

UC San Diego
SCHOOL OF MEDICINE

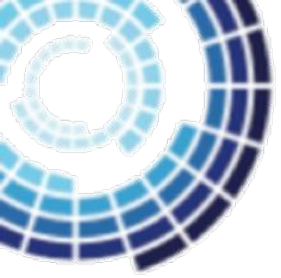
**Department of
BioMedical Informatics**

Scott Phillips

Dr. Olivier Harismendy

Annotation and Reporting of Cancer Bio-Repository Data





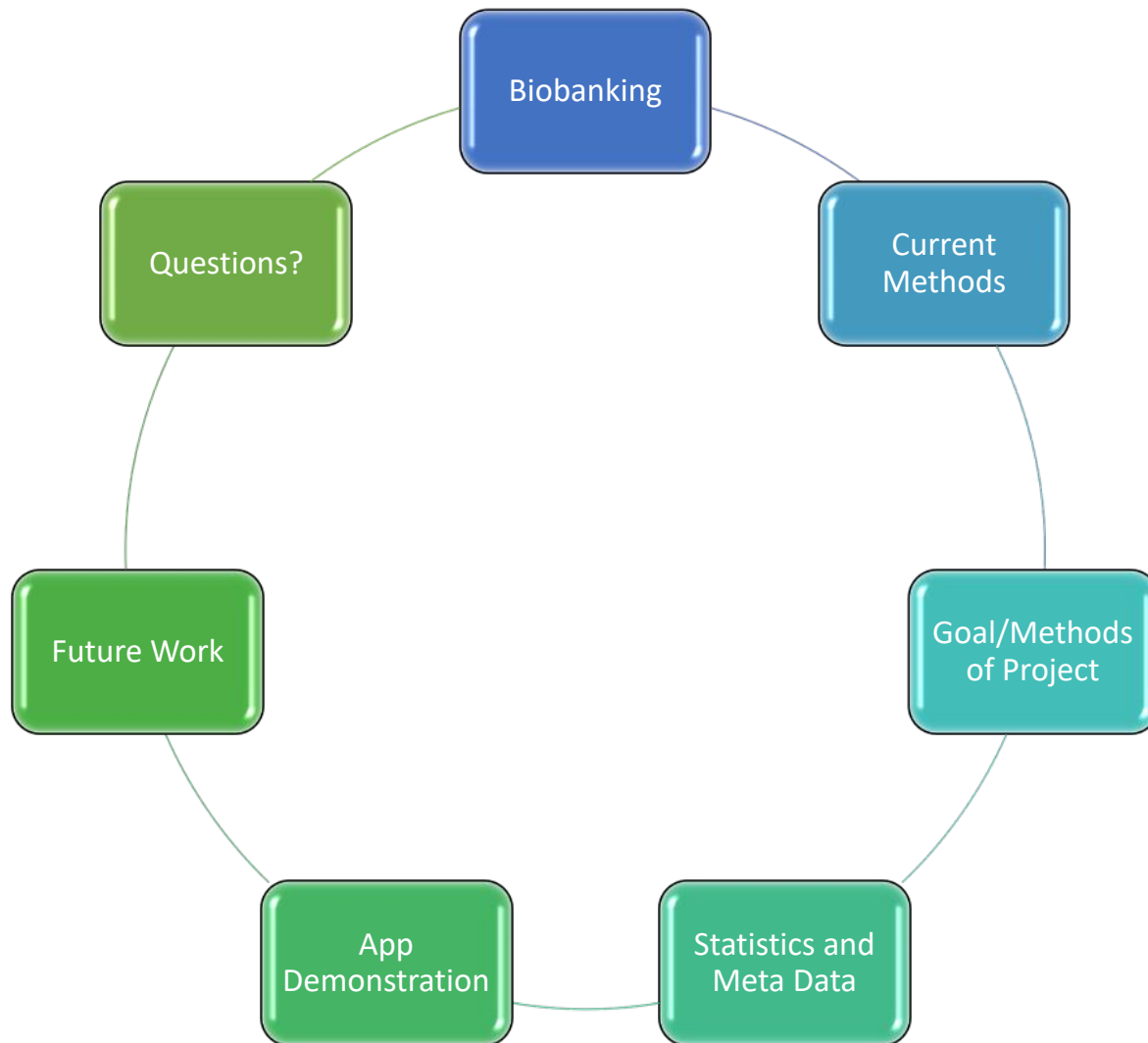
About Me

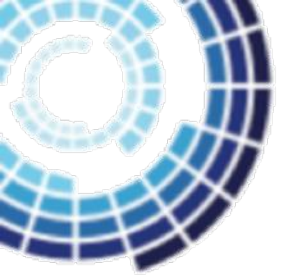
- 4th year Human Biology major
 - Mechanical Engineering minor
- UCSD Fencing
- Graduate School/Med School
- Interested in medical engineering
- Design and send new biotech products to market





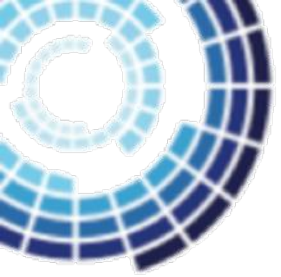
Outline



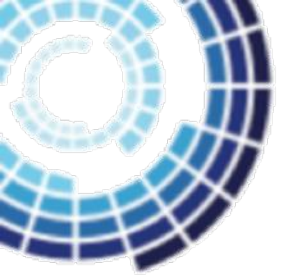


Biobank

- Biorepository storing physical biological samples for research use
- Provides support for translational research
 - Identify unique biomarkers of disease
 - Develop personalized treatment plans
- Ethical questions
 - Privacy concerns for individuals
 - Extent of consent for research usage

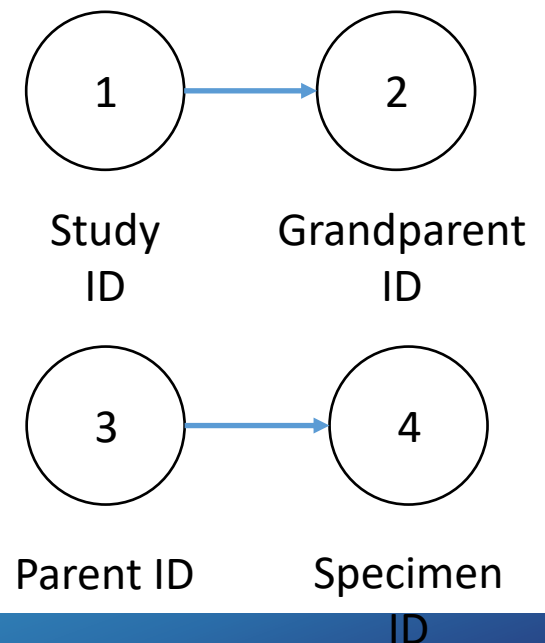
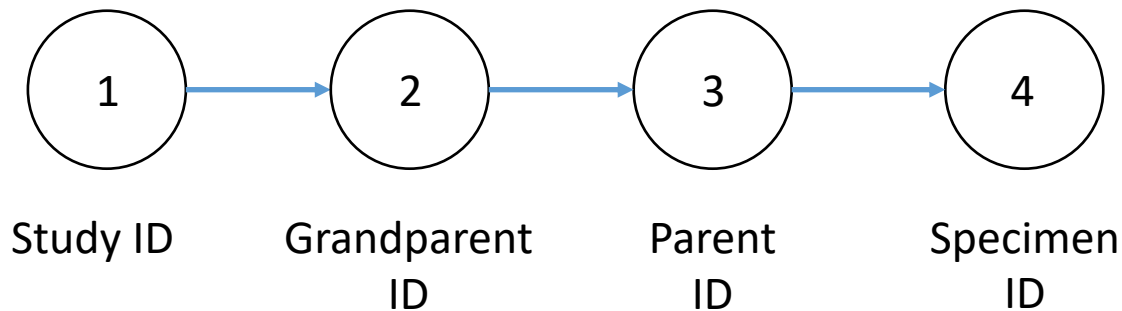


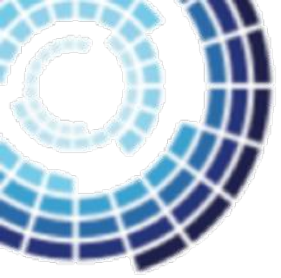
- **Moore's Cancer Center Biorepository and Tissue Technology Shared Resource**
 - Collects samples from cancer patients at Moore's Cancer Center
 - Stores samples for further analysis and research
 - Distributes specimens around the world
 - Uses VELOS data entry/management system
- **Director: Dr. Alfredo Molinolo**
 - Deputy Director: Dr. Jason Sicklick
 - Assistant Director: Dr. Sharmeela Kaushal



VELOS Data Entry System

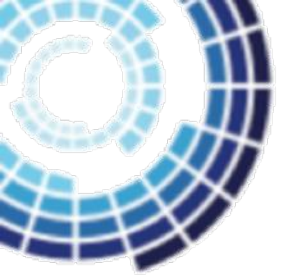
- eSample module within VELOS
 - Not optimized for data extraction or querying
 - Need to aggregate data from multiple sources
 - Technicians' personal data entry preferences
 - Ambiguous/duplicate information
 - Need to de-identify data
 - Unclear data hierarchies





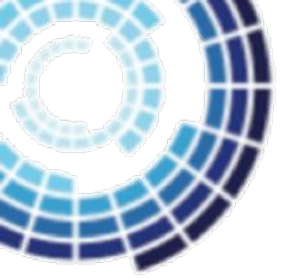
Goal of Project

- Biobank staff need to efficiently identify samples to facilitate research
 - Aggregate, reorganize, simplify data
 - Enable ease of extraction and analysis
 - Specimen identification
 - Metadata about patients
 - Easily understood data hierarchy
 - Large-scale view of entire repository data
- Allow BTTSR staff to intuitively query database



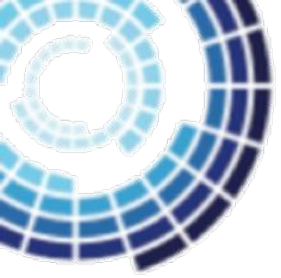
Method of Project

- R
 - Free programming language designed for statistical computing
 - Designed for managing and displaying large data sets
 - Large number of add-on packages for wide variety of applications
- Shiny
 - Add-on library to R
 - Enables user to build web apps through R
 - Javascript, HTML, CSS support
 - Users can host on personal servers or Shiny hosting servers



Statistics and Meta Data

- 246645 samples
- 5490 unique patients
- 6311 unique patient and disease history combos
- 7 unique sample types: Buffy Coat, PBMC, Plasma, Saliva, Serum, Tissue



App Demonstration

www.BANDICAM.COM

Not secure | 35.165.49.70/?view=shiny

http://35.165.49.70/p/0328580d/ | Open in Browser | Publish

BTTSR De-ID'ed Data

Patient General Data | Patient Specimen Data | Specimen Data Table | Plots | Summary Data

Show 10 entries Search:

	studyID	Disease.History	Gender	Race	Ethnicity	Assays
1	AA5734	Cervical Cancer History	Female	White	Non-Hispanic	0
2	AA5774		Female	White	Non-Hispanic	1
3	AA5910		Female	White	Hispanic	1
4	AA5808	Abdominal Cancer History	Male	White	Non-Hispanic	0
5	AA2846	Breast Cancer History	Female	White	Non-Hispanic	0
6	AA5972	Liver Cancer History	Female	Native Hawaiian or Other Pacific Islander	Non-Hispanic	0
7	AA3837	Lung Cancer History	Male	White	Non-Hispanic	1
8	AA5473	Lung Cancer History	Male	White	Non-Hispanic	0

Study ID:

Disease.History:

Collection Date: 2006-11-03 to 2019-03-14

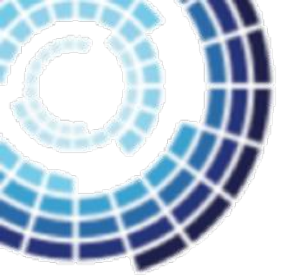
Gender: Male Female

Race:

Ethnicity:

Number of Assays:

Windows Taskbar: 3:21 PM 8/22/2019



Future Work

- Replace VELOS
 - Data extraction and analysis not supported
 - Prone to human error in entry
- Clean and organize data on back-end
 - Write programs to automatically process and sort data
 - Live connection to VELOS database
 - Integrate more data sources into app
- Upload app to Shiny server
 - Enable access for entire medical center

UC San Diego

SCHOOL OF MEDICINE

Department of BioMedical Informatics

Acknowledgements

- Olivier Harismendy
- Dr. Jejo Koola
- Dr. Tim Kuo
- Elizabeth Santillanez
- Nancy Herbst



Questions?