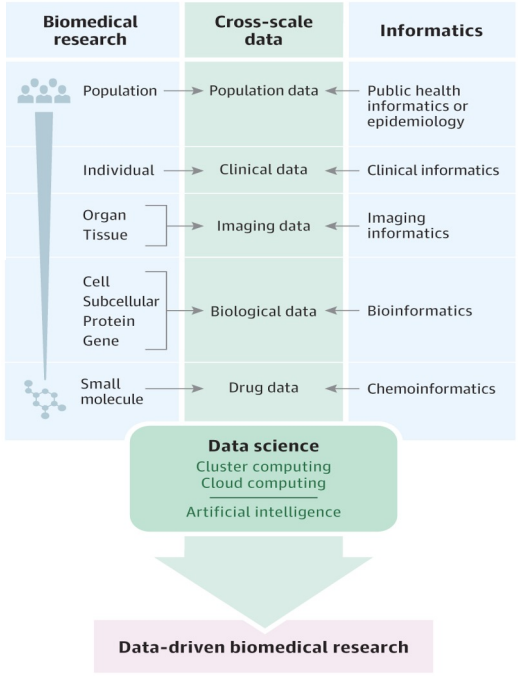
W. Jim Zheng, Ph.D.
Professor
Founding Director, DSICCR
Department of Bioinformatics & Systems Medicine
McWilliams School of Biomedical Informatics
UT Health Science Center at Houston
wenjin.j.zheng@uth.tmc.edu

Lisha Zhu, W. Jim Zheng: Informatics, Data Science and Artificial Intelligence, **JAMA**, 320(11):1103-1104, 2018



Visualization of key data or a resource

Research Interests
Methodologies: Bioinformatics, Data Science, Literature mining, Knowledge representation, Artificial intelligence
Diseases: Cancer (CPRIT RP170668), Brain/Brain Disease (NIH 1U24MH130988-01, NIA 1R01AG066749-01, NIA 5R56AG069880-02), Autoimmune Diseases (DoD W81XWH-22-1-0164)

Strengths or Unique Resources
CPRIT Data Science & Informatics Core for Cancer Research (**DSICCR**) is a CPRIT funded core facility that provides support and collaboration opportunities in bioinformatics, data science, high performance computing and artificial intelligence (see slide #2).

Type of collaborator you seek
All types of expertise in single cell biology who share similar research interest (see slide #3).

Publication List (link & qr code):
https://www.ncbi.nlm.nih.gov/myncbi/18EP7Fevc6fkn/bibliography/public/?sortby=pubDate&sdirection=descending

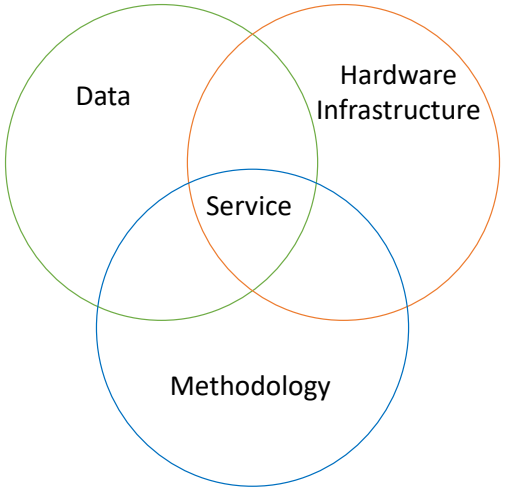
Lab or Faculty website (link & qr code):
https://sbmi.uth.edu/faculty-and-staff/jim-zheng.htm

LinkedIn (link & qr code): https://www.linkedin.com/in/w-jim-zheng-6260b01/

DSICCR – AI and Big Data Research

Four Components of DSICCR

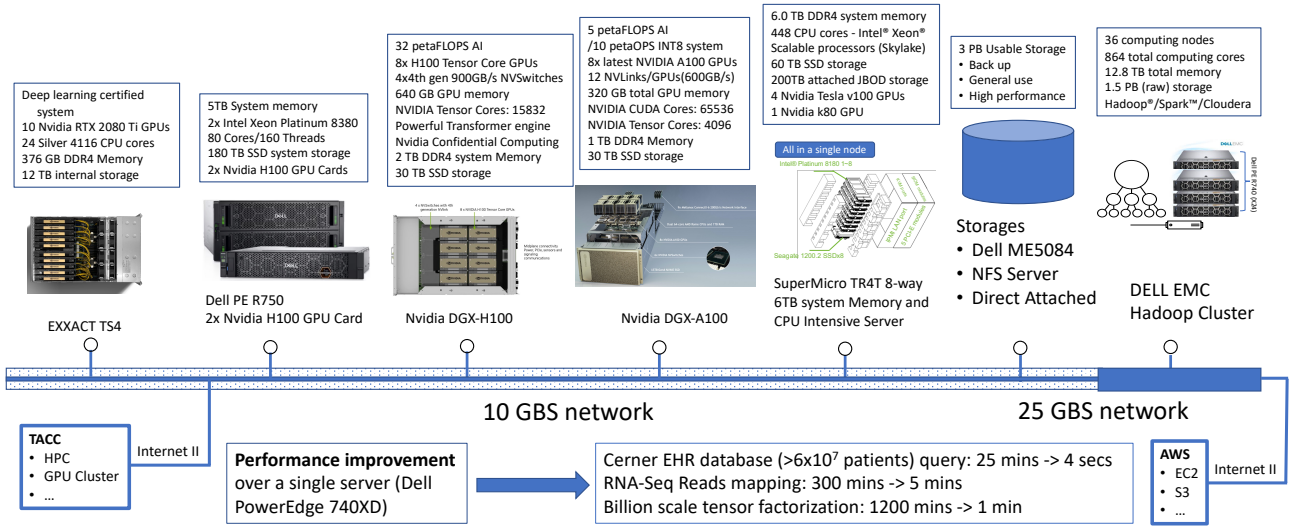
Data
- Data acquisition
- Data management
- Data discovery
- Data retrieval
- Data analysis

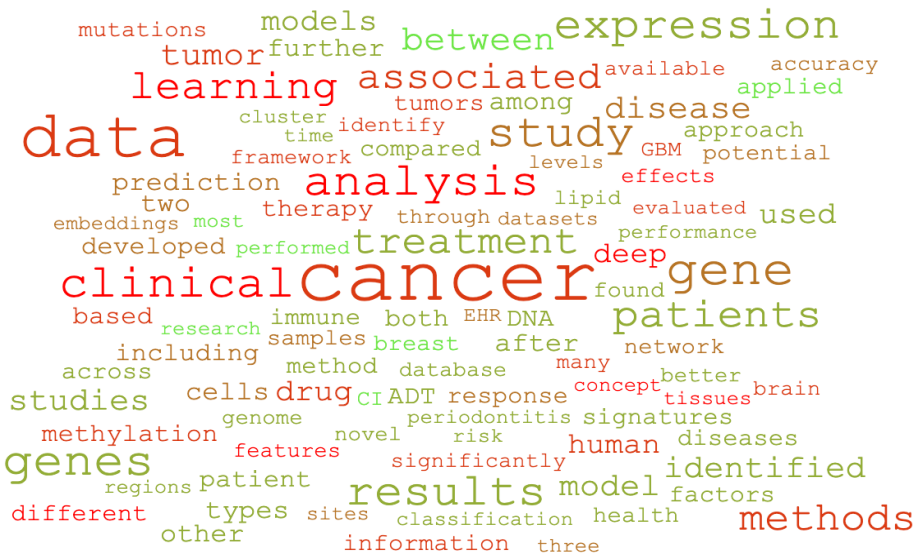


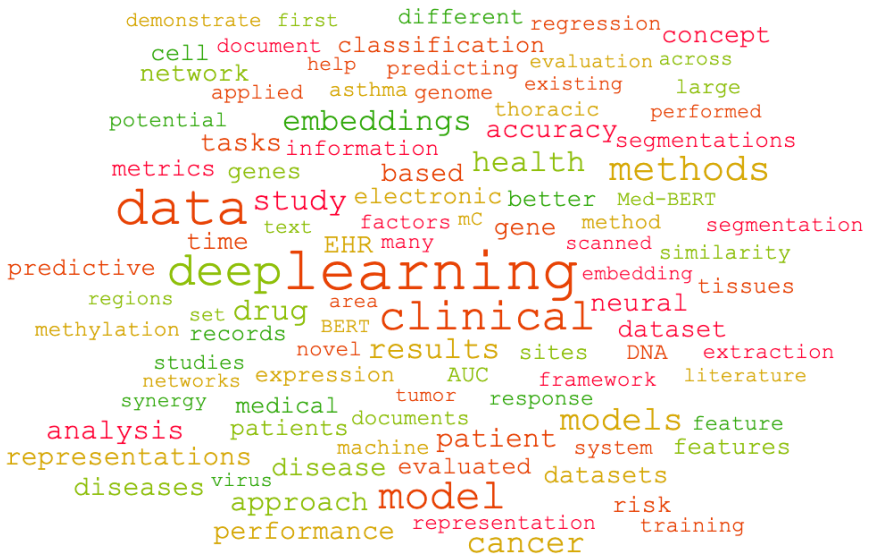
Hardware
- Versatile
- Diverse
- High performance
- High capacity
- High availability

Methodology
- Robust
- Cutting edge
- Diverse
- Cross-scale

Service
- Analysis service
- Collaborative research
- Grant development
- Online accessible tools





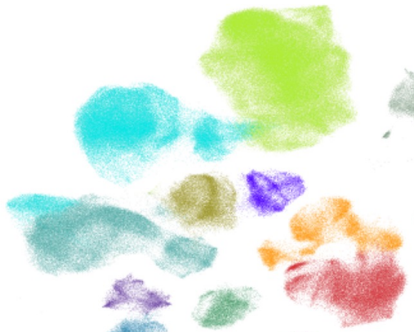


- **$1.5 million** cutting-edge computing hardware.
- **130 publications**
- Supported obtaining over **$60 million** grant awards.

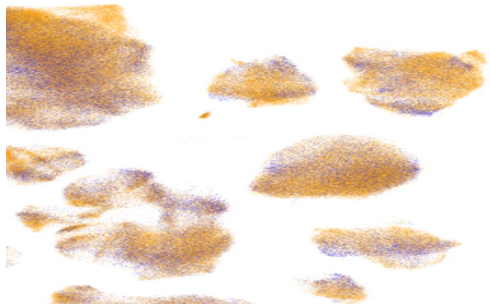
Research Interest from Data Science Perspective:
How to integrate scRNA-Seq data to advance disease research?

Challenges:

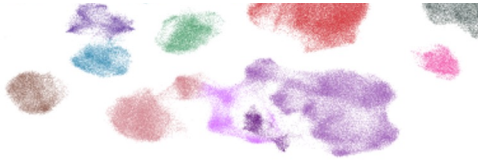
1) Cell cluster annotation
2) Cell type definition
3) The role of literature and other biomedical knowledge
4) Annotating cell clusters and cell types by ontology and standardized terminology
5) Cross data set analysis, data integration
6) Data findability, accessibility, interoperability and reusability (FAIR)
7) Application in research domains: cancer, brain/brain diseases, autoimmune diseases



scRNA-Seq results from lab #1



scRNA-Seq results from lab #2



scRNA-Seq results from lab #3

Figure 1. Data from different groups are often discrete and not integrated. There are HCA, BICAN and many other efforts, but there are new and unique opportunities and solutions.

The purpose of this introduction: Inviting experts with strong track records in single cell biology and similar research interest to develop innovative solutions.