How can developers treat Ovarian Cancer?

Mridula Jayaraman **(MJ)**
Developer
ThoughtWorks Inc.

Overview

- The Players
- Genomics & Cancer
- The Problem
- Our Solution
- Learnings
- Questions

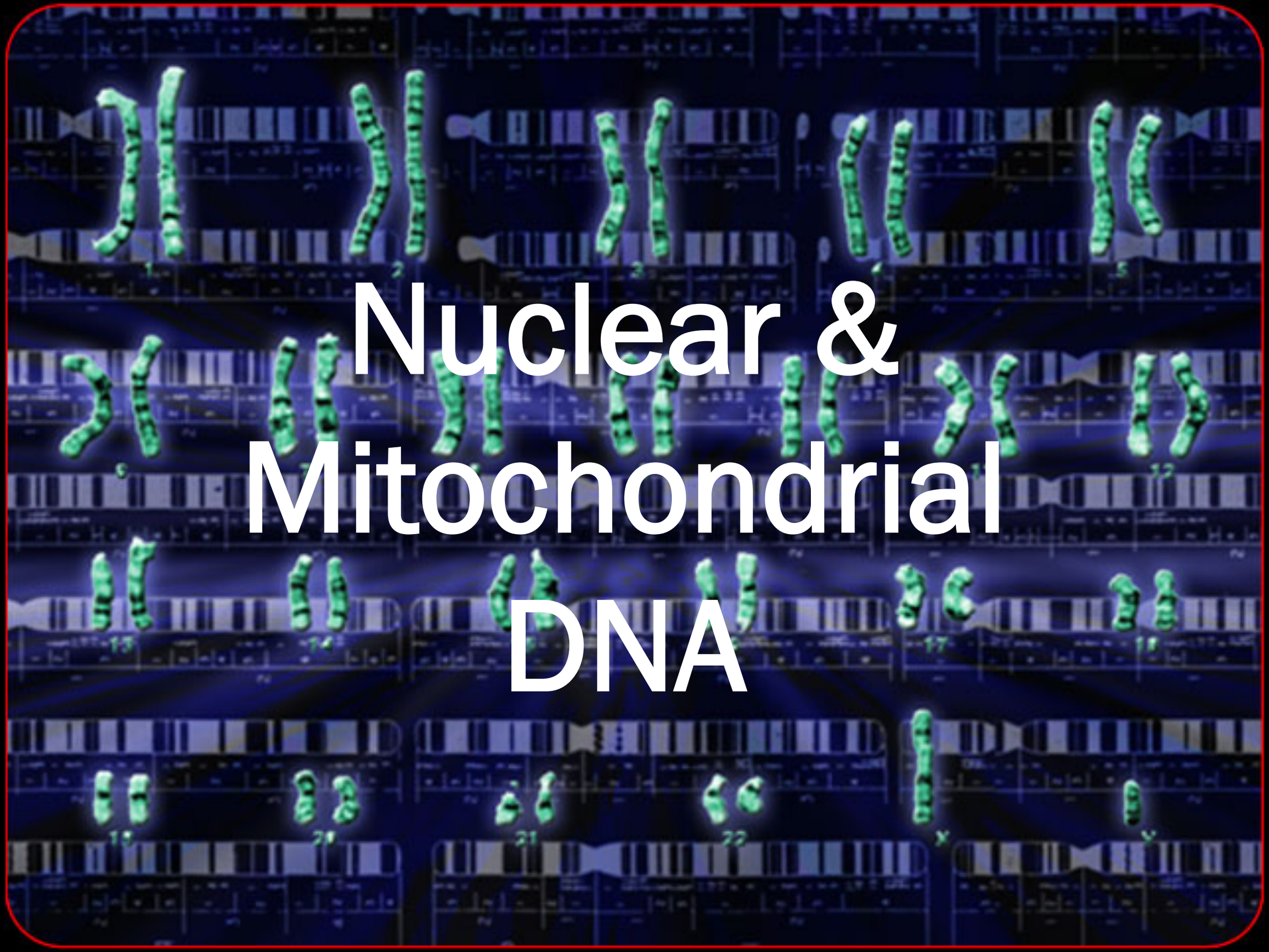
The Players

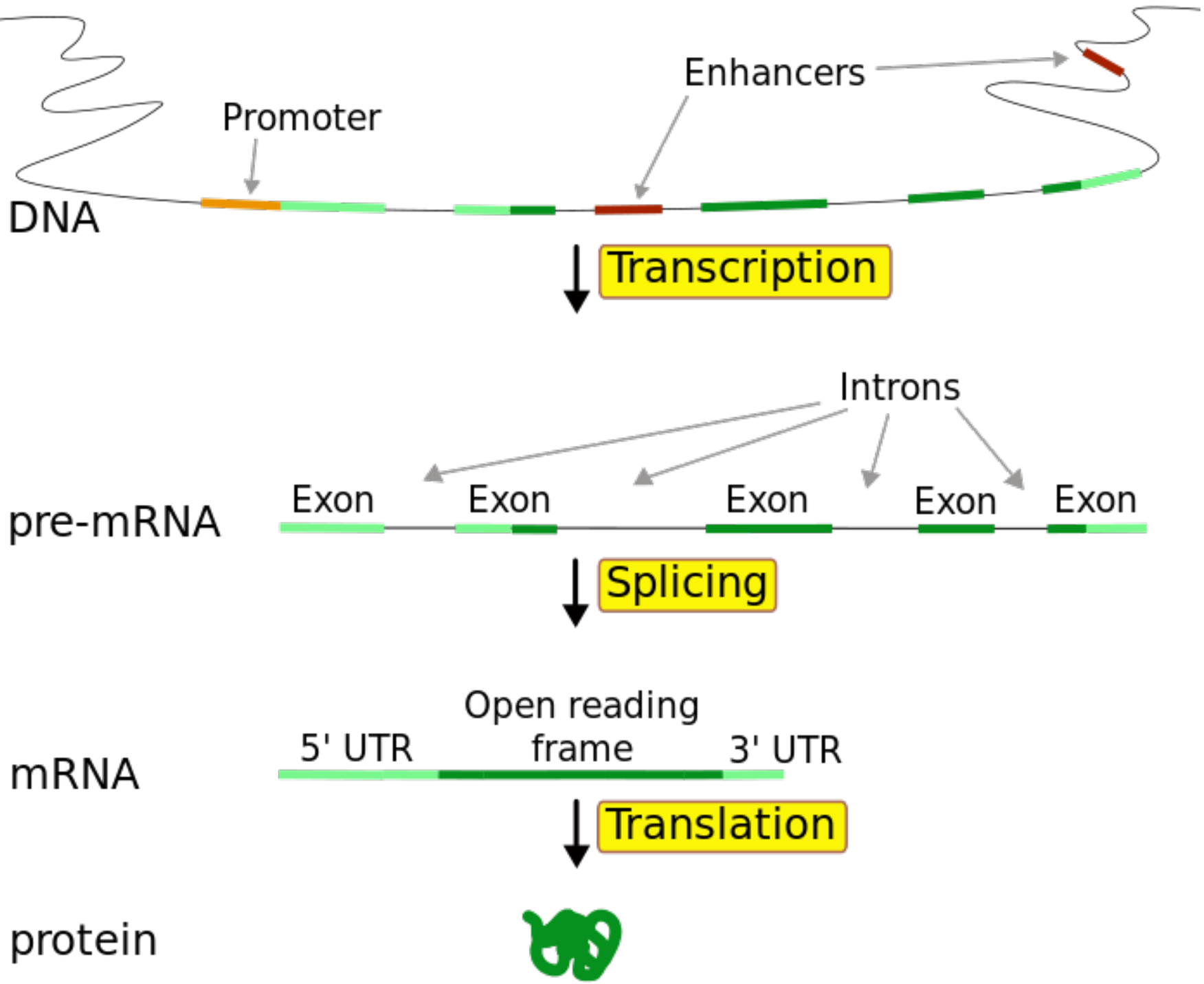
THE CLEARITY FOUNDATION

Powering

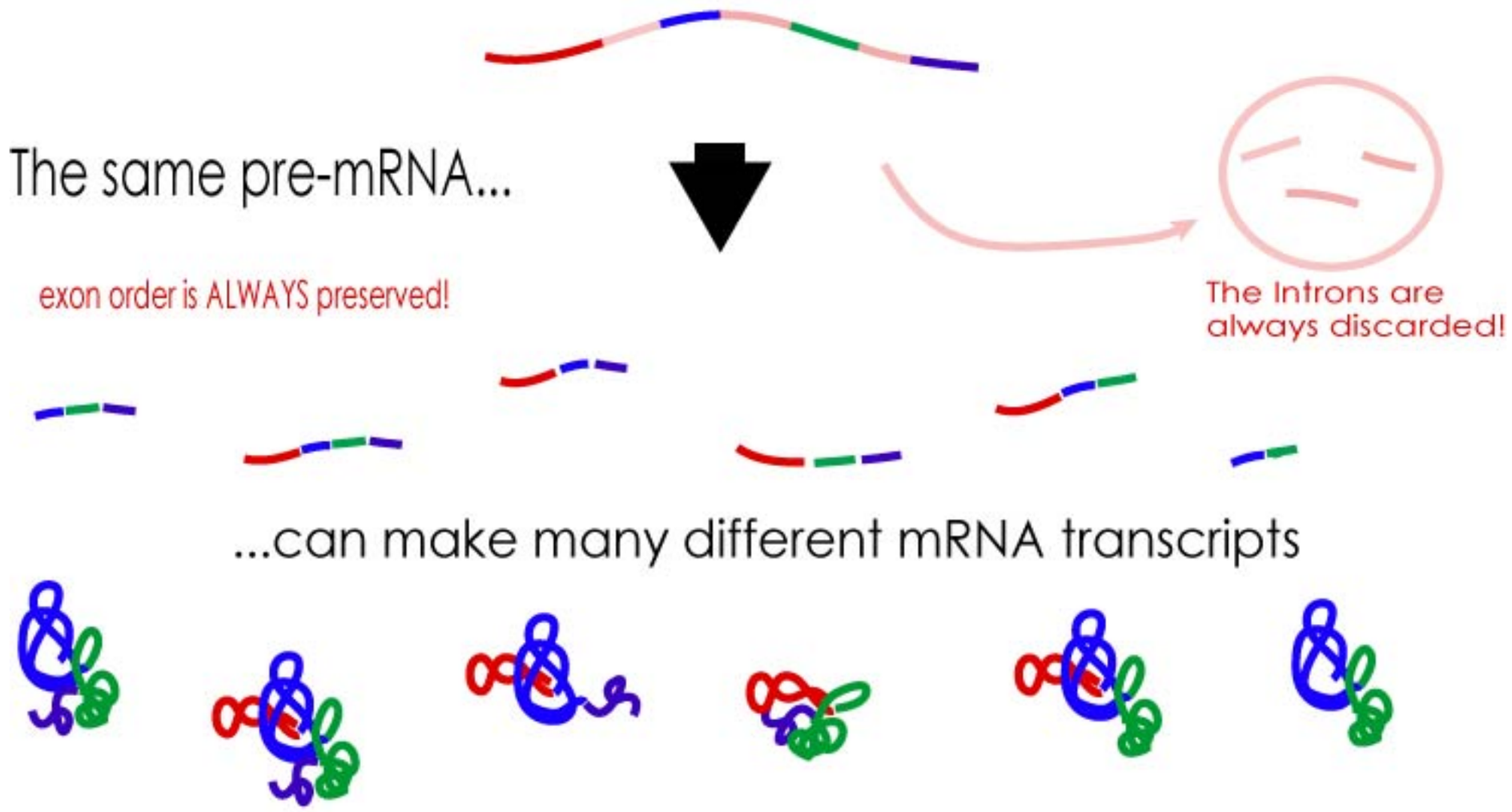
GENOMIC RESEARCH AND MEDICINE

Genomics & Cancer


Nuclear &
Mitochondrial
DNA


DNA
Promoter
Enhancers
Transcription
pre-mRNA
Introns
Exon
Exon
Exon
Exon
Exon
Splicing
mRNA
5' UTR
Open reading
frame
3' UTR
Translation
protein

Alternative Splicing



DNA= A Lot of Data

- 3.2 billion base pairs
- ~ 25000 genes (3% of genome)
- 0.5 Tb sequence data / normal genome

Mutations

Cancer is a Disease of Abnormal Gene Function

- Modifications within the DNA
- **Genotype to Phenotype implications vary**
- Increasingly vital in treatment choices

Ovarian Cancer Treatment

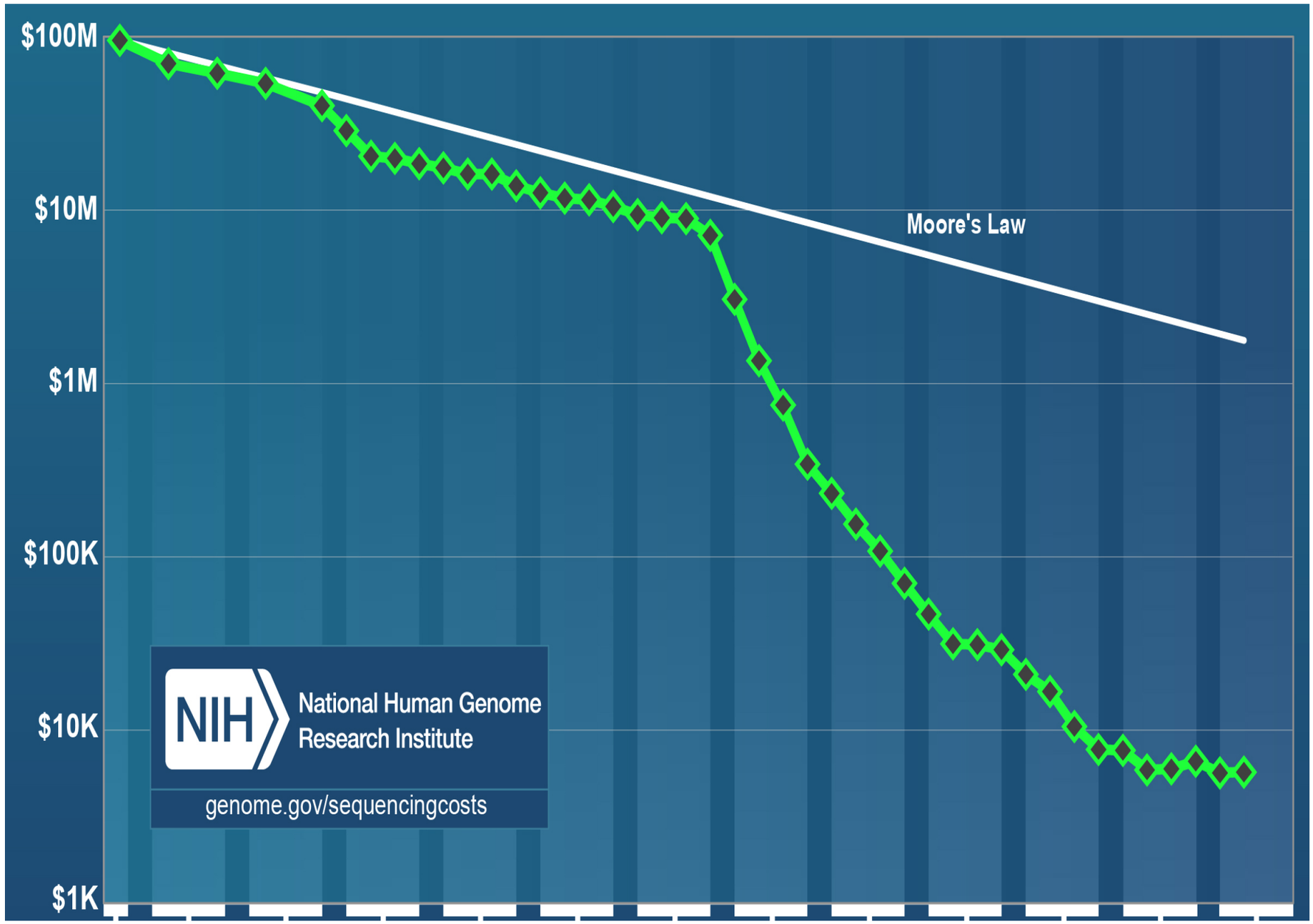
22,000

~ 14,000

Standard of Care

- Organ-based
- Expensive (~ $100,000/round)
- Cancer will reoccur in 75 - 80% women
- Chemotherapy drugs become less effective with time

What is Changing?



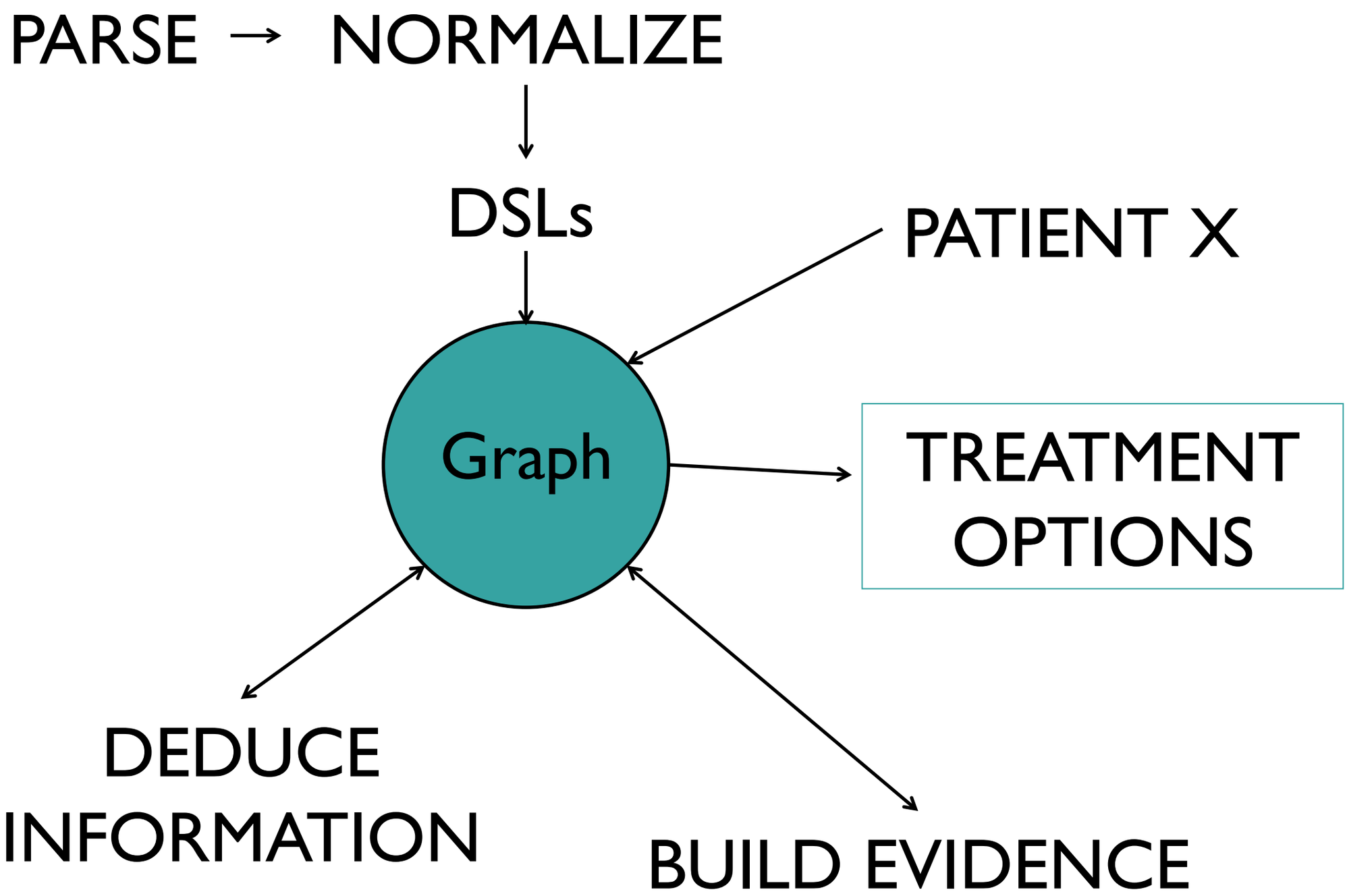
ThoughtWorks®

What Did We Know?

- Lots of data (0.5 TB!)
- Access to cancer/normal sequencing data
- Access to some connections between patient biology & treatments

What Did We Do?

Start building


PARSE → NORMALIZE
DSLs
PATIENT X
Graph
TREATMENT OPTIONS
DEDUCE INFORMATION
BUILD EVIDENCE

PARSE

The Data Reference

- Genes
- Genomes
- Biomarkers
- Drugs
- Treatments

PARSE

LOTs of Data

Patient

- DNA Sequencing data
- RNA sequencing data
- Biomarker test results (IHC, FISH etc)
- Treatments

PARSE Important Data

Treatment Recommendation

- Link between a particular state of “something” in a patient & a viable treatment option

Internal DSLs

Example

Parse & normalize in Ruby

↓

Turn it into a data structure

- Clojure DSL

↓

Nodes and relationships created in db

– Neo4J & Clojure

Neo4J

Deduce treatment options

Patient ->

Biology ->

Implication ->

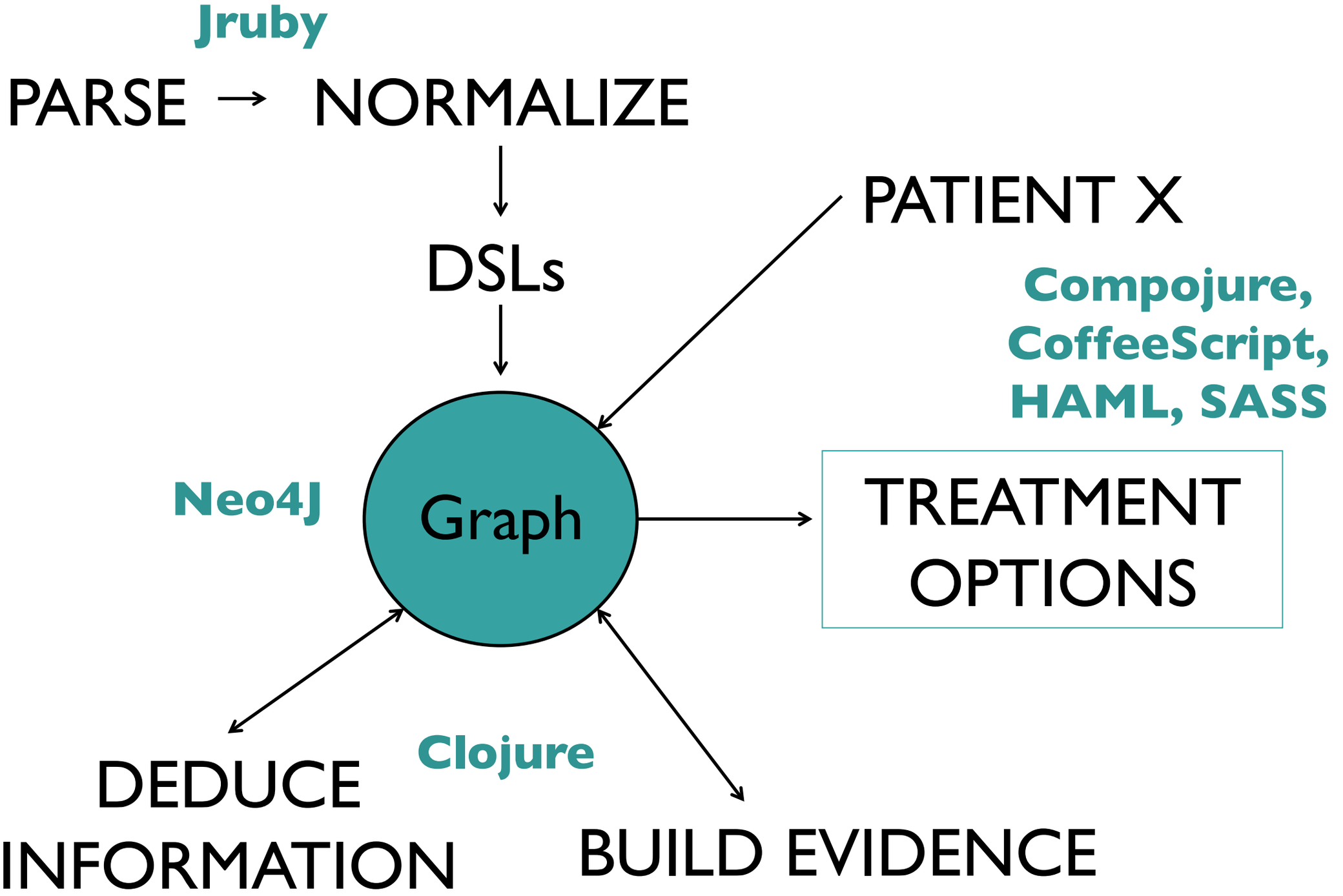
Treatment

Clojure

User-Centric Design

Front-End on Demand

END to END Flow of Our System



Infrastructure: AWS, Python, Puppet

What did we achieve?

- Clearity scaled treatment
- Evidence based system
- Generate alternate drug therapies

Learnings

The HUGE Issues

- A lot of data
- Poor quality data
- Manual intervention
- Code by non-devs/academics
- Recreating the wheel

Mitigations

- Normalize data from many sources
- Feedback & learning

Some Others

- Important to keep up with the domain
- Working with the research community
- LEGAL implications
- Wrong data?

Genome
Splicome
Functome
Alternatome
Microbiome
Nucleome
Foldome
Introme
Exome
Immunome

A long way to go

Questions?

Thanks!

mjayaram@thoughtworks.com

@MridsJayaram