



## KNOWLEDGE MINING

Rancho BioSciences Ph.D. and M.D. scientists excel at knowledge mining that includes:

- Searching for datasets with specific attributes
- Target profile reports
- Building Natural Language Processing (NLP) pipelines and creating training datasets for NLP
- Building disease, genetic variant, microbiome, and drug knowledge bases
- Annotation of publications: interactions, assays, drug activity and toxicity, clinical trials, etc
- Literature search based on customer needs
- Search of Principal Investigators and grants for a research area of interest

## LEVERAGE THE FULL POWER OF YOUR DATA

Whether your focus is drug discovery, diagnostic tests, or foundational research, your company's value hinges on your ability to effectively leverage all available data. Too often, data is isolated by virtue of living on different systems, existing in non-standard formats, or lacking necessary curation. Fixing this problem requires a detailed understanding of both the data itself and the data science techniques required for downstream analysis. Rancho's experienced team of PhD

scientists has both. We can help you integrate your data, mine it to gain insights, and realize its full potential.

Before writing a single line of code, we will meet with you to understand your needs. Our scientists will design a fit-for-purpose solution that may involve custom, open source, and commercial components.

Quality control is built into our workflows to ensure that we deliver a robust solution to you. We will wrap up the project by presenting our findings and sharing all associated code along with complete documentation. We are also available to provide training on new tools or long-term project support.

The Rancho Data Science team can help you with:

- **FAIR**ification of large datasets to ensure data from disparate sources and formats is Findable, Accessible, Interoperable, and Reusable
- Interactive data visualization, including R-shiny, to help you explore your data and make hypotheses
- Python, R, and other scripting needs
- Machine learning
- Statistical analysis



## DATA MANAGEMENT

### - Data management

is an IT practice encompassing an organization's practices across the data life cycle

### - Data governance

is a business practice that defines how data is processed across the organization to ensure that data is FAIR (compliant, private etc)

### - Data modeling

is a process of creating conceptual, logical, or physical data models for one's data

Life Sciences companies accumulate data with ever-increasing pace, and data that they use to create new treatments or new diagnostics is more diverse and complex than in arguably any other domains. With multiple data sources, data elements, stakeholders, and data consumers it gets progressively harder and harder to scale and to maintain control over data life cycle. There is a strong awareness in the scientific community that more data must produce more insight and knowledge, but this goal has been difficult to attain. Life Sciences businesses are now relying on the practice of data management and data governance as best business practice to wrangle data assets and realize value from data.

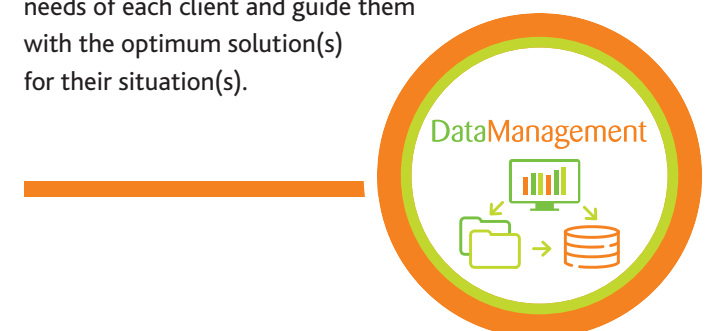
There are many resources that deliver data management and data governance solutions. There are frameworks, platforms, and solutions; there are tools that one can use and knowledge bases to consult with. **We are different.** We provide people – scientists with subject matter expertise trained in data modeling – who can help. Our people have dozens of years of experience in their respective field of study and participated in multiple projects where they did data modeling or provided data governance support. We have real-life experience, and we can support your projects with minimal ramp-up time by leveraging data models we have already built and know how.

### WHAT CAN BE DELIVERED?

- **Data Models:** delivered as ERDs (Entity Relationship Diagrams), in many cases with data dictionaries. We use a variety of tools to deliver data models, Lucidcharts and ER/Studio by IDERA, Inc. being most common

- **Data governance as a process:** we provide scientists trained in data modeling and experienced with data governance projects to support your team operationally. We bridge scientists and IT departments and provide oversight on the data governance process at your organization. We develop community outreach to educate your team on best practices. We write SOPs, post them, and check that people adhere to the recommendations

**Documentation, strategy documents, investigation of third-party tools:** we help clients develop short and long-term data strategies. Often, young companies are focused on research and work hard to develop a new drug or a diagnostic. They find themselves in need to quickly become ready for a regulatory finding, filing (e.g. SEND) but lack experience in-house for this stage of growth. The Rancho team also has the experience to support clients at this stage of their growth. Being a services company, we are platform agnostic and can assess the individual needs of each client and guide them with the optimum solution(s) for their situation(s).



DataManagement

