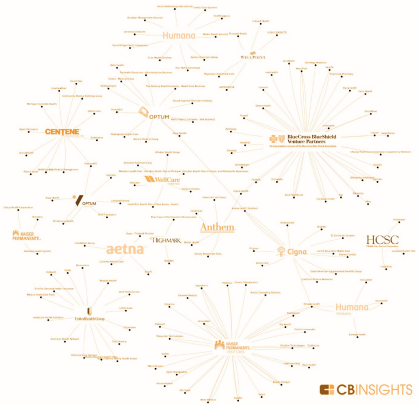
EntityGraph

Beyond BSG



Welcome Back! Business Social Graph (BSG)

Project Gandalf

Josh Trach & Robert (HackDay13)

Technical Background

Credits to previous projects that lays out the foundation:

- Business Relationship Classifier

Microsoft	WS03	17178150	1	NULL	At this rate , Microsoft and its partners would need to upgrade 20,000 in...
Box	Grishin Robotics	18831150	1	NULL	and Box -- with additional participation from new investors Morgan Stanl...

- Operating Metric Extraction

Netflix	worldwide subscribers	16000000.00	Netflix , as you 've likely heard , is now in over 50 % of all U.S. households and
Nium	transactions	2000000000.00	Nium 's platform reportedly handled $2 billion in transactions during Q1 2020 .

- Suggested Tags

id_company	company	id_cluster	cluster_score	group_leader	group_tags
52,882	Uber	2	7.060251236	rideshare	rideshare
52,882	Uber	1	6.0795321465	ride-sharing	ride-hailing, ride-sharing
52,882	Uber	0	4.613509655	car sharing	car sharing, carpooling
52,882	Uber	4	2.1531009674	food delivery	food delivery
52,882	Uber	3	1.936373353	car rental	car rental
52,882	Uber	5	1.7905124426	limousine	limousine

1

One step further

What goes on the EntityGraph?

NODE TYPE

ENTITY:

company/investor (cbi entities)

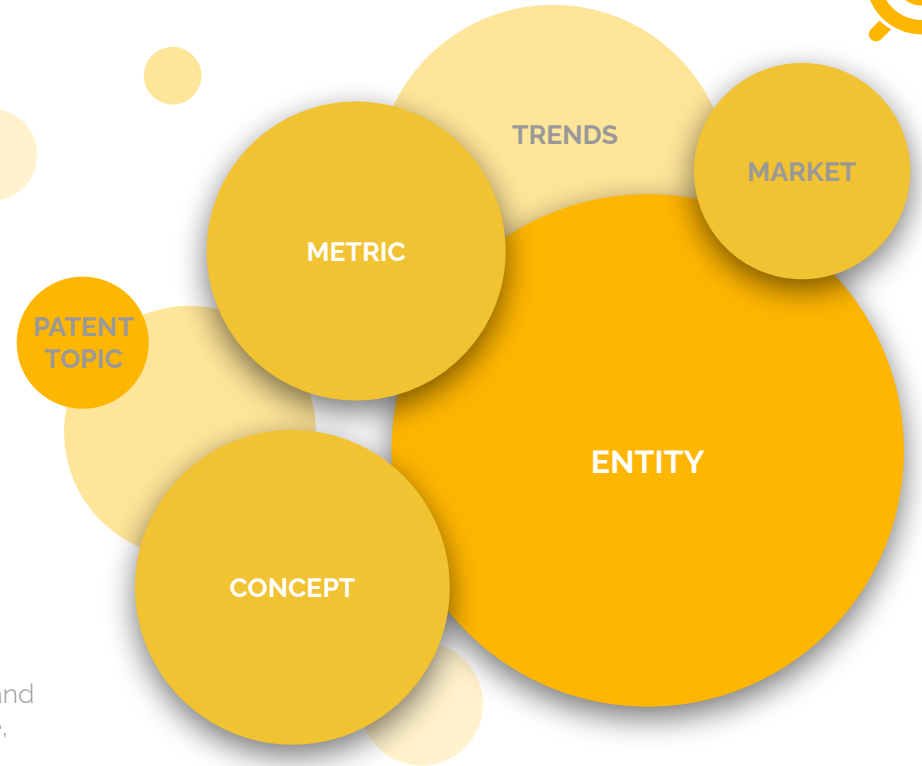
CONCEPT:

suggested tags

METRIC:

operating metrics

Node attributes: (status, funding/valuation size, industry and type) for entity node, word embeddings for concept node, numeric type for metric node.



EDGE TYPE

ENTITY-ENTITY

`PARTNER`, *`COMPETITOR`* are added as business relations.

To be more practical, consider adding more BR types, e.g. Vendor-Client, etc. Or target company-investor relationship

ENTITY-CONCEPT

Based on whether Tag as Concept is learnt from the ENTITY itself or ENTITY similar peers and confidence score of this association.

`STRONG_SELF_ATTR`, *`SELF_ATTR`*,

`LEARNT`,

`WEAKER_LEARNT`

ENTITY-METRIC

Based on metric amount, we bucket relation as:

`TRILLIONS`, *`BILLIONS`*, *`LESS_THAN_1M`*, *etc.*

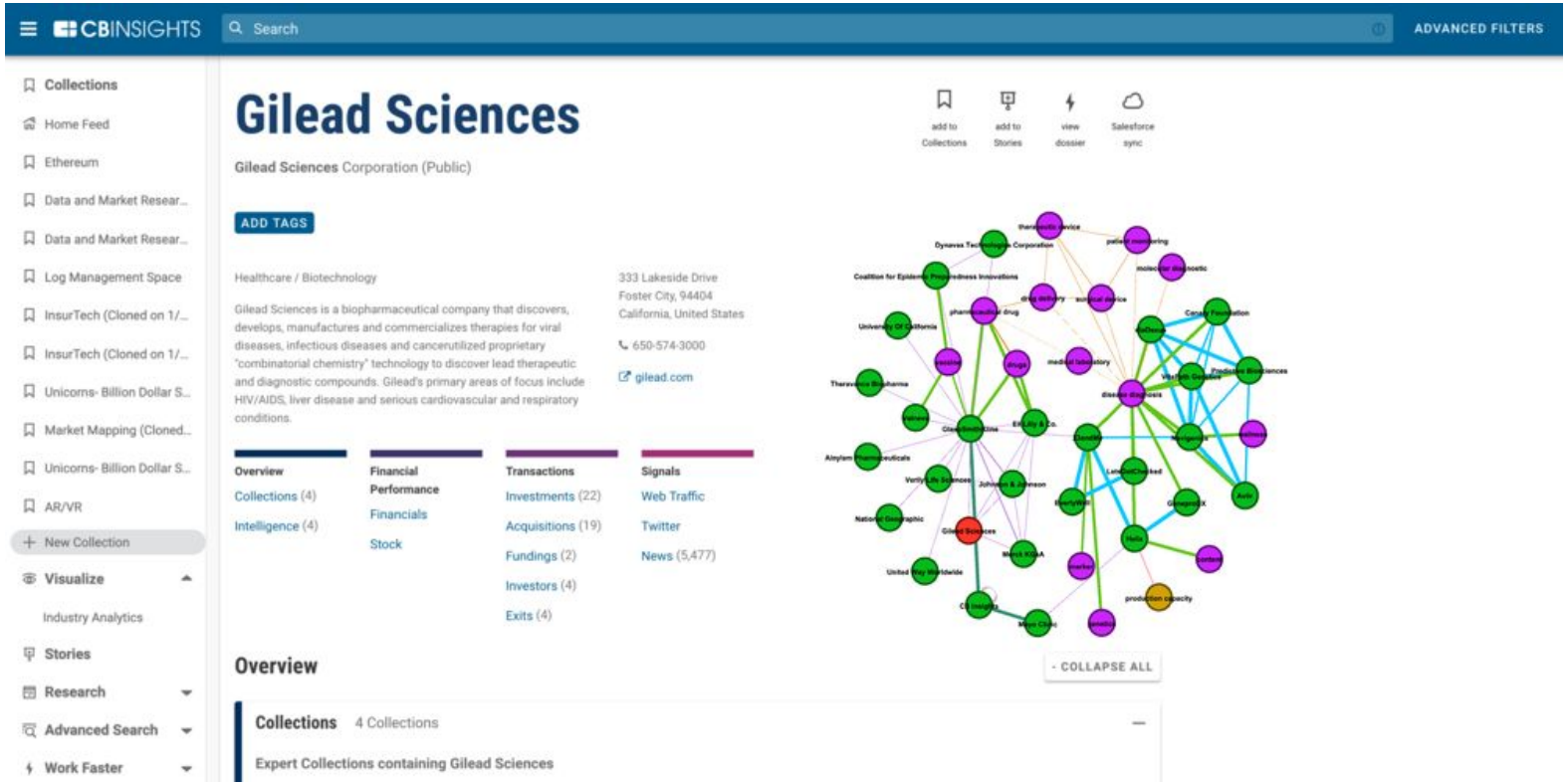
METRIC-METRIC

For Metrics that are relevant, e.g. `revenue-`ads revenue`, we define *`RELEVANT`* edge type

Case 1:

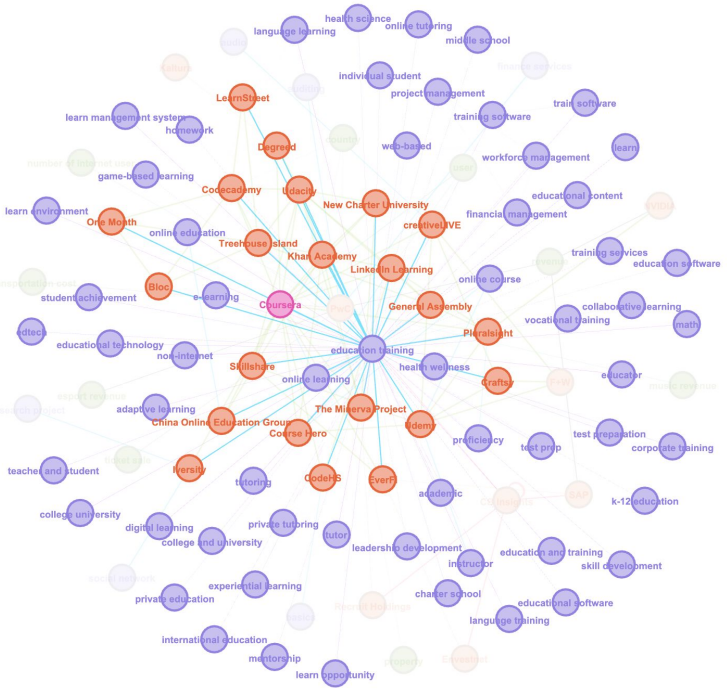
Profile Visualization

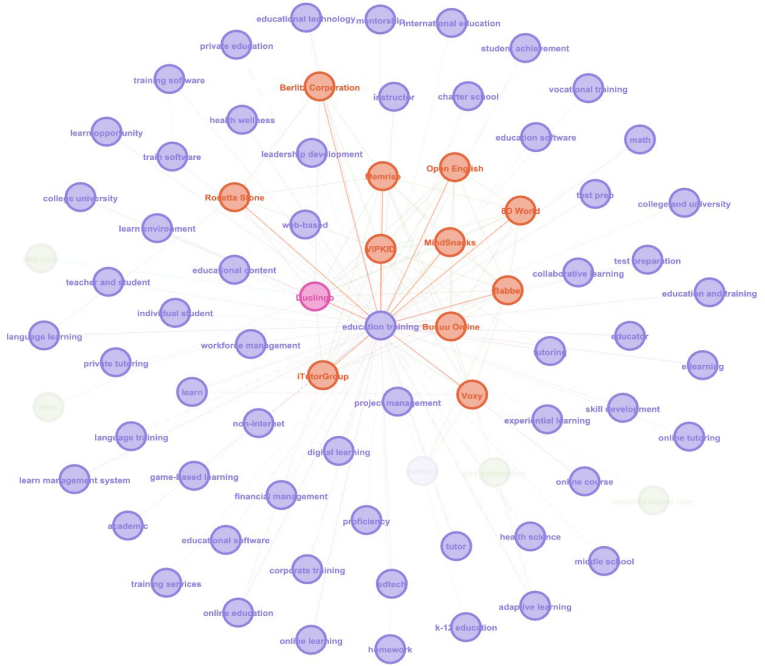
Pick one Node on the Graph to View All Nodes and Edges it has connection with.

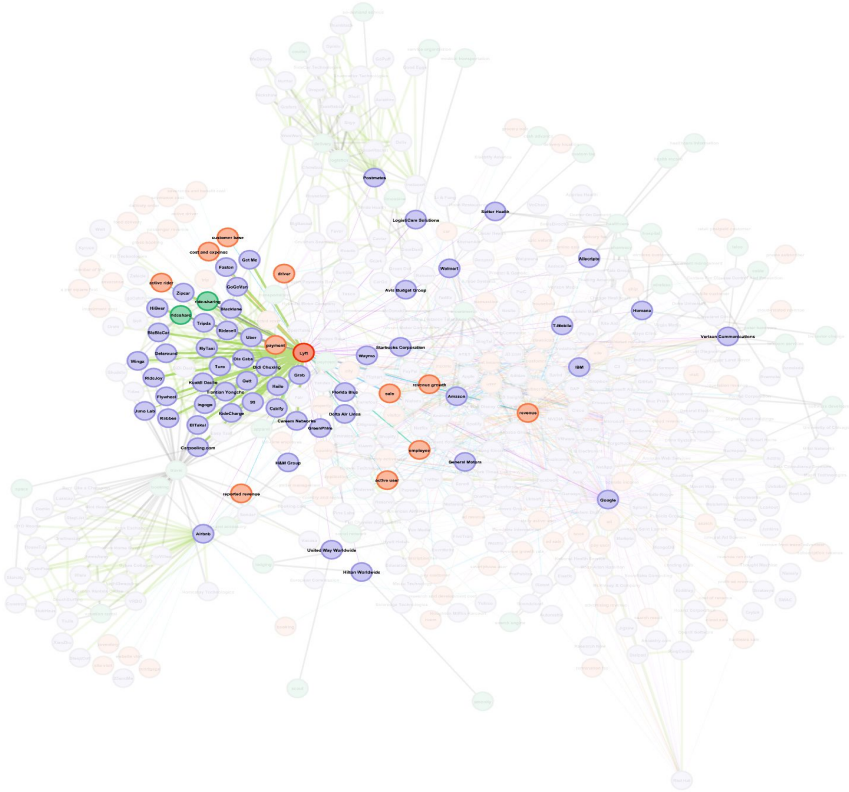


SubGraph view at one Entity Node & defined Search Depth

Reveal Competitive Strategy



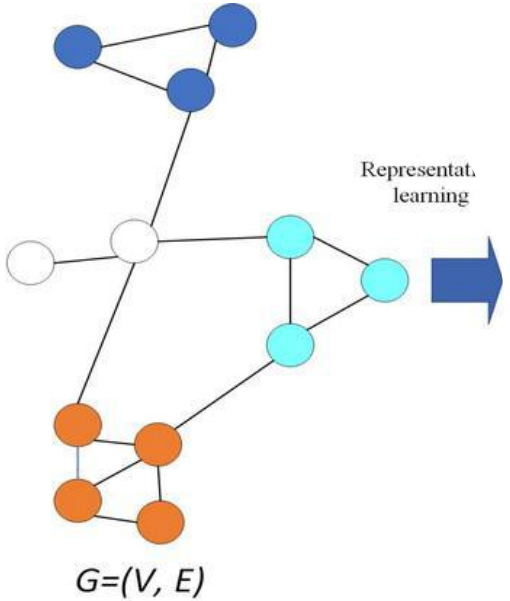




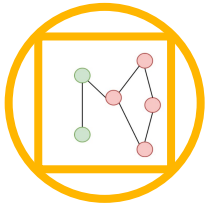
SubGraph View on Node with More Connections

Case 2:
Node Embeddings

A node embedding vector capture latent/hidden information about the nodes and edges, and can be used for (semi-)supervised downstream tasks.

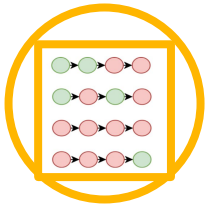


Node2Vec with DeepWalk



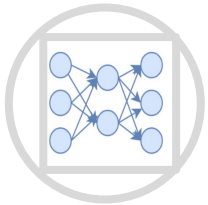
Initializing Whole Graph

Create a Unidirectional Graph.
With Node and Edge Information.



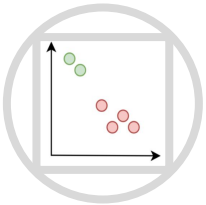
Sampling Random Walk

A graph is sampled with 32 to 64 random walks with a length of about 20-100 steps.



Training skip-gram Model

Skip-gram network accepts a node from the random walk as a one-hot vector as an input and maximizes the probability for predicting neighbor nodes



Computing embeddings

Embedding is the output of a hidden layer of the network.

Node Similarity

	Relevant Concept	Relevant Company	Relevant Investor	Relevant Operating Metrics
23andMe	blood test	Helix	Central Bank of Nigeria	website visit
	personalized medicine	diaDexus	LFM Capital	production capacity
	marker	VitaPath Genetics	Elisa	consumer transaction
Lyft	public transportation	BiTaksi	Tata Motors	food delivery
	car rental	99	Carl Icahn	active driver
	daily commute	Wingz	IPM Group	gross booking

Case 3: Know your client

Client Segmentation

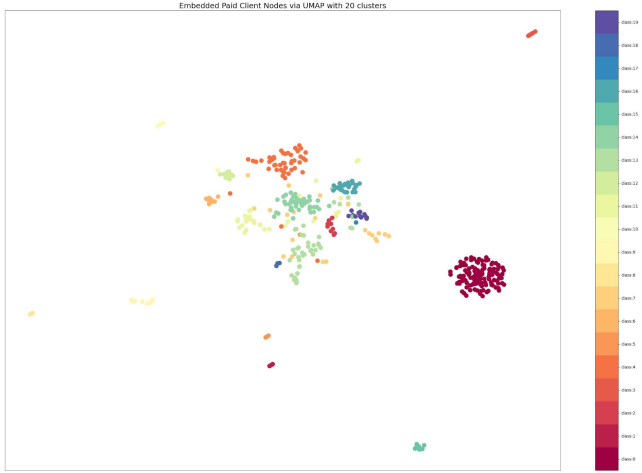
Clustering all paid clients of CBI based on their node vectors.

Identify Prospects

Look for potential clients that are similar to CBI paid client per cluster

Warmer call to prospect, sales opportunities

Similar Entity to CBI paid Client



Thanks!

Any feedbacks?

Find me on Slack, Zoom or Next Office Hour!

Gitlab Repo: 139/w2/HACKDAY-13-entitygraph