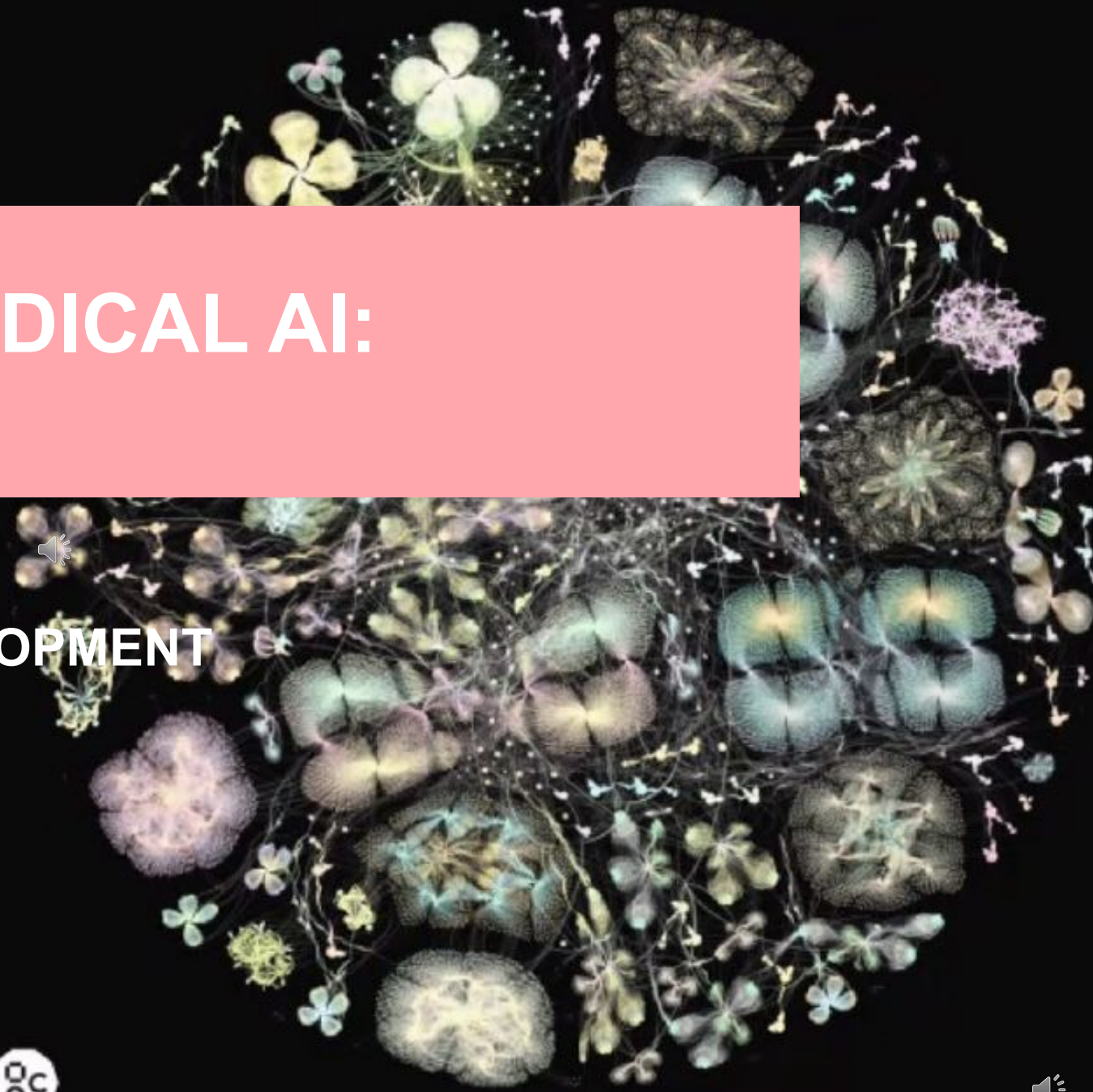
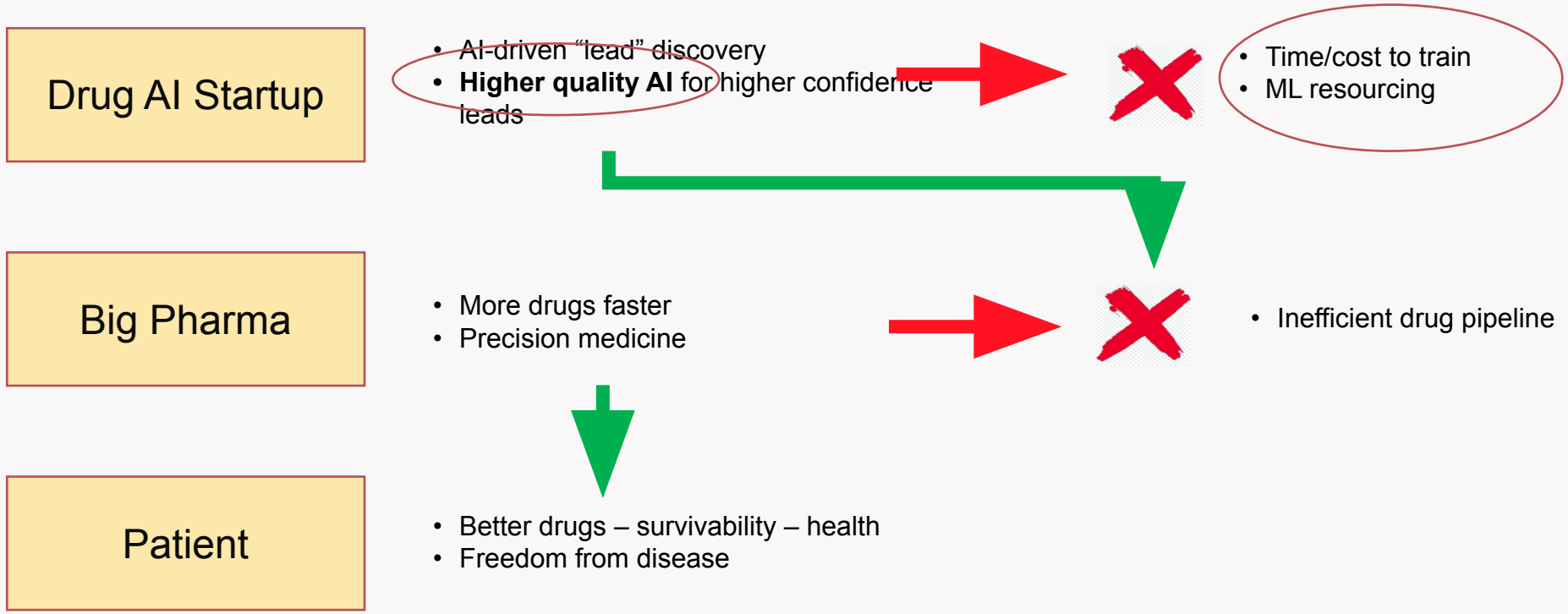
USE OF GNN IN BIOMEDICAL AI: DRIVING AI QUALITY



TOM WILSON, VP BUSINESS DEVELOPMENT
HEALTHCARE AI

GRAPHCORE

SOLUTION CHAIN: PHARMACEUTICAL



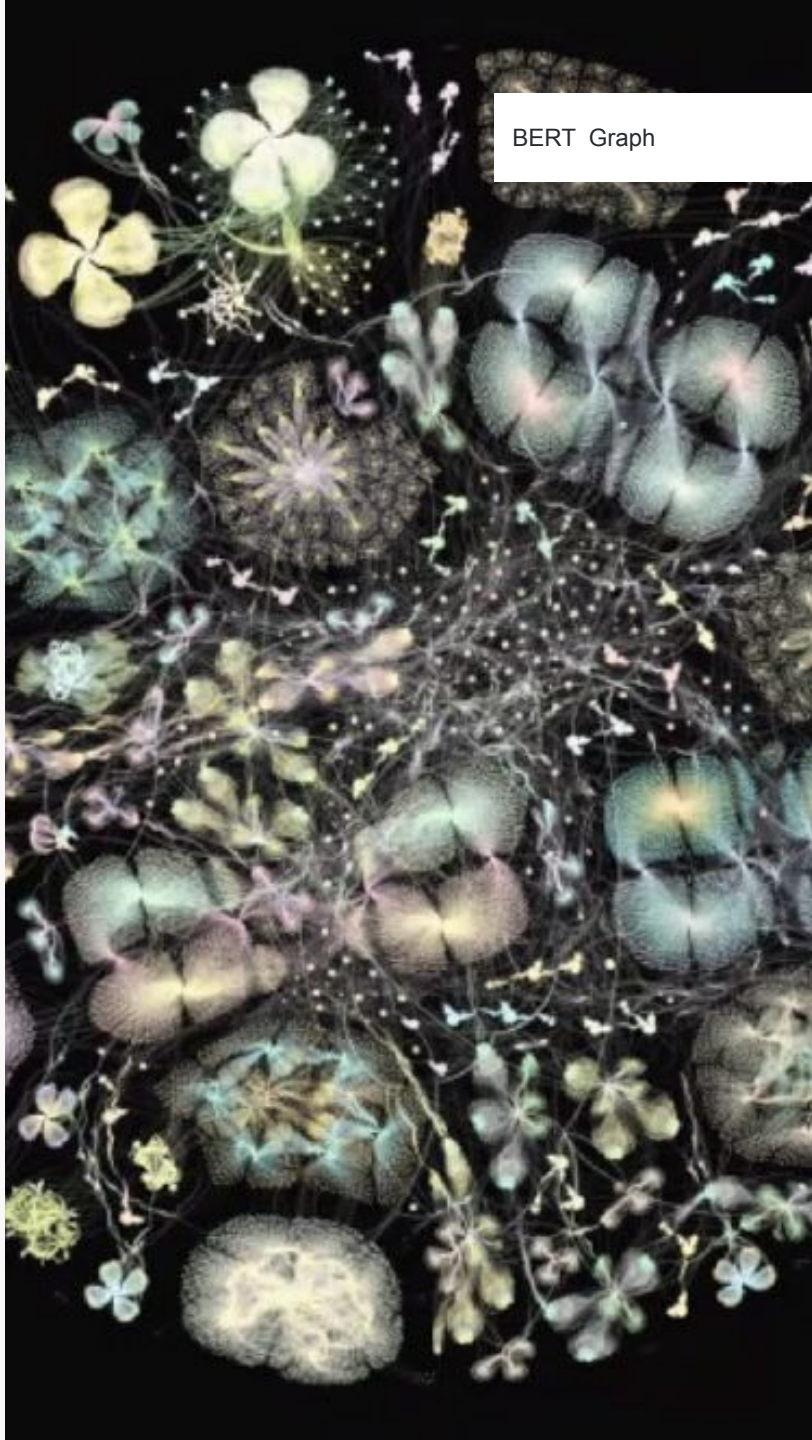
COMMON CHALLENGES

Time to Train and/or Training Costs

- Long Time to Train restricts time for experimentation and productivity
- Training Costs are exceeding initial projections due to:
  - Increasing complexity and size of Knowledge Graphs (KGs) and datasets
  - Churn in dataset content requiring more frequent pre-training
  - Access to CSP Processor instances may be a bottleneck

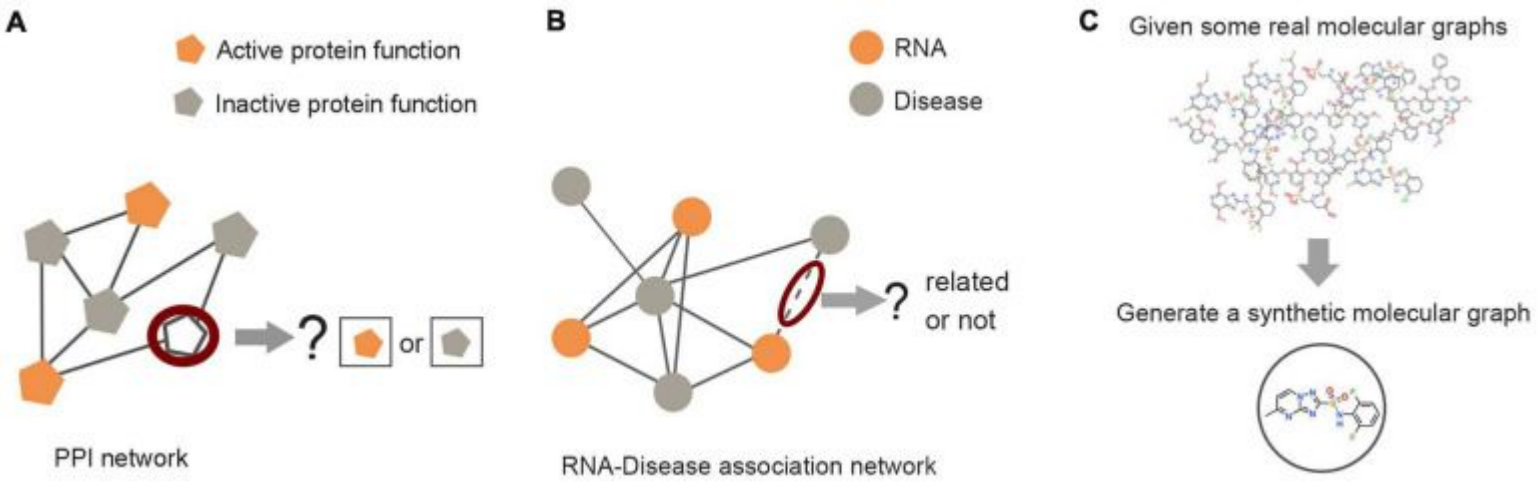
Significant Shortage of ML Resources and Experience

- Highly competitive AI startup environment vs a relatively limited talent pool
- More use cases and experiments than can be resourced
- Lack of familiarity with new ML areas like Knowledge Graphs, Graph Autoencoders, GNNs, etc.



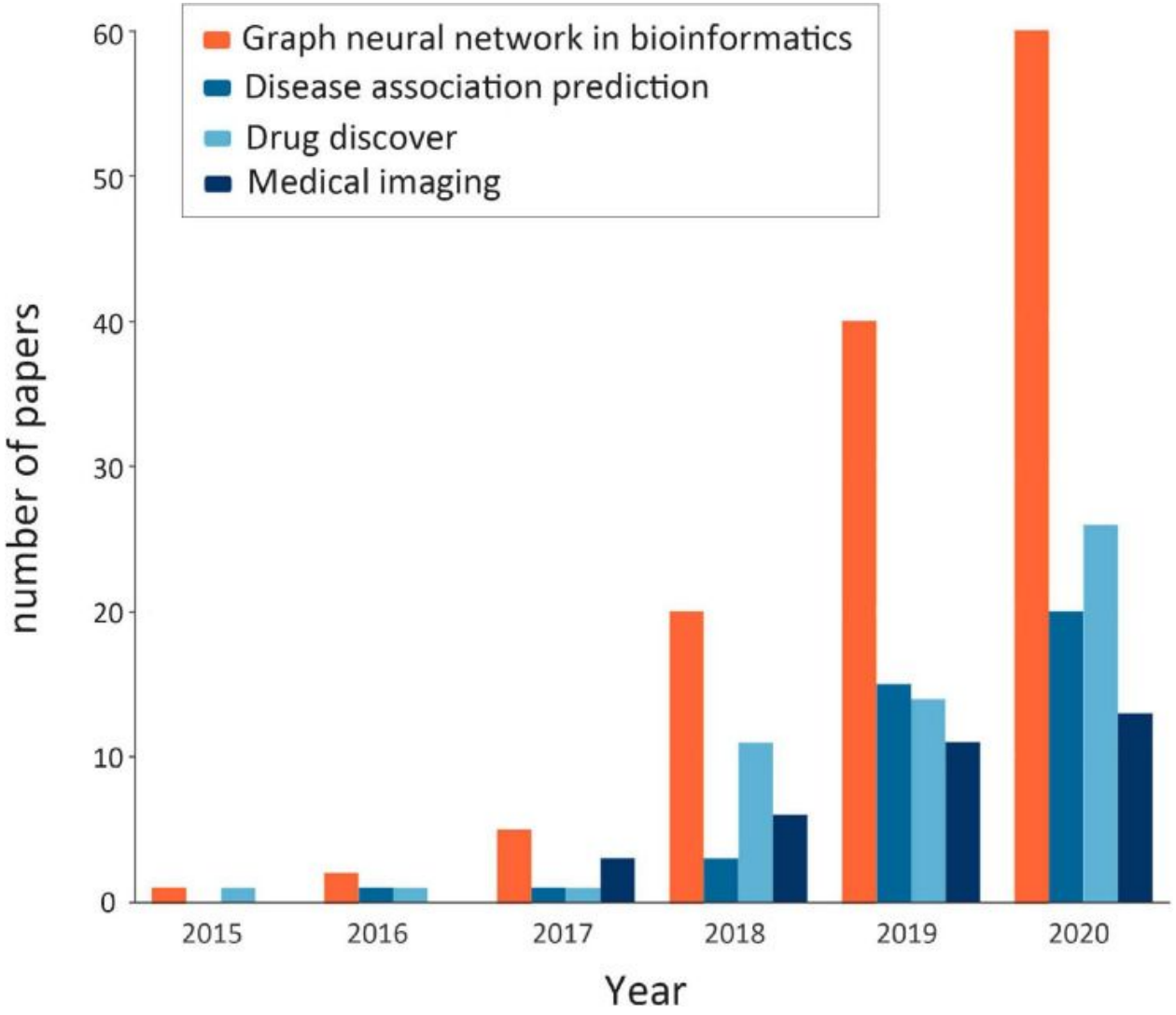
GRAPH NEURAL NETWORKS

- **Optimized to learn from graph-shaped data and ideal for modeling the dependencies between nodes in a graph**
- **Three main uses**
  - A - Node classification: E.g. protein function classification in PPI.
  - B - Link Prediction. E.g. new drug disease relationships
  - C - Graph generative task E.g. generate a synthetic molecular graph
- **Much of the data of interest can/should be represented as a Graph structure: molecules, interactions, pathways, drug-disease interactions, knowledge graphs in general….**
- **Application of better learning from graph-shaped data = Higher Quality AI**



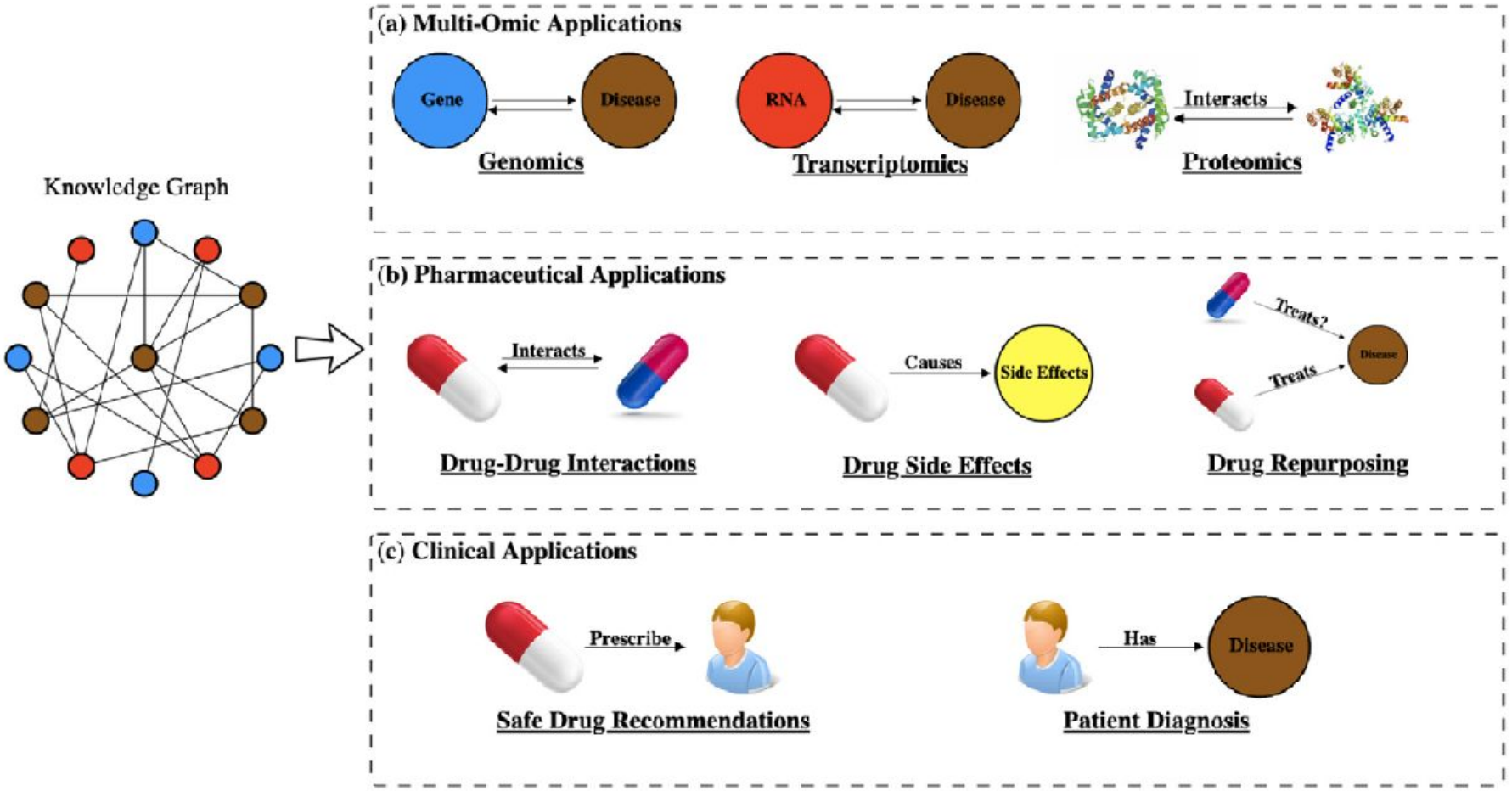
From Zhang et al, 2020: Graph Neural Networks and Their Current Applications in Bioinformatics (nih.gov)

INCREASING GNN INVESTIGATIONS



From Zhang et al, 2020: Graph Neural Networks and Their Current Applications in Bioinformatics (nih.gov)

KNOWLEDGE GRAPHS



Constructing knowledge graphs and their biomedical applications; David N. Nicholson a, Casey S. Greene 2020

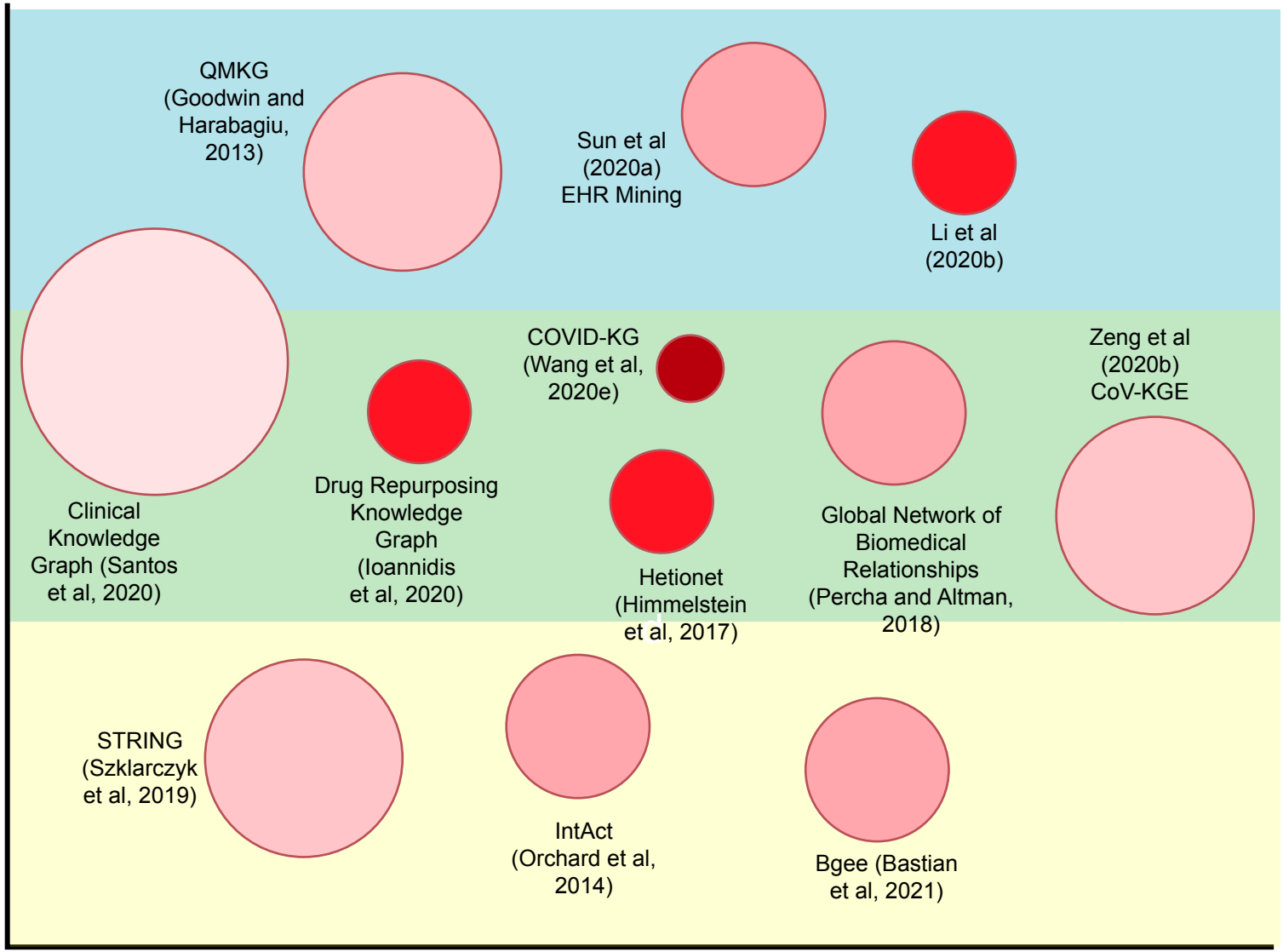
KNOWLEDGE GRAPH OPEN-SOURCE EXAMPLES

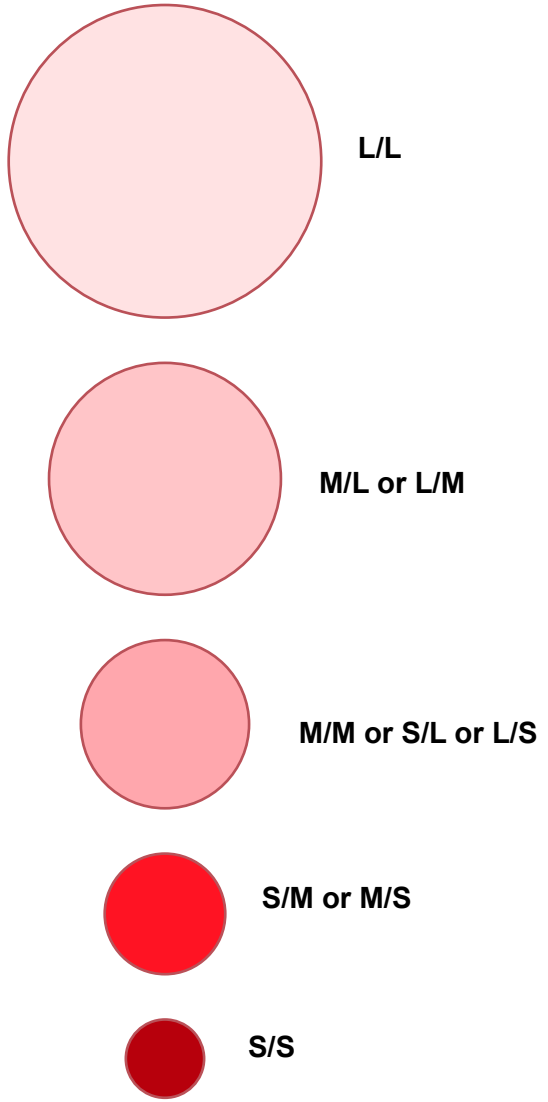
Node/Edge Sizing
Large - > 10M to > 100M
Medium - >100K to 10M
Small – 10K to 100K
N/E used below

Clinical Applications
Diagnostic Recommendations, Insurance

Pharmaceutical Applications
Precision Medicine
Drug Reuse, Literature Mining

Multi-omics
Drug Discovery, Genomics Analytics

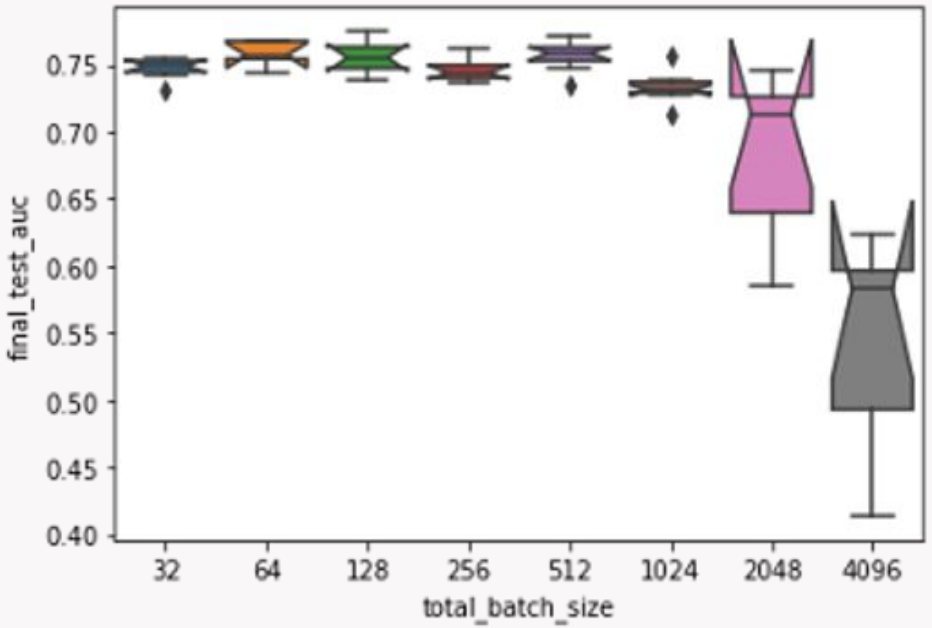


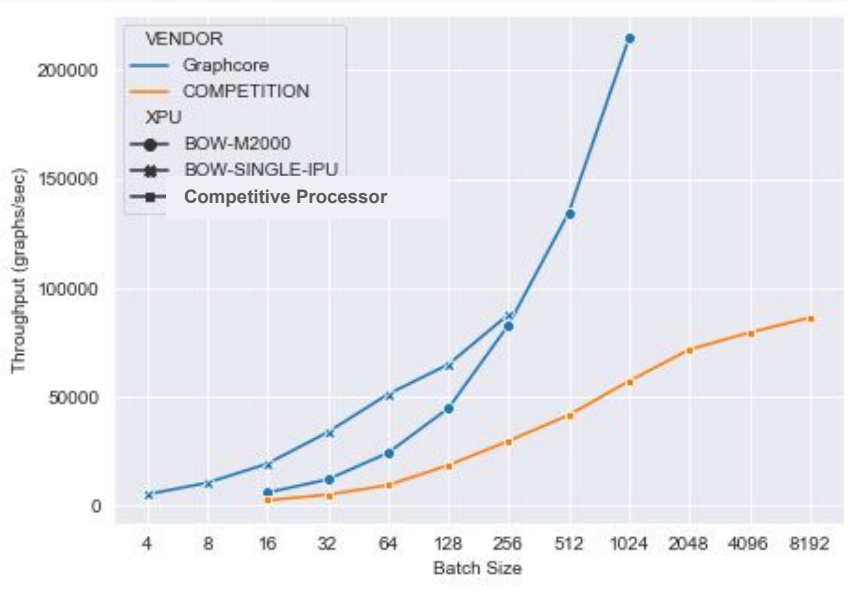


Derived from https://graph-neural-networks.github.io/static/file/chapter24.pdf

GRAPH ISOMORPHISM NETWORK FOR OGBG-MOLHIV

- The Graph Isomorphism Network[1] (GIN) is a popular architecture, here used as a baseline for the ogbg-molhiv (binary molecule classification) task.
- 41K graphs, 25.5 nodes per graph, 9-dimensional input node features
- 1.7M model parameters
- Single chip and BOW-M2000 (4x IPU) @ FP16 vs Competitive processor (contact tomw@graphcore.ai for details)
- Large batch size harms training accuracy for fixed #epochs.
- Also note an efficiency cross-over for single IPU vs M2000 beyond ~ 256 BS

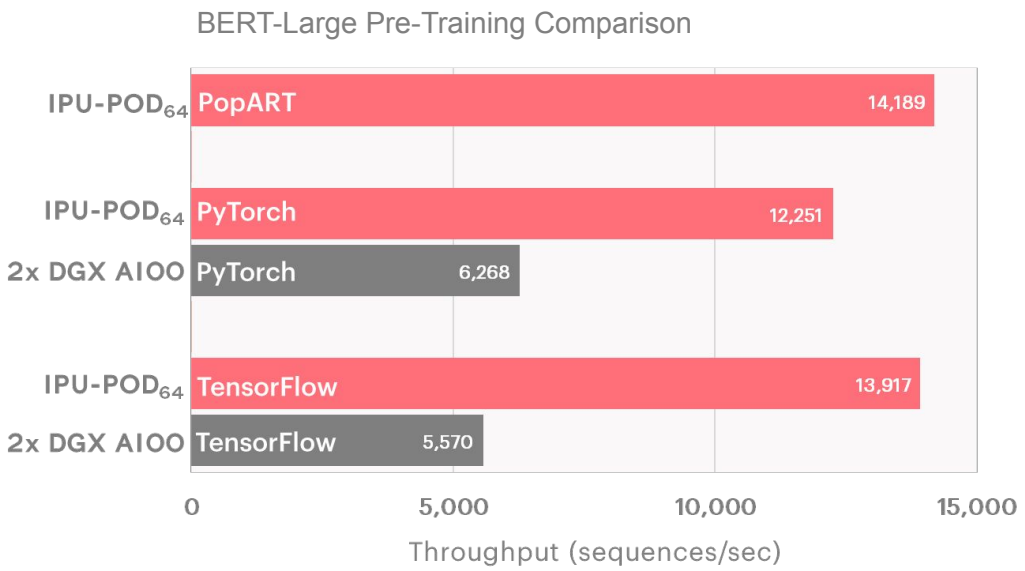




[1] Xu, K., Hu, W., Leskovec, J. and Jegelka, S., 2018. How powerful are graph neural networks?. *arXiv preprint arXiv:1810.00826*.

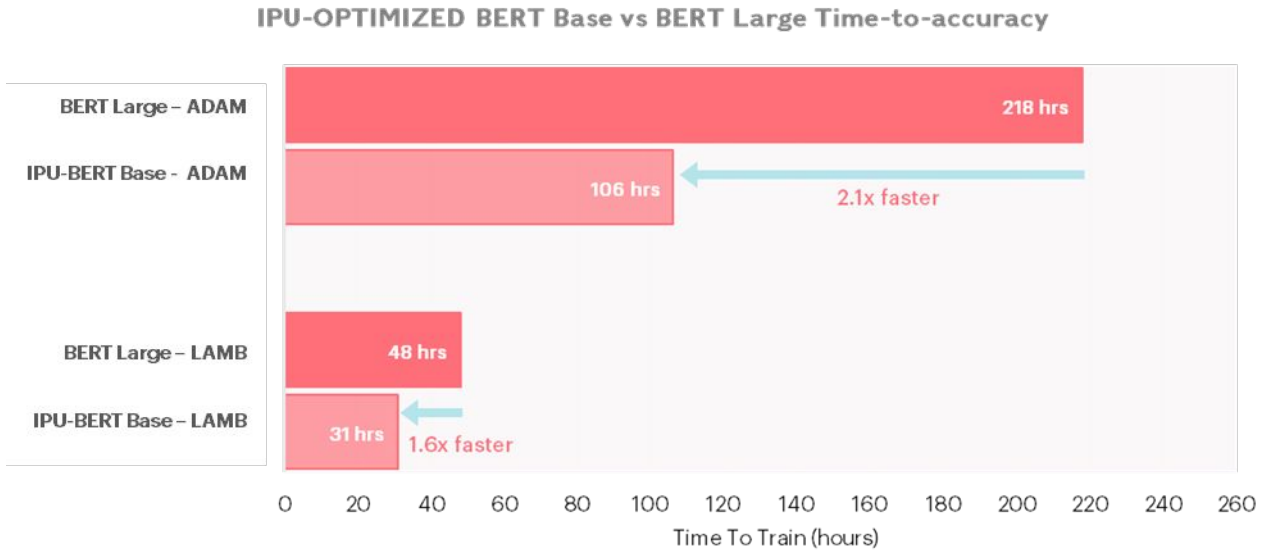
TRANSFORMER TRAINING PERFORMANCE

“Vanilla” BERT-Large



NOTES: (updated 20th Dec 2021) | headline speedup based on TensorFlow vs TensorFlow throughput comparison
BERT-Large Phase 1 Pre-Training Throughput (SL128)
IPU-POD64 (16x IPU-M2000) | SDK 2.4.0 | https://www.graphcore.ai/performance-results
DGX A100 320GB (A100-SXM4-40GB) | Mixed Precision | Using 97% perf scaling based on published BERT results
DGX A100 PyTorch https://developer.nvidia.com/deep-learning-performance-training-inference 1st Dec 2021
DGX A100 TensorFlow - https://ngc.nvidia.com/catalog/resources/nvidia:bert_for_tensorflow/performance

ML-Optimized BERT-Large



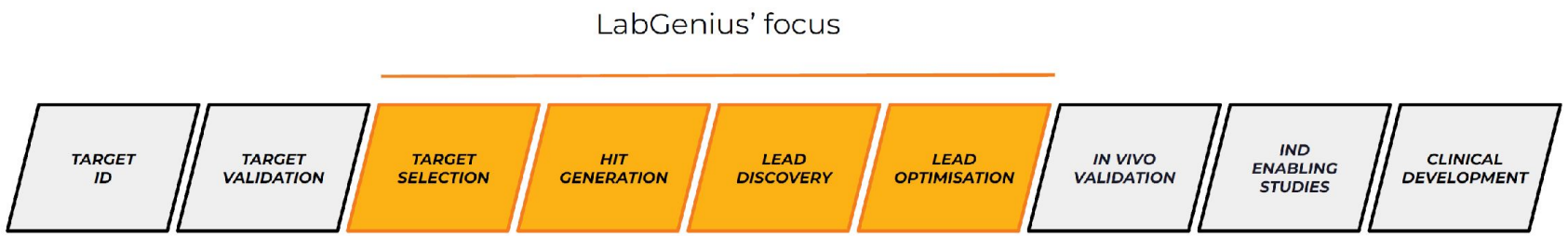
IPU-Optimizations leverage IPU architecture capabilities
adds grouped convolution module
2x reduction in # parameters
2x faster time to train
improved accuracy of results

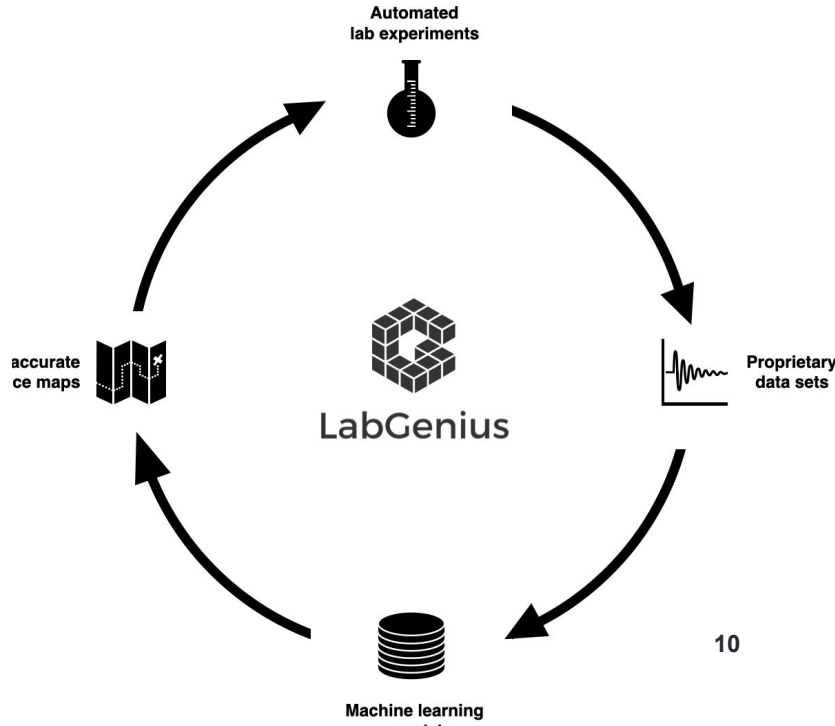
PREDICTING RELATIONSHIP BETWEEN PROTEIN STRUCTURE AND FUNCTION: LABGENIUS

"We use about a hundred million protein sequences and train our model on the IPU using BERT, adapted from the natural language processing field. Previously, we were using GPUs to perform the same task. It took us around a month or more to get a trained model. So, the idea of iteratively improving the protein representation by tweaking the architecture and/or the training data composition was pragmatically out of reach for us. As soon as we started using IPUs, it felt like our hands were untied. The turnaround time now for training a model from scratch is less than 2 weeks, using hundreds of millions of protein sequences - at least twice as fast."

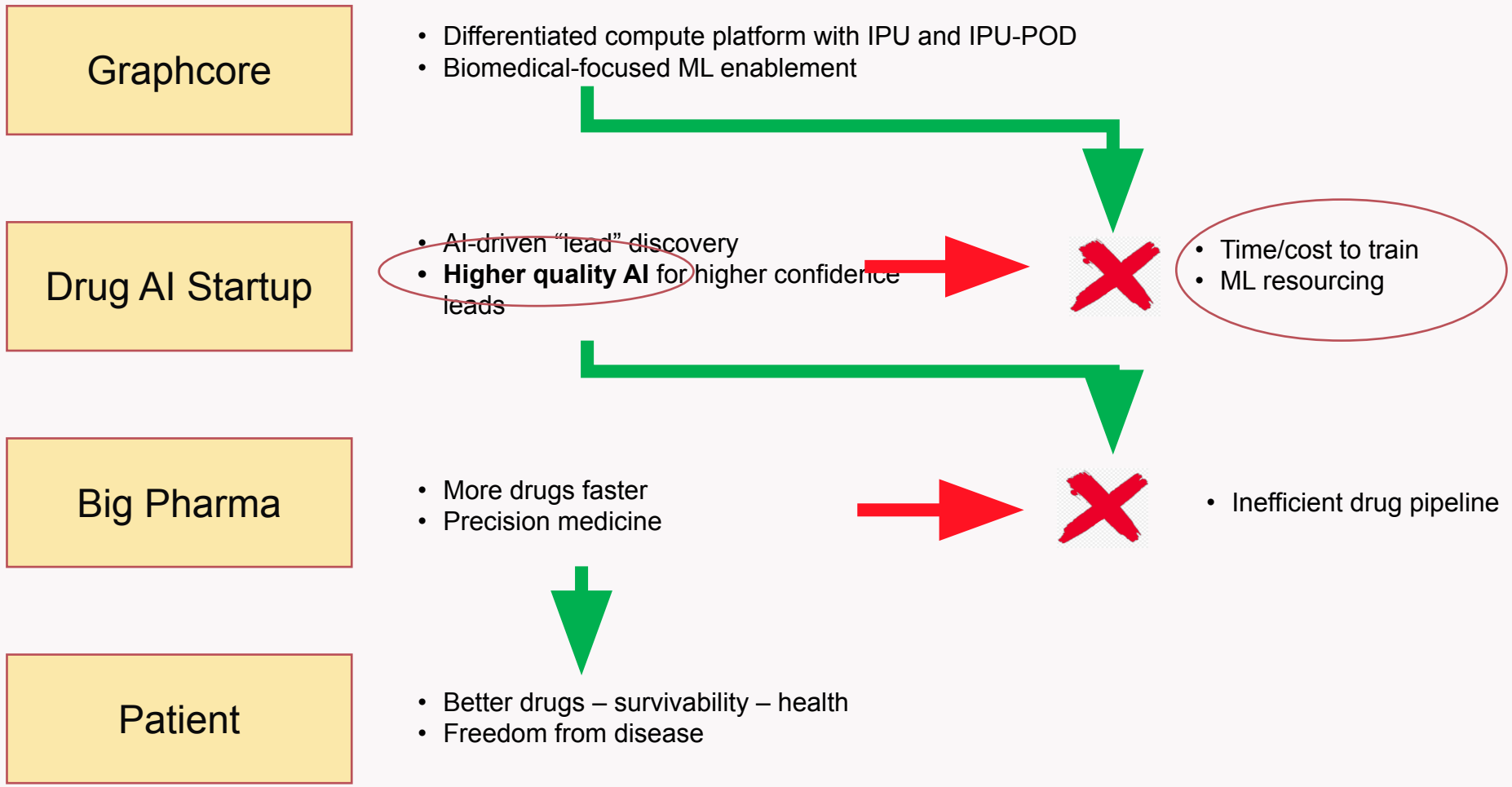
Katya Putintseva, Data Science Lead at LabGenius

LabGenius





SOLUTION CHAIN: PHARMACEUTICAL



THANK YOU

Tom Wilson
tomw@graphcore.ai