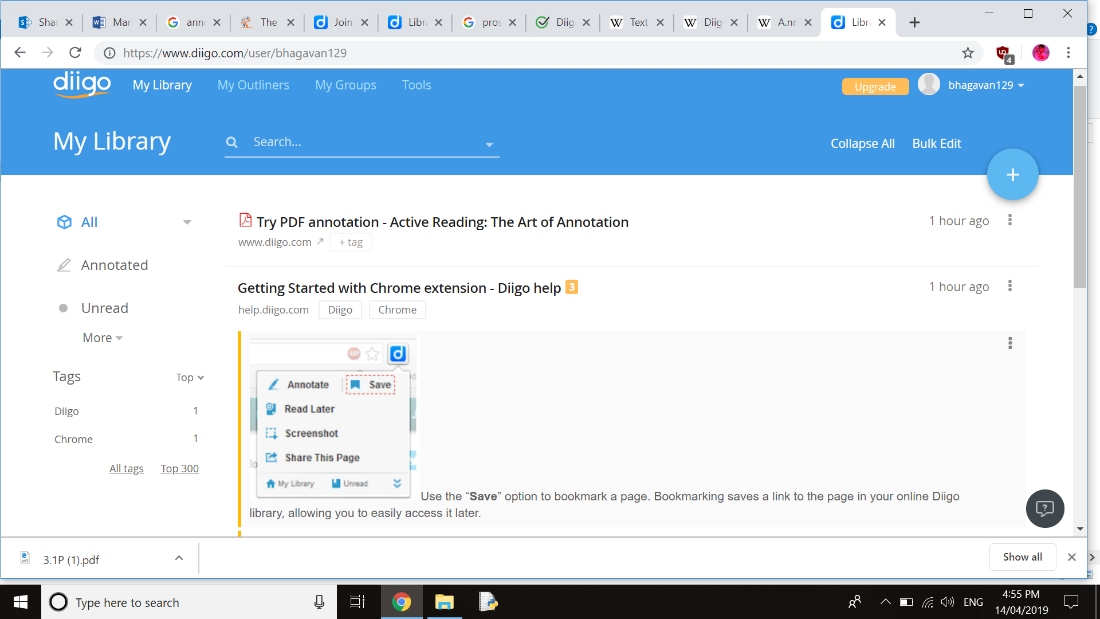
***Price***

The tool of is freely available for anyone to use and can be downloaded by anyone.

*Diigo*

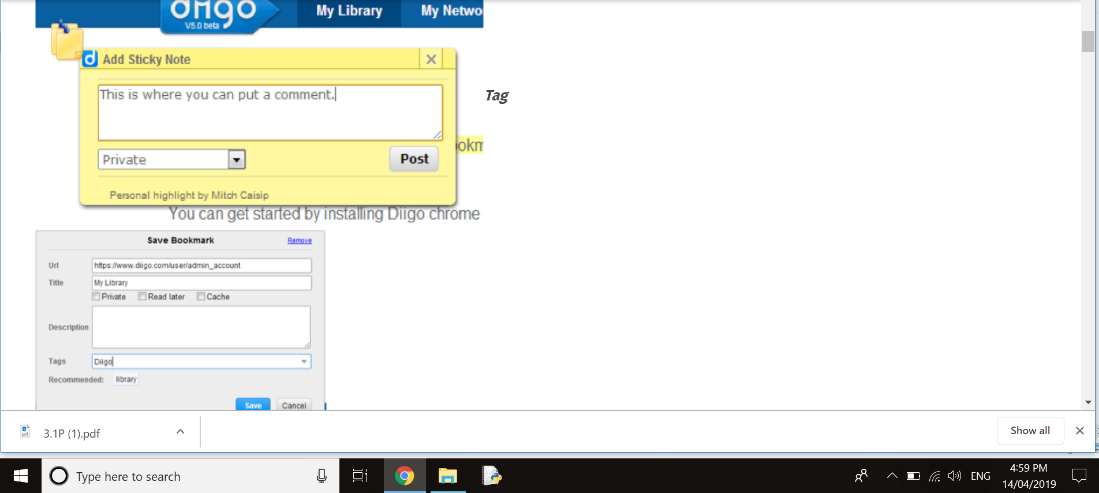
***Description of the tool***

Diigo is one of the most commonly used tools for web annotation. It allows you to bookmark and tag web pages. In addition, it allows the user be able to highlight a part of a webpage if you wish to emphasize something and attach sticky notes to it.



***Features of the tool***

One of the most significant features that Diigo has, is the advanced search feature. This feature allows the user to rummage through the text on pages which the user has bookmarked including, tags, titles and URL’s. The user can then identify a specific word or phrase using the highlighting feature. Another of the key features of Diigo, is the ability to comment on other users notes, which can either be set to public or private.



***Features that align with our aim***

Comparatively with our project, the general process of highlighting and annotating the text is similar. Although, in our project the aim to hyperlink the annotated text to provide assistant to the user.

***Features that would need to be altered***

As outlined above there are several features that align with our such as the general annotating and sifting parts of text for further examination, although we need to include additional features to the Diigo tool, as we require the extra step of hyperlinking words with definitions.

***Benefits of the Tool***

* Customization options in this tool and social feature help students learn key web skills like curating and sharing online content.
* Easy to install extension and start knowledge from around the internet it is a great capturing tool.
* The features it offers to collect articles and text snaps from the web and pdf are quite nice and can be really useful.

***Limitations of the Tool***

* In this the user need to be logged into Diigo and stay within one browser tab to use the system. And there isn’t much guidance on selecting sources.
* The premium versions are a bit pricy.
* As a free user you are given limited usage of the software including a limit on how many words you can highlight and bookmarks. If it is a tool that is used very often, it is worth spending.

***User Scenarios***

* The user needs to open the Diigo website first.
* Then can search for the annotated words.
* The user can highlight the words or paragraph in the web page.
* The user is able to comment on the other users notes which is set for public or private.

***Price***

* Diigo membership varies in three ranges standard, professional, business. These are applicable for the members in Diigo. The price varies according to the ranges.

*ZipBoard*

***Description of the tool***

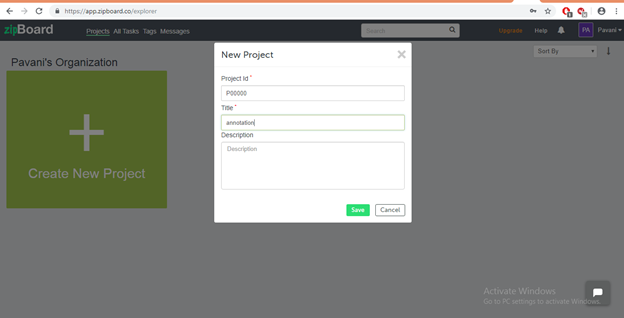
ZipBoard is a bug tracking tool that enables users to take screenshots, annotate and collaborate with web pages. Additionally, it has the capacity to assigns task while browsing users web-based products.



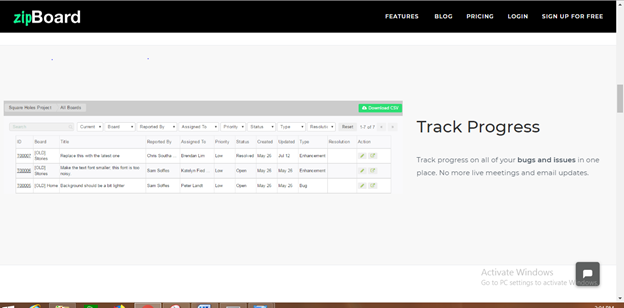
***Features of the tool***

Zipboard embodies many features including but not limited to, the main annotation and collaboration features. The main features are outlined below with corresponding images:

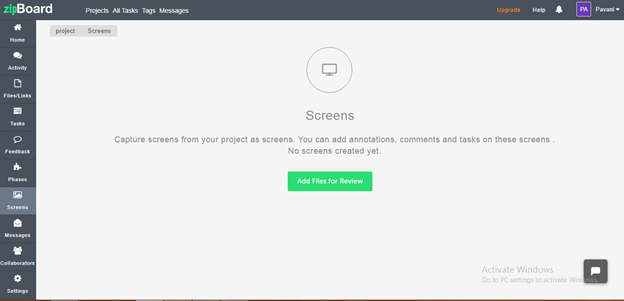
1. “Create new project” needs to create a project first for the annotation. Then we can annotate and track the progress.



1. The Annotation feature helps in tracking the progress to stay on top of the projects and complete tasks.



1. The Unlimited collaborations feature means there are no limits to the amount of people collaborating on the project.
2. Annotate the words in the screen, can comment and assigns tasks on these screens.



***Features that align with our aim***

There are similar annotation features of Zipboard that align with our aim. Our project highlights the words that are annotated and provides hyperlink, whereas Zipboard allows the user to annotate the words from the web pages in and can provide or derive their own meanings.

***Features that would need to be altered***

The basic annotation feature of Zipboard would need to be built upon additional features added to align with our aim. There would need to be the addition of hyperlinking to definitions, rather than the manual feature that Zipboard uses.

***Benefits of the Tool***

* Free web page and image annotation tool
* Helps in collaborative working
* Can annotate the words by ourselves.

***Limitations of the Tool***

* Only free signup for one project.
* Add on costs.
* No offline access

***User Scenarios***

* The user has to create account in ZipBoard.
* Then the user needs to create the project for annotation.
* Later the user can annotate the words or images.

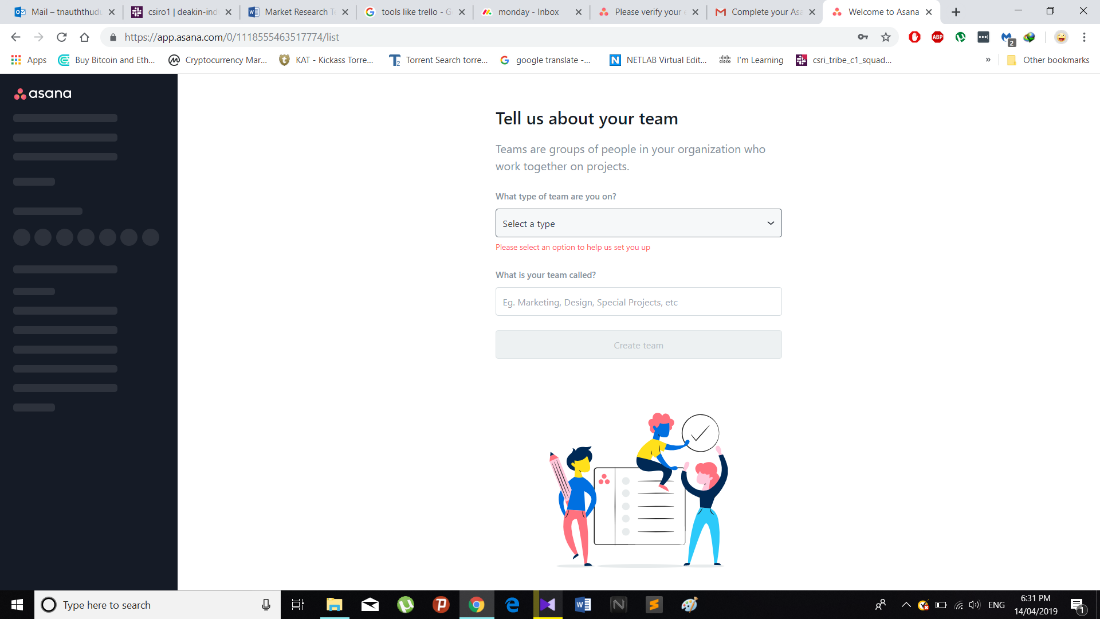
***Price***

The first project is free and then a monthly subscription is required for $39per month. The same concept applies for the team version, however the monthly is cost $79.

*Asana*

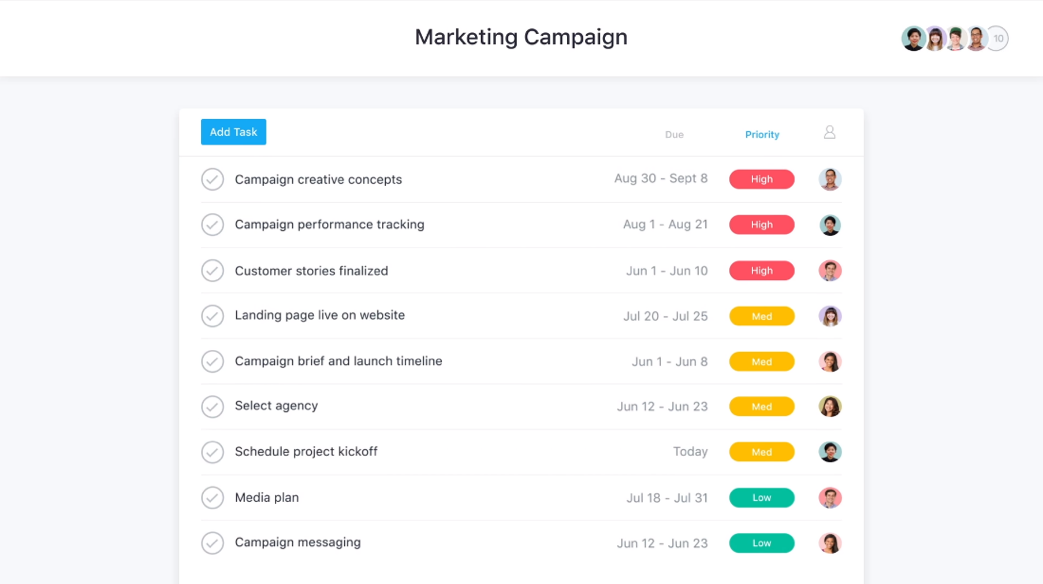
***Description of the tool***

Asana is tool that is predominantly used for application development and is designed for team communication during development.



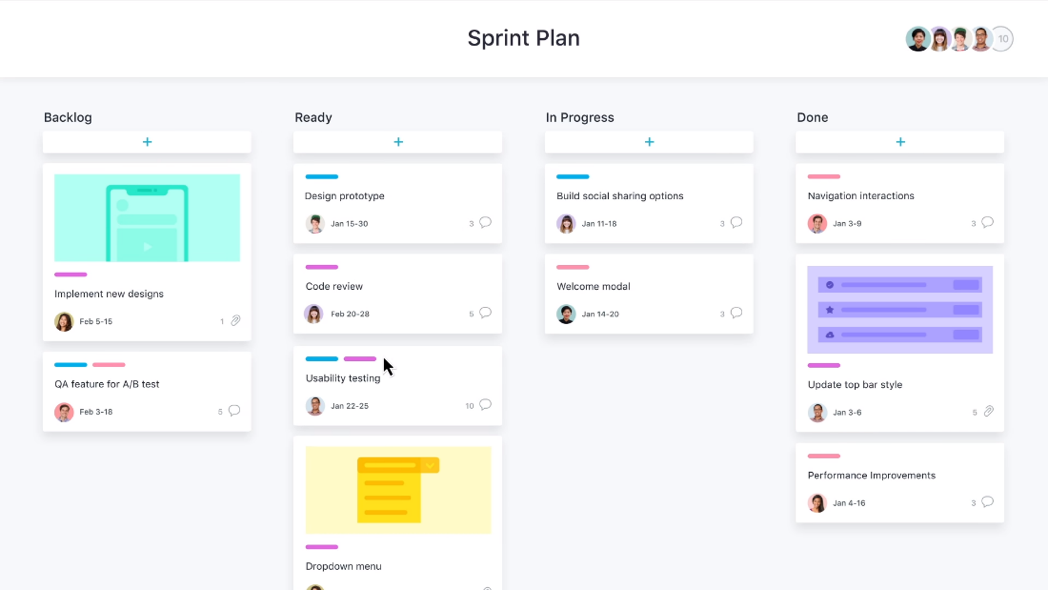
***Features of the tool***

Asana has many of features for developing tools. One of the features is that an each and every members of group can edit and add contribute to the code.



***Features that align with our aim***

One of the features that align with our aim is that is can easily divide work load . This tool is very helpful for dividing work load for each team members. Not only that, sprint plan is very helpful for creating subtask.



***Features that would need to be altered***

There are no any features that would need to be altered to align with our project aim and deliverables.

***Benefits of the Tool***

* The tool is excellent for group collaboration
* The tool has many features that assist development

***Limitations of the Tool***

* This tool has only one-month trail.
* User need to log in

***User Scenarios***

* User doesn’t need to sign in for using the tool
* Any member in the team can edit and code the project
* The user can easily divide the work load.

***Price***

The first month is free and after first month the price is $129 for the full pack.

FEASABILITY ANALYSIS

Based on initial understanding of the basic requirements of the tool that is to be developed and the research undertaken into other relevant tools, it appears that the development of analytical text tool is feasible. However, as shown above in the document there are a range of similar tools that meet many of the requirements of the tool, which could be used instead of committing to the cost of developing a new similar tool. There are however many benefits of developing a new tool from scratch in order to ensure that all the requirements are met and all the features align with the tools purpose. A cost/benefit analysis would be a good step in determining the certainty of a financial payoff, however simply based of the information available at this point in time, a new tool could be foundationally modelled off the most similar existing tool Bioportal. As discussed above Bioportal has all required features of the project, but is simply limited by the range of libraries available. It could be adjusted and developed to supersede the original and meet all requirements. This conclusion however is based solely off the information provided and focused independent research and therefore cannot take into consideration the availability and longevity of finances.

REFERNECES

**References for Bioportal**

Bioportal.bioontology.org. (2019). *Welcome to the NCBO BioPortal | NCBO BioPortal*. [online] Available at: https://bioportal.bioontology.org/ [Accessed 5 May 2019].

**References for LightTag**

Medium. (2019). *Announcing LightTag — The easy way to Annotate text*. [online] Available at: https://medium.com/@TalPerry/announcing-lighttag-the-easy-way-to-annotate-text-afb7493a49b8 [Accessed 6 May 2019].

Lighttag.io. (2019). *LightTag - The Text Annotation Tool for Teams*. [online] Available at: https://www.lighttag.io/ [Accessed 6 May 2019].

TechCrunch. (2019). *LightTag is a text annotation platform for data scientists creating AI training data – TechCrunch*. [online] Available at: https://techcrunch.com/2018/05/11/lighttag/ [Accessed 6 May 2019].

**References for TagTog**

"AI-enabled Text Annotation Tool | PDF, Markdown, CSV, html, tweets, & many more types of Documents", *Tagtog.net*, 2019. [Online]. Available: https://www.tagtog.net/. [Accessed: 08- Apr- 2019].

**Reference for Asana**

"Use Asana to manage your team's work, projects, & tasks online · Asana", *Asana*, 2019. [Online]. Available: https://asana.com/?noredirect. [Accessed: 09- Apr- 2019].

**Reference for monday.com**

"monday - team management software | monday.com", *monday.com*, 2019. [Online]. Available: https://monday.com/. [Accessed: 16- Apr- 2019].

**Reference for ZipBoard.com**

Free website annotation tool for web developers, designers | zipBoard. (2019). Available at: https://zipboard.co/?[Accessed15Apr.2019].