

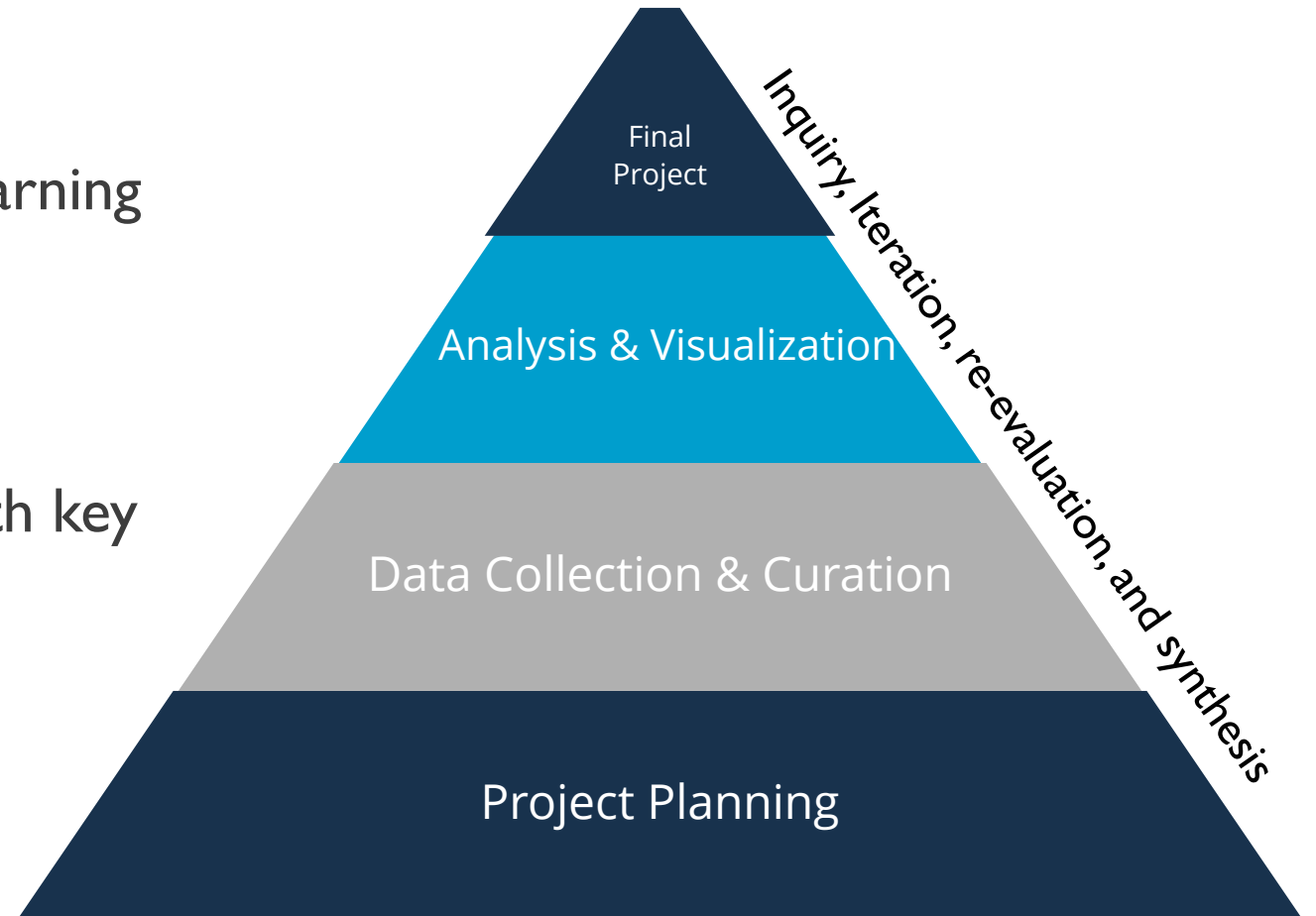


**MORE THAN JUST A
PAPER: TRANSFERRABLE 21ST CENTURY SKILLS
IN TEXT & DATA MINING RESEARCH**



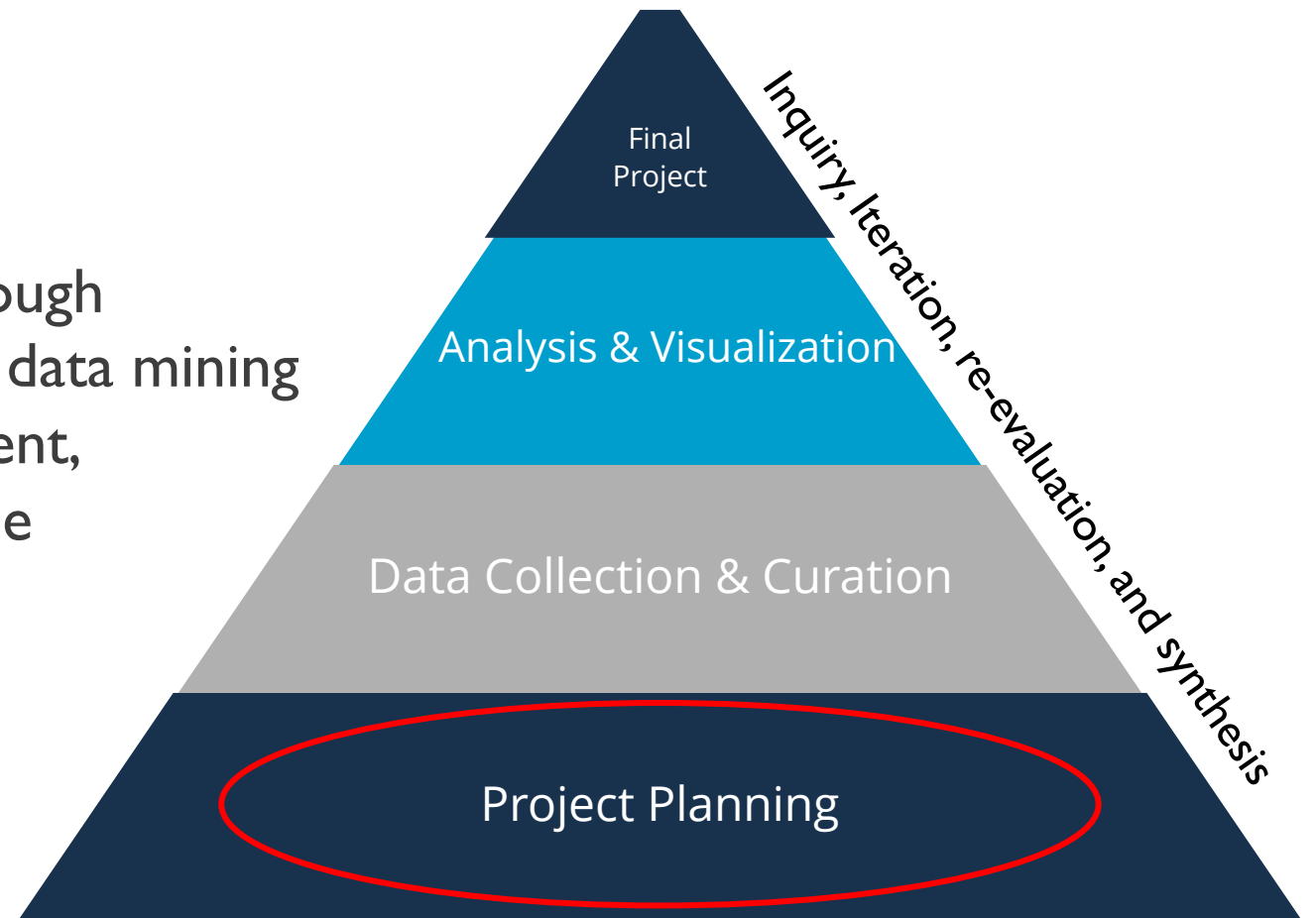
SITUATING TEXT AND DATA MINING FOR LEARNING

- The power of project-based learning
 - Digital Literacies
 - Higher-order thinking skills
 - Competencies alignment with key text and data mining facets



PROJECT MANAGEMENT SKILLS FOR RESEARCH SELF-REGULATION

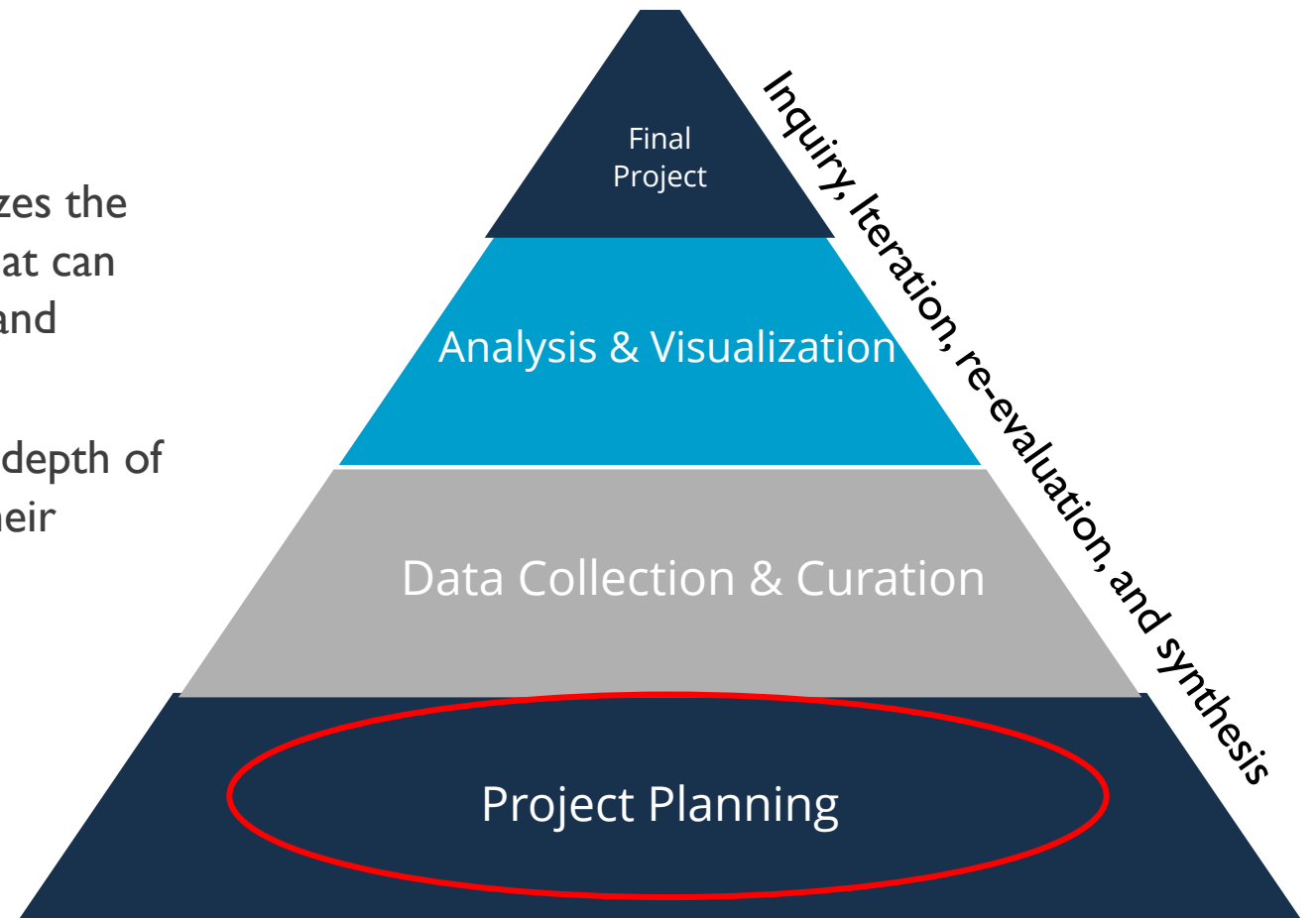
- Project Management Skills
 - Increases self-regulation through challenging rigor of text and data mining by promoting risk management, collaboration, scope, and time management.



IDEATION OF RESEARCH INQUIRY

The rigor of text and data mining research prioritizes the need for researchers to have a depth of interest that can drive how their curiosity. This impacts the inquiry and perseverance they participate in.

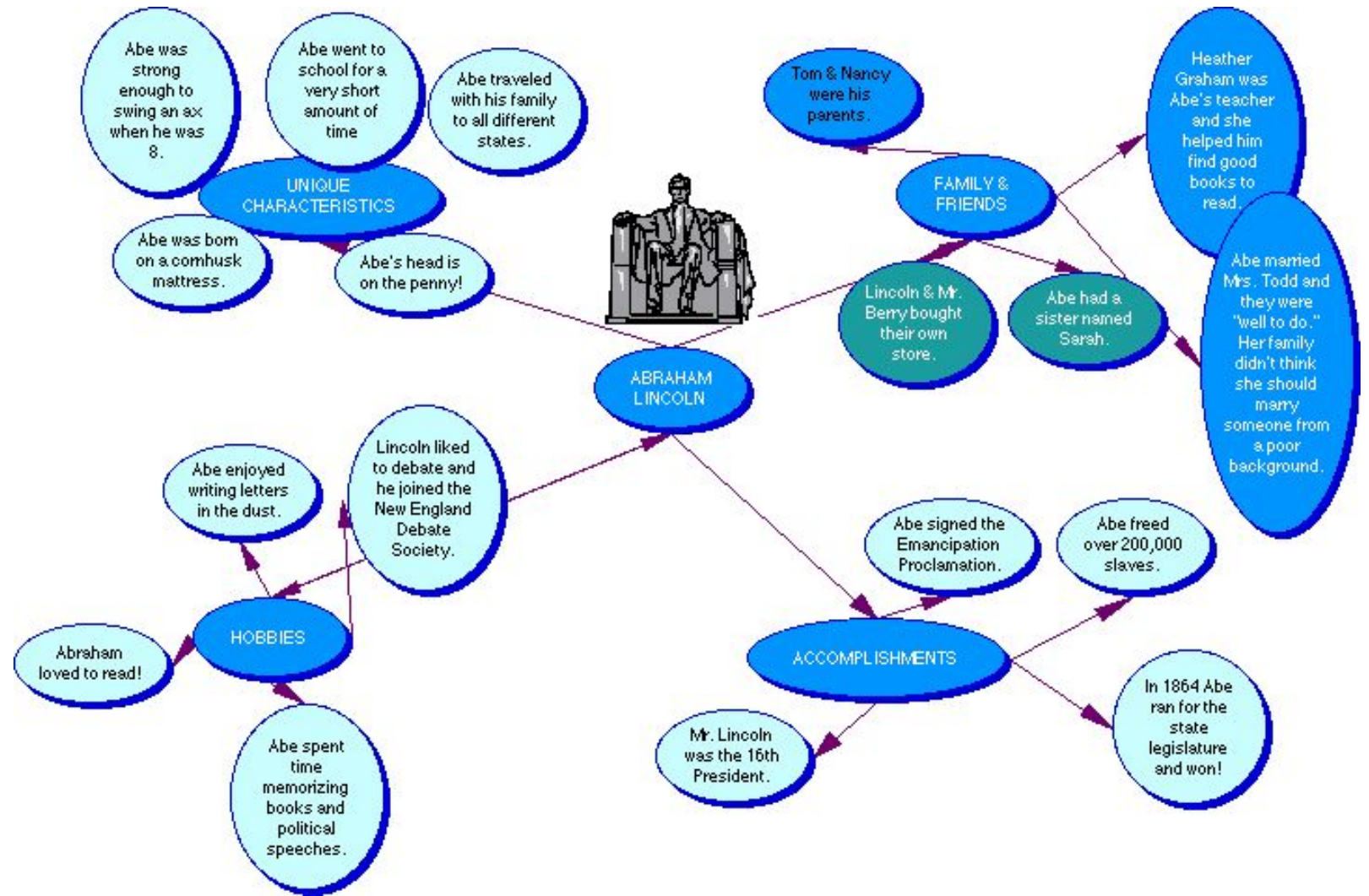
- Concept Mapping topic of interest to illustrate depth of research opportunities, but also the scope of their pursuit.



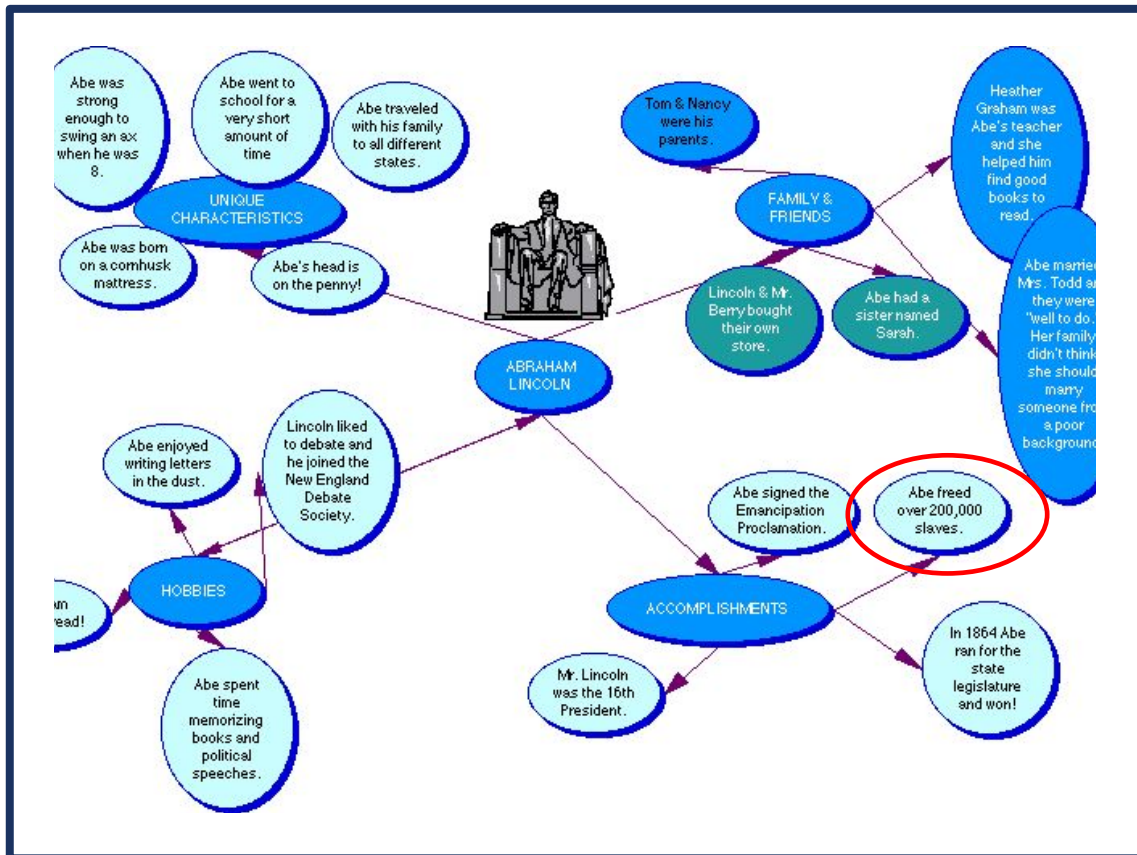
IDEATING ON TOPICS

■ Aligned literacies and critical thinking skills:

- Basic recall to drawing connections among ideas
- Visual Literacy
- Close to Distant Reading



DEVELOPING RESEARCH QUESTIONS

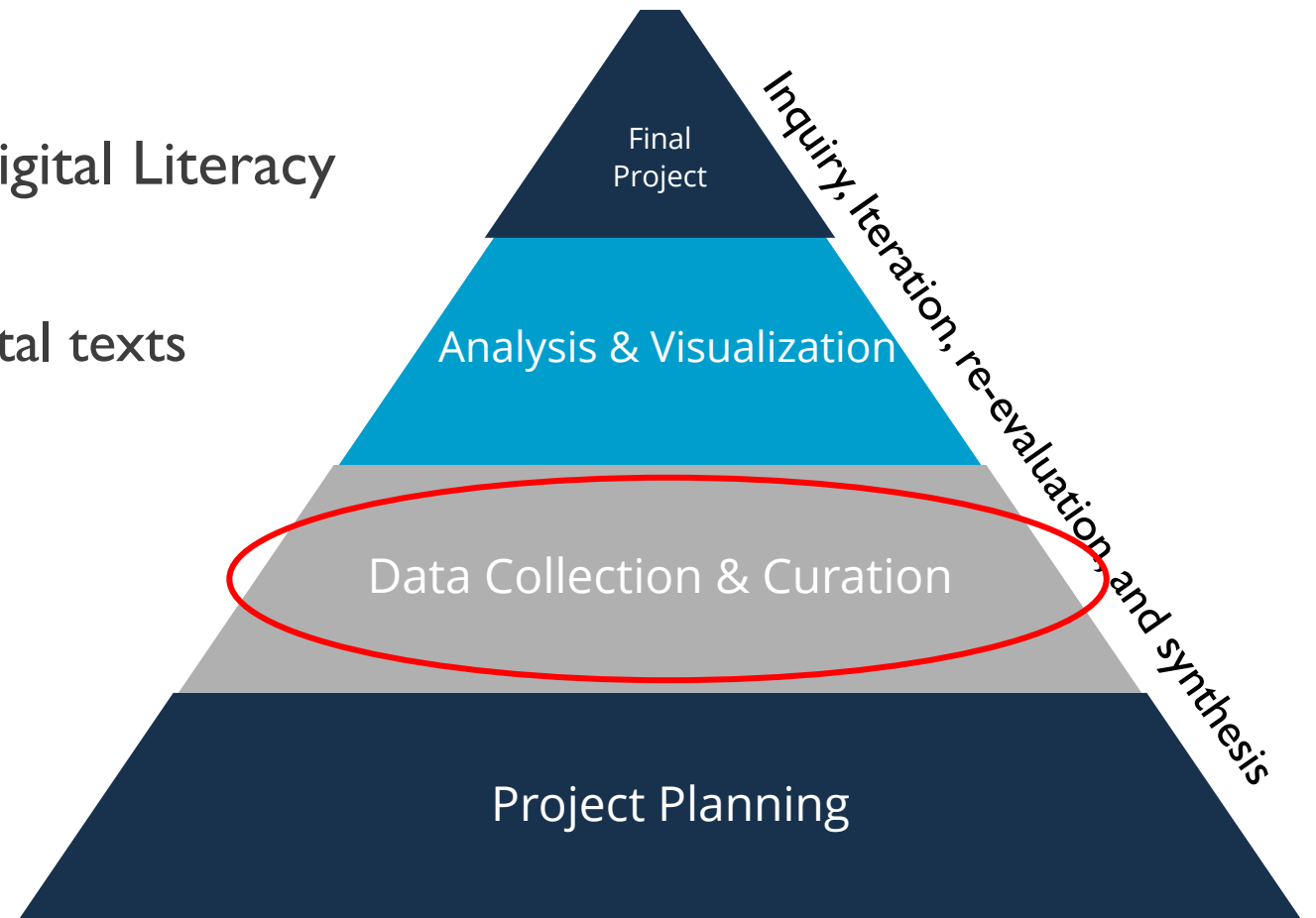


Asking the Right Questions

- Going from Topic to Questions
 - Start with the Who
 - Abraham Lincoln
 - Why
 - Why did Lincoln find it important to free slaves?
 - How
 - How did Lincoln allude to the freeing of slaves *in his speeches*?
 - This type of question addresses tone, motivation, themes in the texts of his speeches.
 - In what span of time/critical period/event
 - Textual evidence from 1809-1865?
 - Final Question: **How did Lincoln allude to freeing of slaves in his speeches between 1809-1865 and why did he find it important to free slaves?**

PERSISTENCE IN DATA COLLECTION AND CURATION

- Persistence, Iteration and Digital Literacy Skills.
 - Understanding OCR of digital texts
 - Effective Searching
 - Creating Content Sets
 - Cleaning Texts



CRITICALLY THINKING ABOUT METHODOLOGY

When selecting a tool...

HOW do I wish to answer my research question?

Am I looking for the quantity of vocabulary words?

Am I looking for **how** words are used?

Am I looking for themes across my content set?

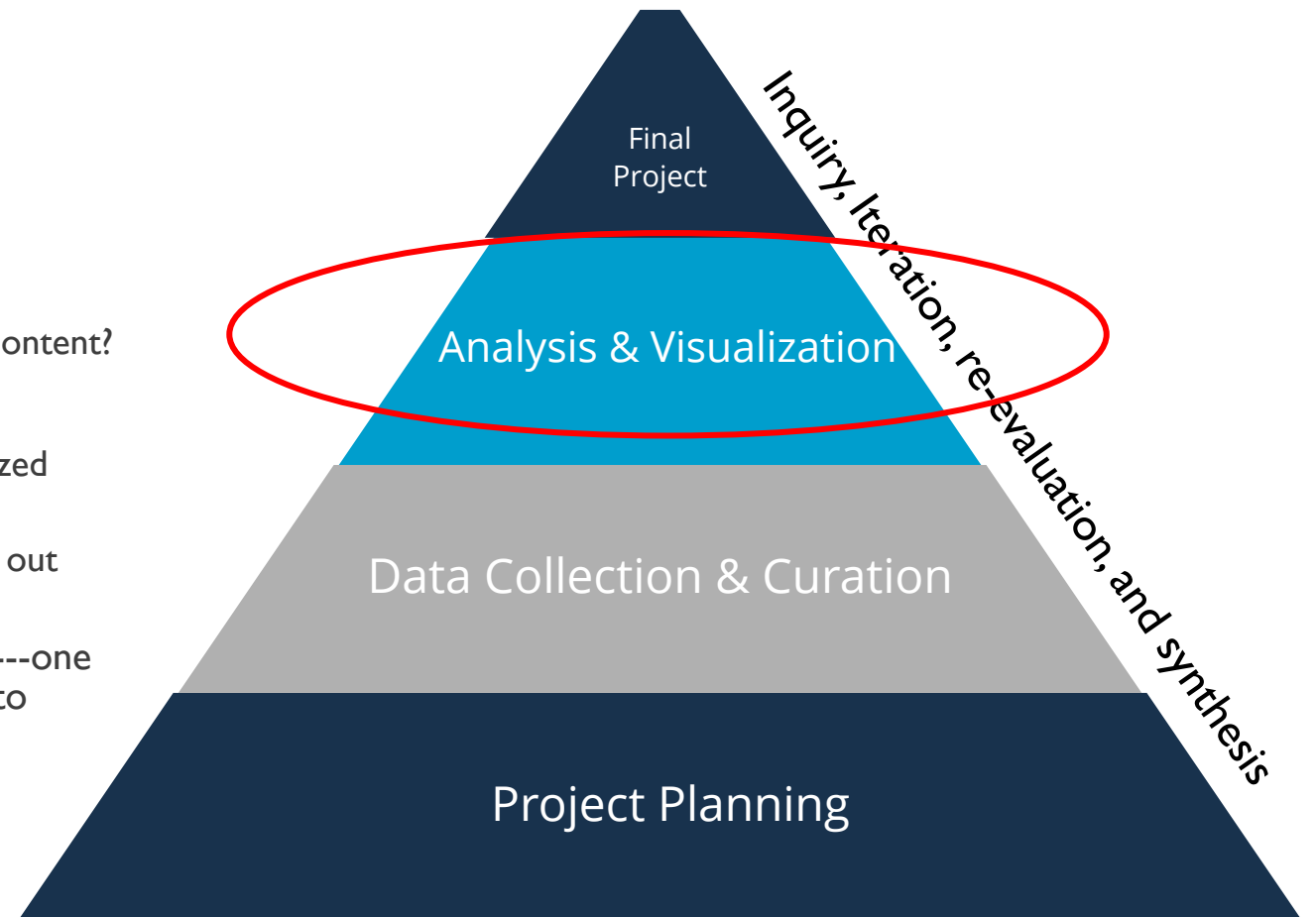
Am I looking for the tone (i.e., sentiment) relevant to my content?

Does my content require a level of cleaning?

Will the type of content I have curated be effectively analyzed with these selected tools?

How will the metadata affect my ability to effectively carry out the analysis?

Set yourself up for success: Get into the iteration mindset---one round of analysis with any given tool is rare, and less likely to produce the results you're looking for.



QUESTIONS FOR INTERPRETATION OF RESULTS

When evaluating results...

HOW is the content revealing itself as output?

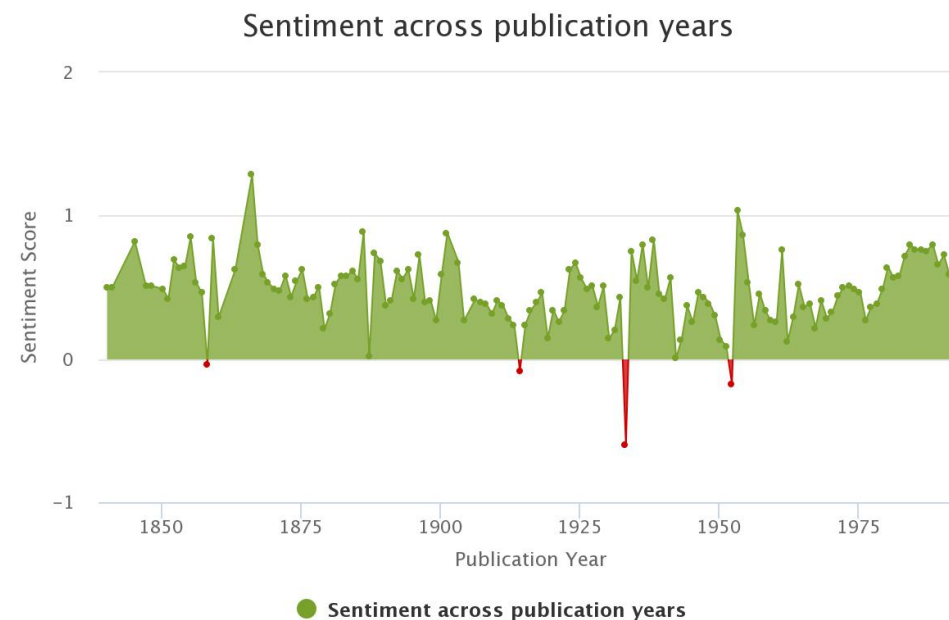
Why do the results appear the way they do?

Given what I already know about the research topic, what makes sense or does not make sense?

Do I see any outliers and odd characters/words that are impacting interpretation or the visualization?

Do I need to iterate in my cleaning configurations to reach a clearer set of results?

Should I use close reading of some of the content in order to understand the vocabulary, sentiment, or themes?



LEARNING MILESTONE: DEVELOP RESEARCH QUESTIONS

Activity

Create a concept map for your topic. Be sure to branch out at least three layers.

Select a concept on your outer most layer and perform some preliminary research to determine if the level of specificity of this area is actionable for research.

Then, do the “Who, Why, How, When” activity to develop your research question using that preliminary research and reflection.

Evaluate your research question and write down what you anticipate will be needed for cleaning, tools, and close reading to plan and scope your project effectively.