

# Measuring Observable Influence & Impact of Scientific Research Beyond Academia

SFU

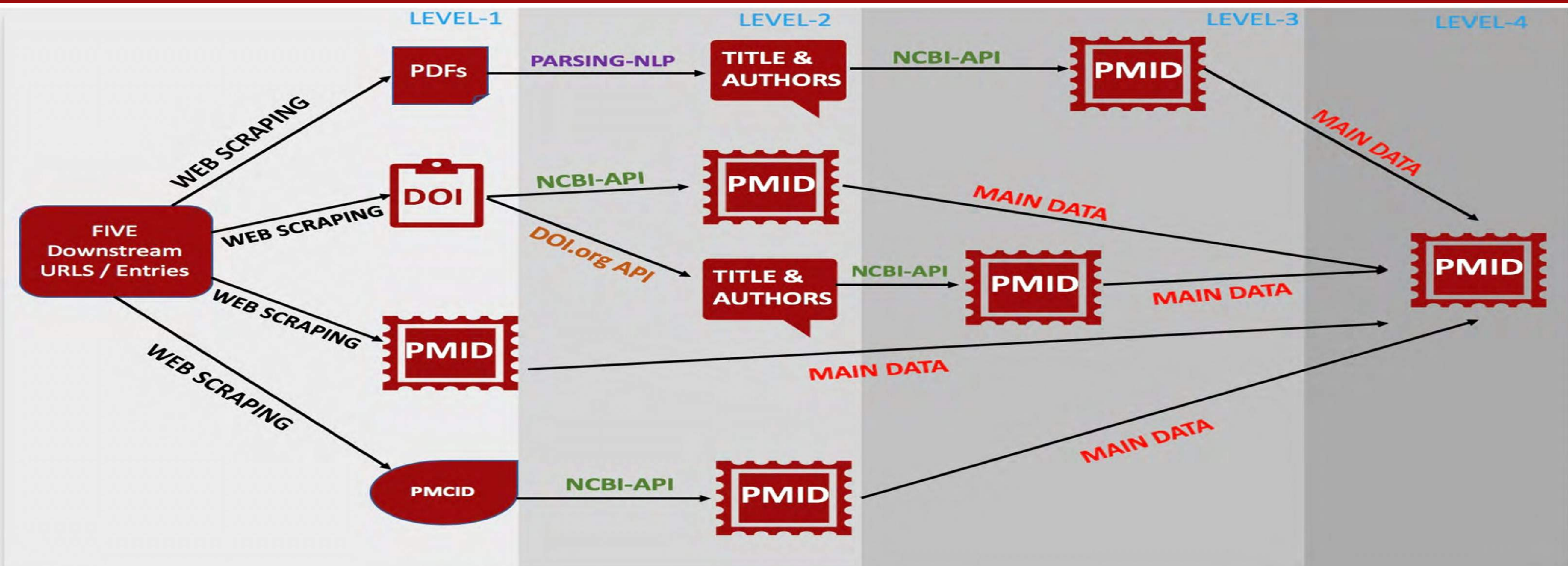
Team: abraca-data- Honghui Wang | Shray Khanna | Chhavi Verma

## MOTIVATION

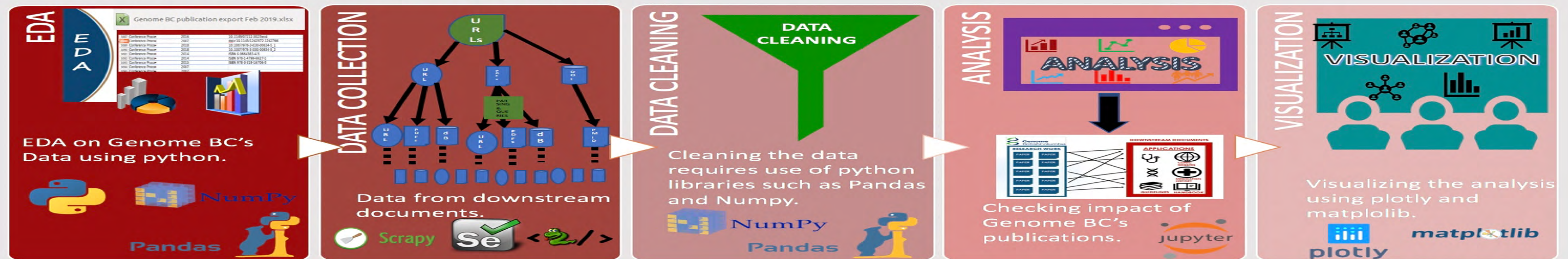
- > Checking impact of research publication in Healthcare.
- > Make healthcare research more impactful and beneficial to people.
- > Help researchers understand Real World impacts of their publication in Healthcare.

## CHALLENGES

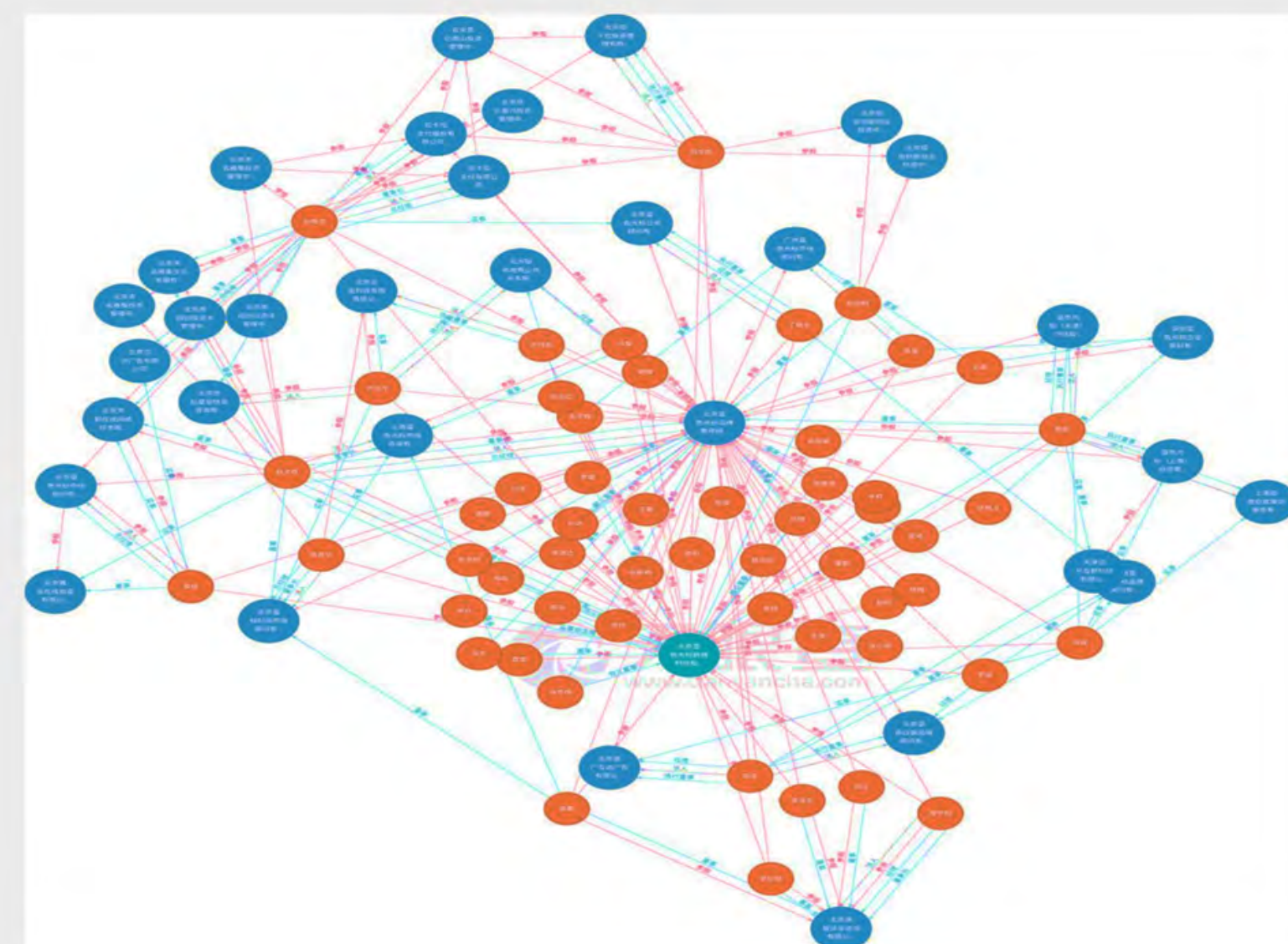
- > Data Collection is the biggest challenge we face.
- > References from downstream data is collected after going through each level.



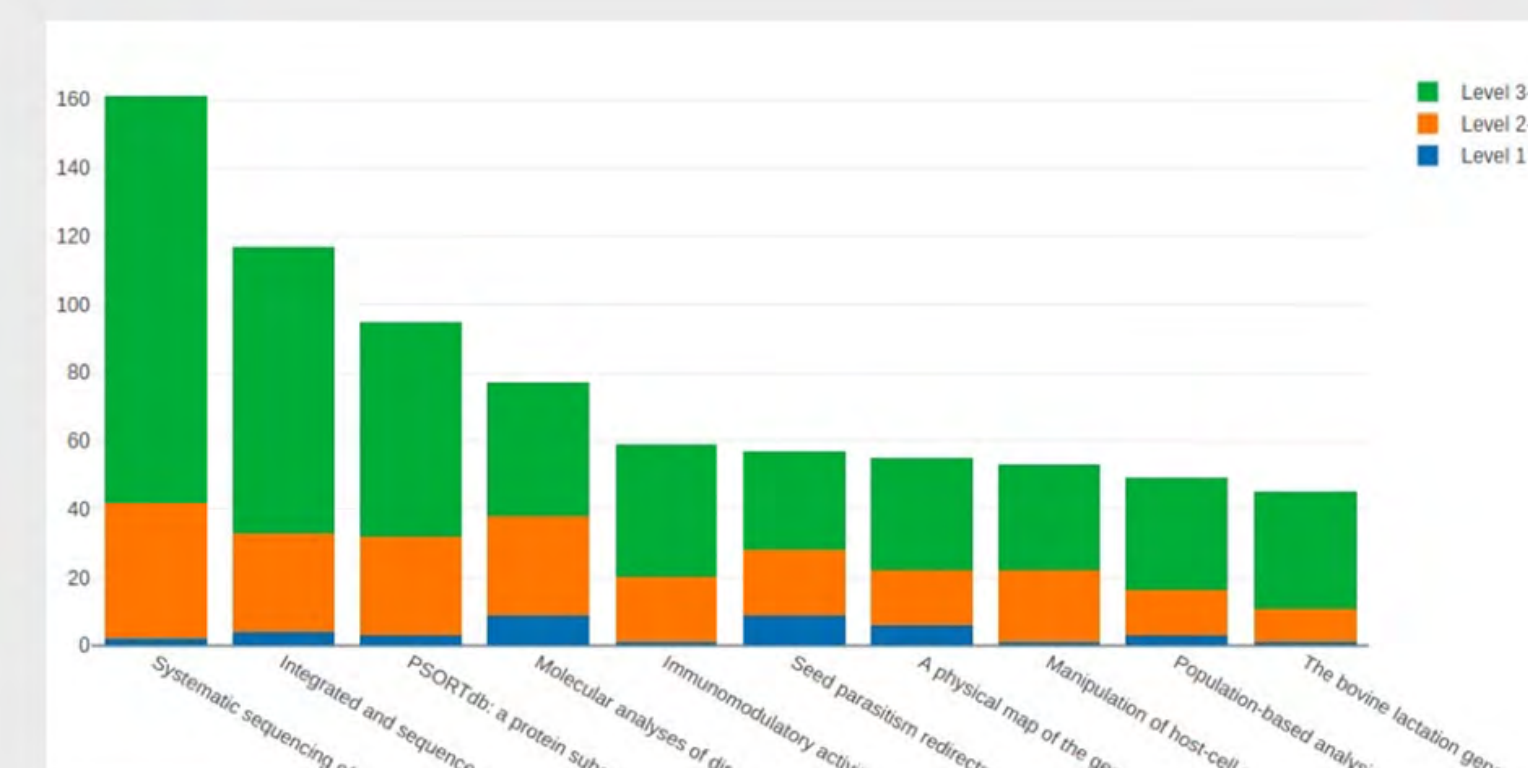
## DATA SCIENCE PIPELINE



## RESULTS AND VISUALIZATION



Left Picture :-  
The red circles denote the publications of Genome BC  
While the blue circles denote the entities from the five downstream documents we have.  
The more the connections of a Genome BC publication, the more impactful the publication is.



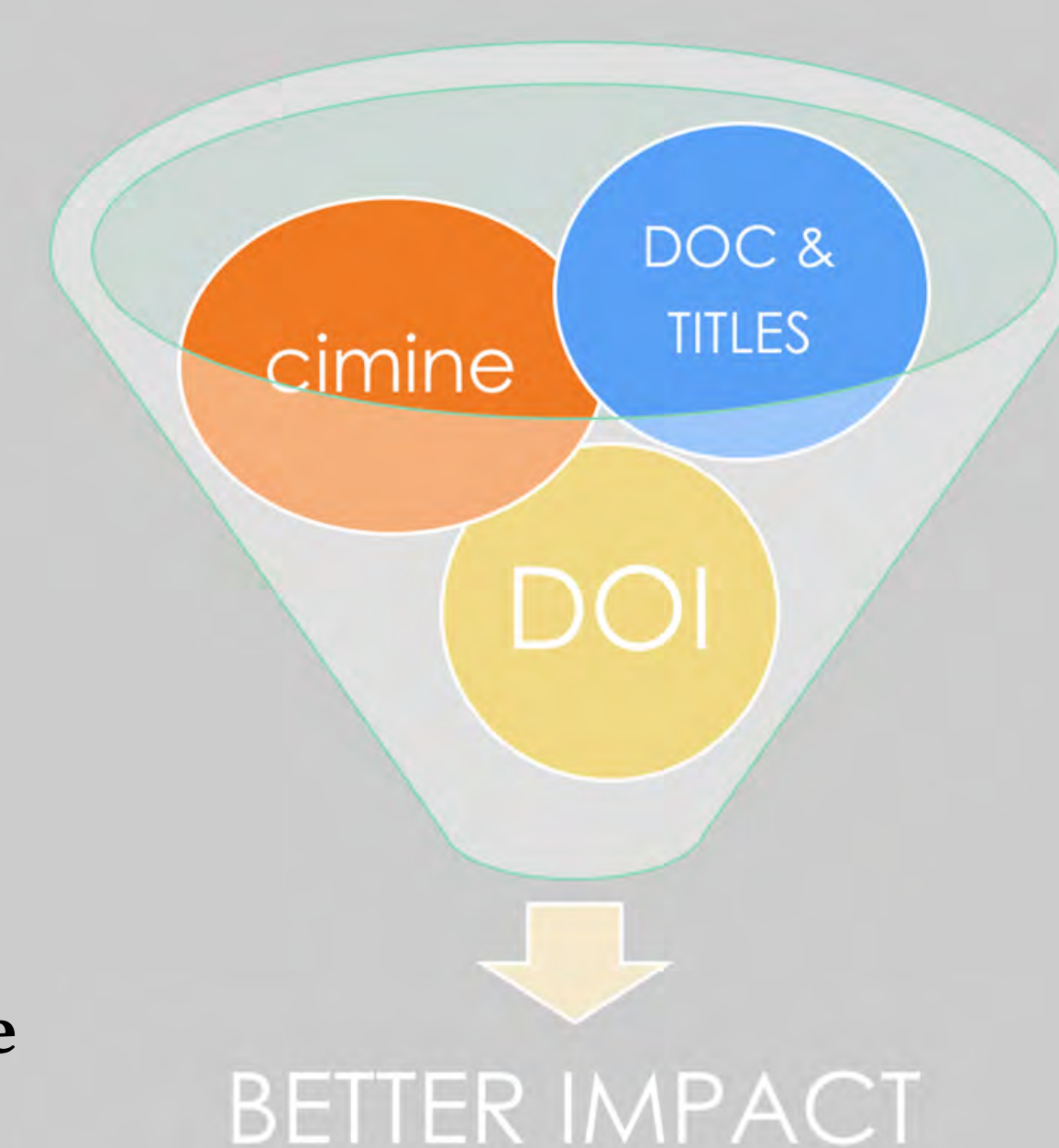
Right Picture :-  
We are observing the depth of a particular publication of Genome BC  
The y axis shows the total count  
A directly referenced paper will have more green colour

## FUTURE WORK

Use more clues, like DOI, Titles, pptx and docx files, to make connection between downstream documents to academic papers. This requires more time and labor as every clue needs a specific rule to extract the info.

Design an interactive framework allowing user to design the rules for the situation that our parser can not figure out.

Use commercial tools like Cernine to parse the pdf file more precisely. This needs financial support from Genome BC.



## ACKNOWLEDGEMENT

We thank Genome BC for being our mentor throughout the project. Thank you to Dr. Jiannan Wang and Dr. Steven Bergner for their support throughout the Programming for Big Data Course.

## REFERENCES

1. Ravenscroft J, Liakata M, Clare A, Duma D. Measuring scientific impact beyond academia: An assessment of existing impact metrics and proposed improvements. PLoS one. 2017 Mar 9;12(3):e0173152.
2. Chenet M. Identify and extract entities from bibliography references in a free text (Master's thesis, University of Twente).
3. Connan J, Omlin CW. Bibliography extraction with hidden markov models.
4. Tkaczyk D, Szostek P, Fedoryszak M, Dendek PJ, Bolikowski L. CERMINE: automatic extraction of structured metadata from scientific literature. International Journal on Document Analysis and Recognition (IJ DAR). 2015 Dec 1;18(4):317-35.