Design Knowledge Representation with Technology Semantic Network

Serhad Sarica & **Jianxi Luo**

saricas@ihpc.a-star.edu.sg luo@sutd.edu.sg

Motivation & Hypothesis

- A long and complex technical text may be hard to comprehend
- Topic mapping methods provide only high-level associations / disassociations.
- How do we map / associate concepts within the designs?
- A **semantic network trained on technical knowledge** may provide more meaningful network representations of technical texts.

Semantic Networks in ED

WordNet, ConceptNet vs. TechNet

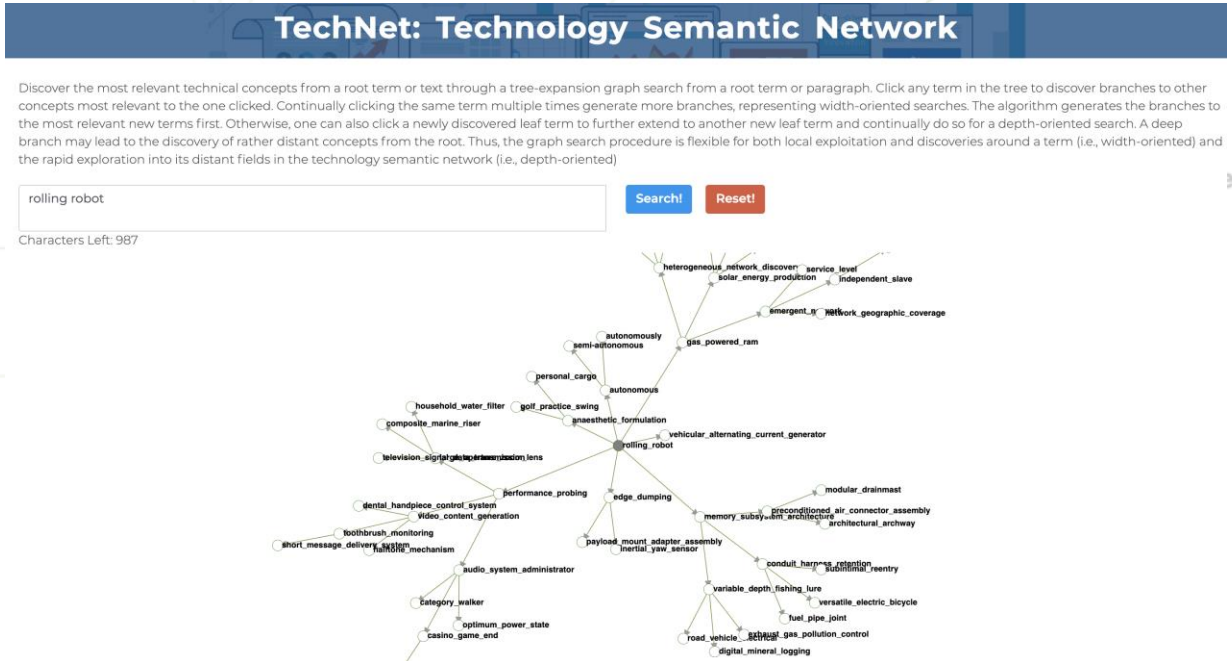
	Construction method	**Data Source**	**Relations**	**Quantitative Relations**
WordNet	Hand-built	Lexicographers	Synonym, hyponym, meronym, troponym, antonym	Shortest path, information content
ConceptNet	Unsupervised	DBPedia, Wiktionary, WordNet, OpenCyc, common webcrawl, common-sense, …	34 types of qualitative relations such as RelatedTo, FormOf, IsA, PartOf, etc.	Cosine similarity of word-embeddings trained on ConceptNet
TechNet	Unsupervised	USPTO Patent Database	-	Cosine similarity of word-embeddings trained on patent texts

TechNet
Technology Semantic Network

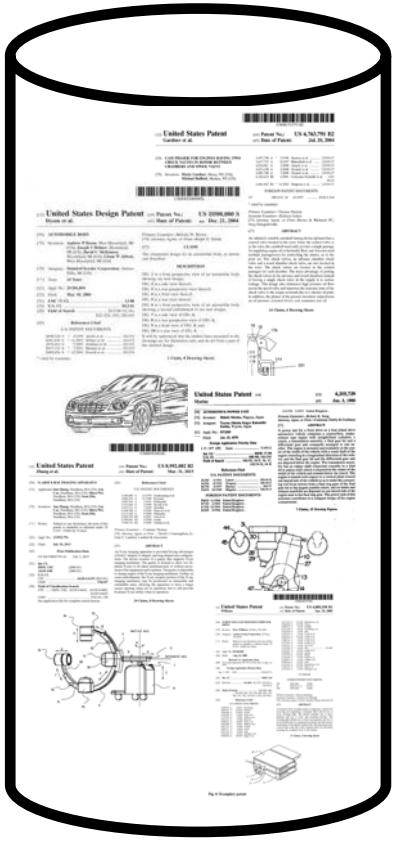
- More than 4 million terms
- Cosine similarity
- ~10^{12} possible relations
- Infrastructure: Public APIs

  www.tech-net.org

  www.github.com/SerhadS/TechNet



Data Source
USPTO Database
Granted: 1976 – Oct. 2017



~5.6M
Titles &
Abstracts

Term Extraction
NLP Tools

Punctuation Removal

Phrasing

Denoising

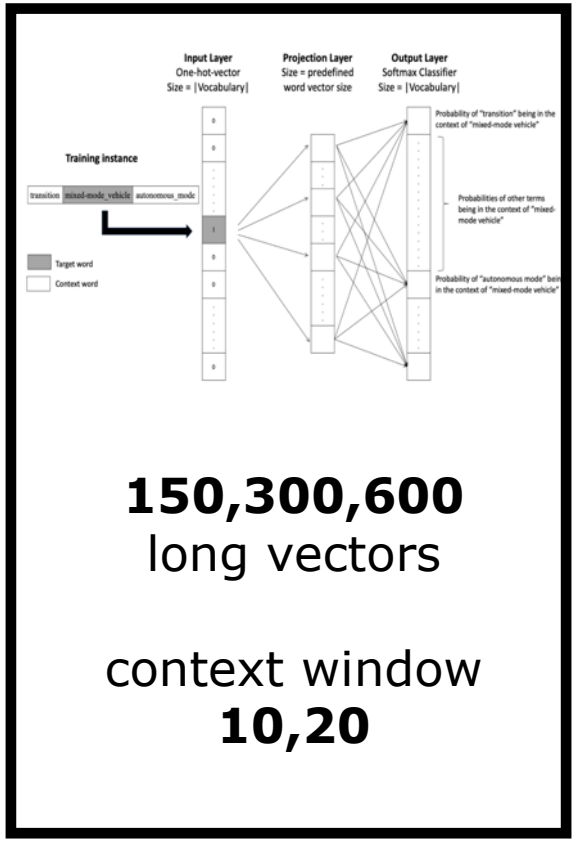
Lemmatization

Stopwords Removal

Tokenization

~26M
Tokenized
Sentences

Term Vorization
Language Models



150,300,600
long vectors

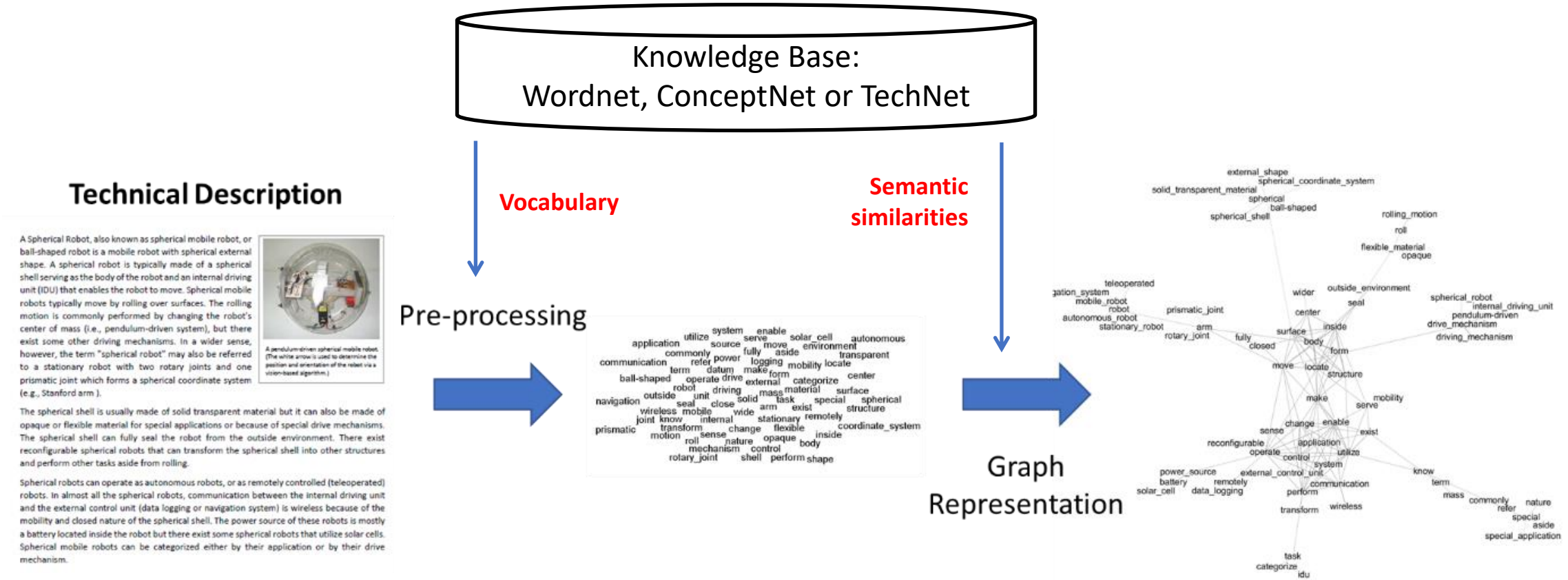
context window
10,20

Semantic Networks in ED

A Spherical Robot or ball-shaped robot is a mobile robot with spherical external shape .

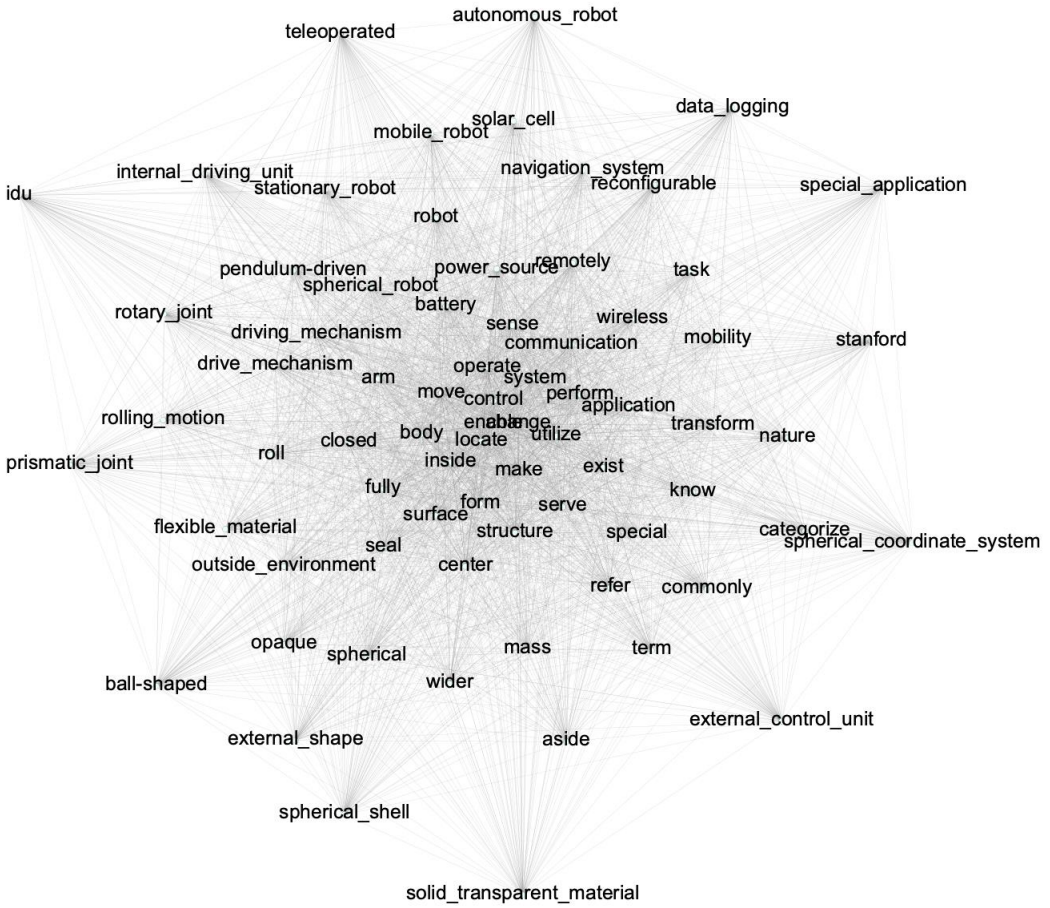
WordNet	**ConceptNet**	**TechNet**
Spherical	Spherical	Spherical robot
Robot	Robot	Ball-shaped
Ball-shaped	Ball-shaped	Robot
Mobile	Mobile	Mobile robot
External	External	Spherical
shape	shape	External shape

Overview of Methodology



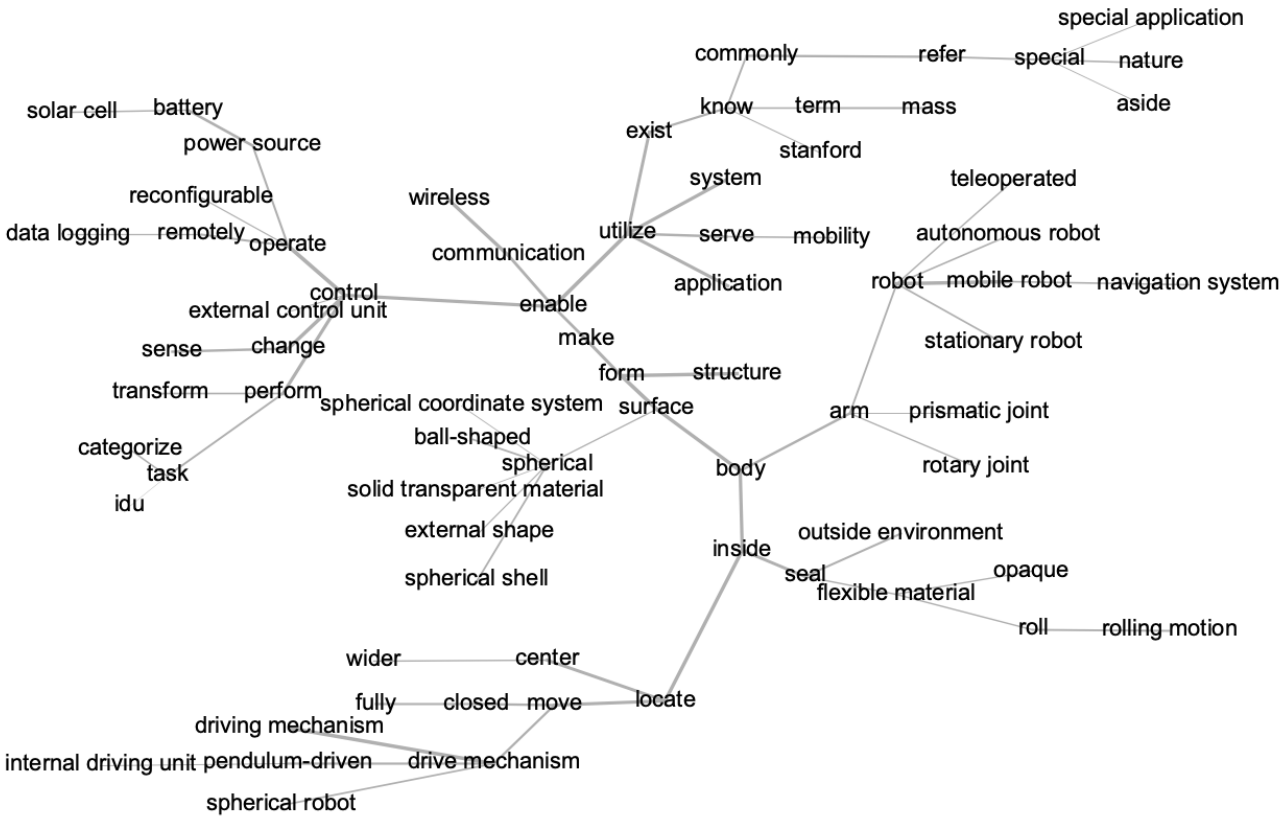
Graph Representation Approach

1) Link the terms by edges weighted by referring to terms' semantic similarities in corresponding knowledge base (i.e. either WordNet, ConceptNet or TechNet)



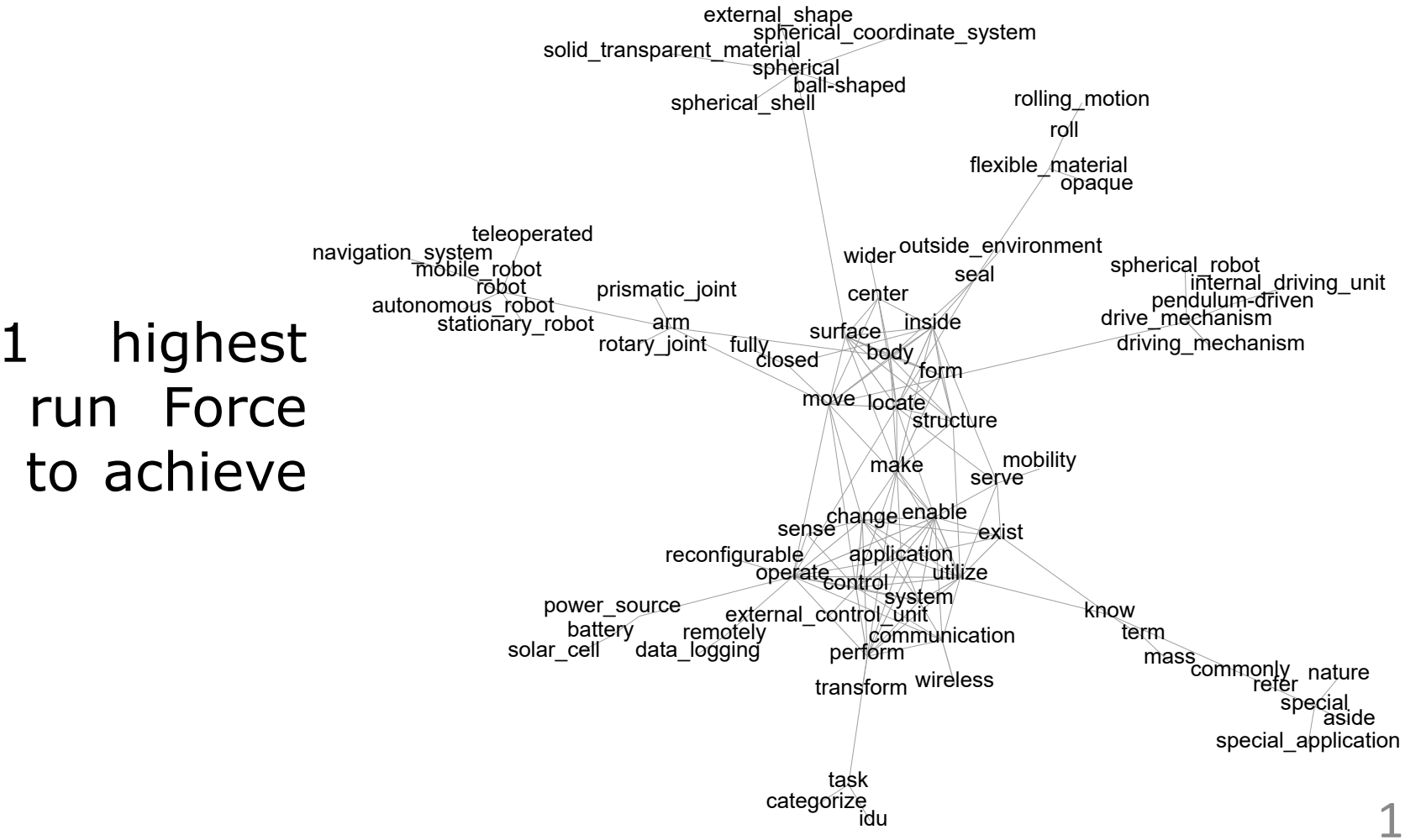
Graph Representation Approach

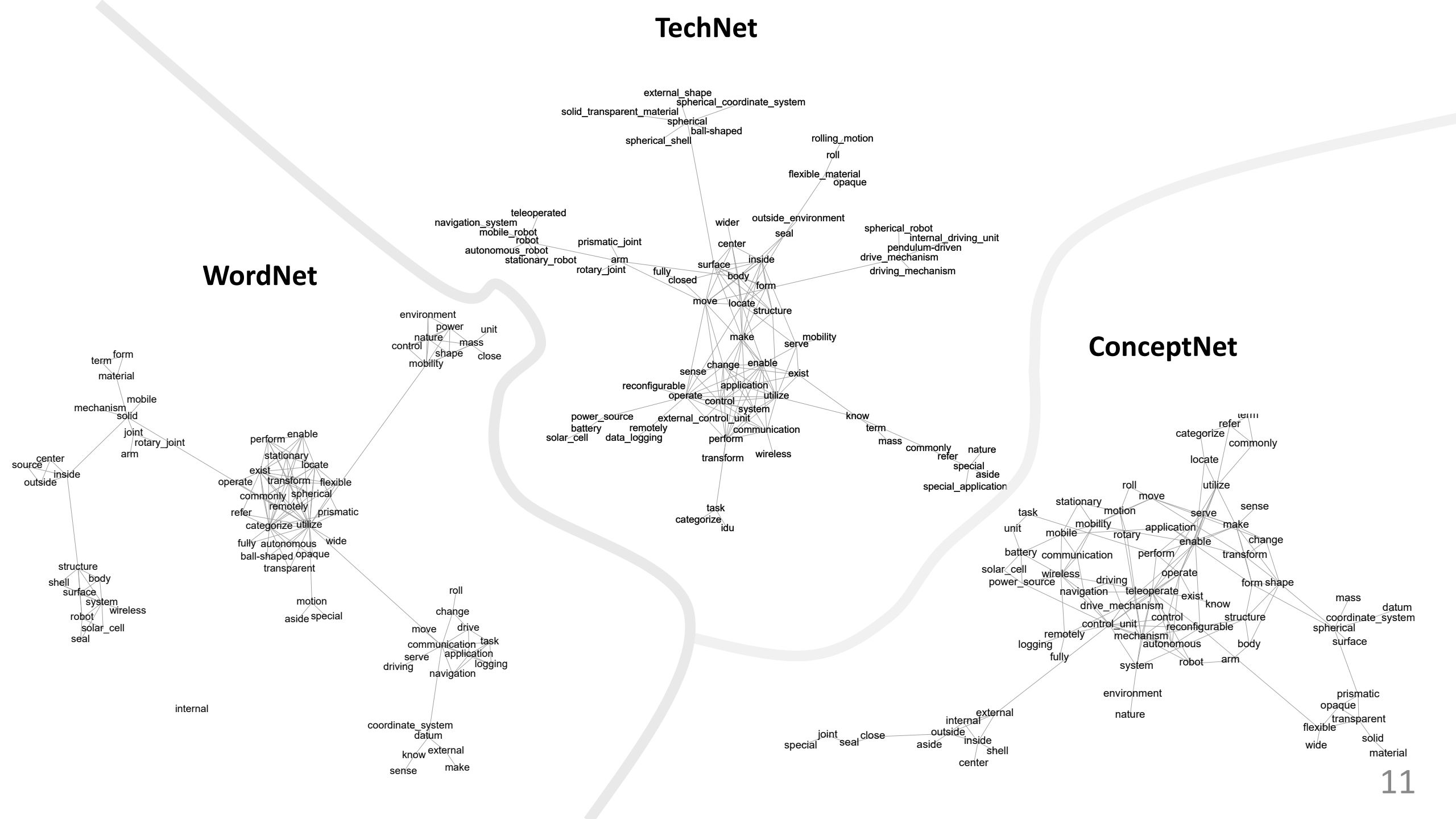
2) Generate maximum spanning tree of the graph



Graph Representation Approach

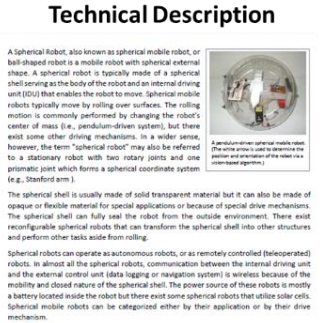
3) Add next N+1 highest weighted edges and run Force Atlas layout algorithm to achieve final visualization




TechNet
WordNet
ConceptNet
external_shape
spherical_coordinate_system
solid_transparent_material
spherical
ball-shaped
spherical_shell
rolling_motion
roll
flexible_material
opaque
teleoperated
navigation_system
mobile_robot
robot
autonomous_robot
stationary_robot
prismatic_joint
arm
rotary_joint
wider
outside_environment
seal
center
inside
surface
fully
closed
body
form
spherical_robot
internal_driving_unit
pendulum-driven
drive_mechanism
driving_mechanism
move
locate
structure
make
mobility
serve
change
enable
sense
exist
reconfigurable
application
operate
control
utilize
system
power_source
battery
solar_cell
data_logging
remotely
external_control_unit
communication
perform
transform
wireless
know
term
mass
commonly
nature
refer
special
aside
special_application
task
categorize
idu
environment
power
unit
nature
mass
control
shape
close
mobility
form
term
material
mobile
mechanism
solid
joint
rotary_joint
arm
center
source
inside
outside
structure
shell
body
surface
system
robot
wireless
solar_cell
seal
perform
enable
stationary
exist
locate
operate
transform
flexible
commonly
spherical
remotely
prismatic
refer
categorize
utilize
fully
autonomous
wide
ball-shaped
opaque
transparent
motion
aside
special
internal
roll
change
move
drive
communication
task
application
serve
logging
driving
navigation
coordinate_system
datum
external
know
sense
make
refer
categorize
commonly
locate
utilize
roll
move
stationary
motion
sense
task
serve
mobility
application
make
unit
mobile
rotary
change
enable
battery
communication
perform
transform
solar_cell
wireless
operate
power_source
driving
form shape
navigation
teleoperate
exist
know
drive_mechanism
mass
datum
control
structure
coordinate_system
control_unit
reconfigurable
spherical
remotely
mechanism
surface
logging
autonomous
body
fully
system
robot
arm
environment
nature
prismatic
opaque
transparent
flexible
solid
wide
material
external
internal
joint
close
outside
special
seal
aside
inside
shell
center

Survey Setup

1) Reading comprehension Summarization



>50 words summaries

2) Evaluation of individual visualizations

How well does this graph represent the **specific** design of "spherical robot"?

3) Redundant question asking the best visualization

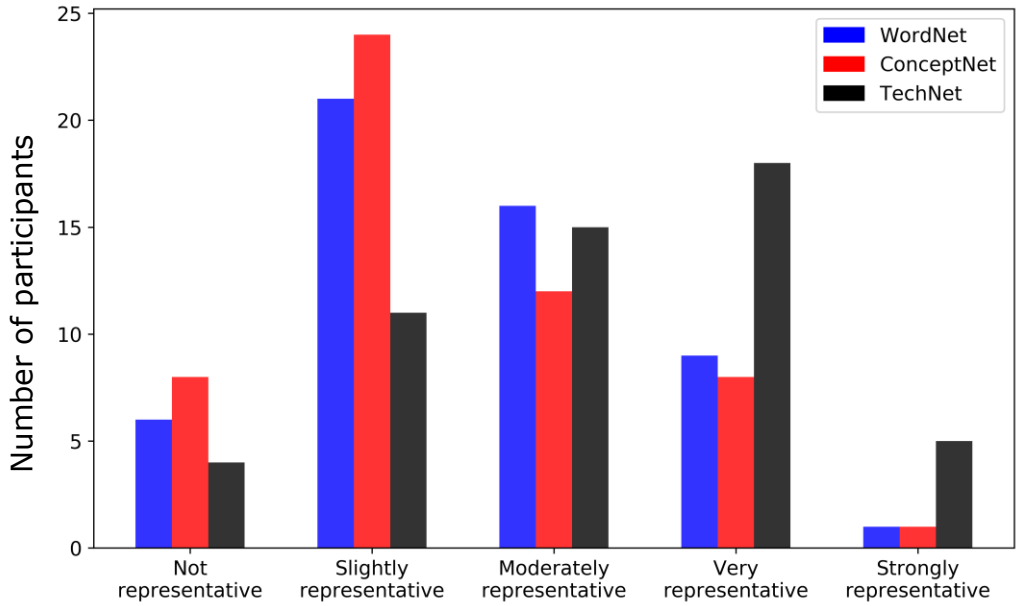
Which of the three graphs is the best representation of the **specific** design of "spherical robot"?

Participants

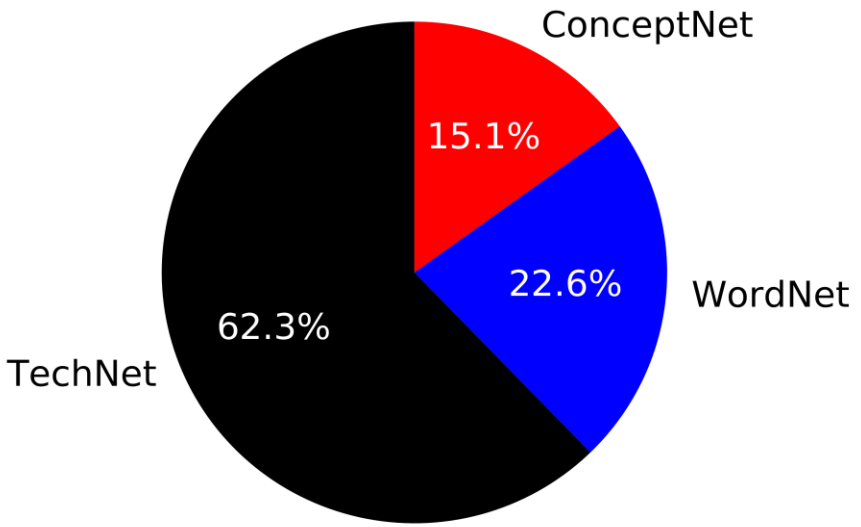
- 56 participants
- PhD students from SUTD and NTU, who have engineering backgrounds
- Conducted informal interviews with 25 of participants
- 3 participants removed since they admit they did not pay attention to the questions asked

Results

How well does this graph represent the **specific** design of "spherical robot"?



Which of the three graphs is the best representation of the **specific** design of "spherical robot"?



Participant comments

- TechNet can create distinct groups of highly relevant concepts

  *"The nodes of the graph include **Multi-Word Units** (MWUs - e.g., spherical coordinate system) that are **important for the reproduction of entities** in the text description. The entities that comprise **MWUs carry a specific meaning in this context**, which is lost when these are decomposed into single word units"*

  *"**Visually** it splits my attention into **a few key components** which I associate as different parts that can be used to **describe a spherical robot**"*

- Comments point out engineers' tendency to favor more modular and detailed representation

Findings and Limitations

- Knowledge base with a technical focus seems to support better visualizations of a technical text for technical eyes comparing to common-sense knowledge bases

- A first step to explore visual summarization techniques based on semantic networks and knowledge graphs
-
- Limitations:
  - Depends on specific graph filtering and visualization methods
  - Graphs are not the only possible visualization technique
  - There exists other potential knowledge bases


Thank you!