**VAST 2024 MC3 Data Description**

Data dictionary and nodes for MC3:

Graph Description:

* Directed multi-graph, allowing multiple edges between nodes
* 60520 nodes
* 75817 edges
* 4782 connected components
* Possible node types are: {Person, CEO, Organization, Company, FishingCompany, LogisticsCompany, NewsCompany, FinancialCompany, NGO)
* Possible edge types are: {Shareholdership, BeneficialOwnership, WorksFor, FamilyRelationship}
* The graph format is a JSON format generated by Python’s [network.node\_link\_data()](https://networkx.org/documentation/stable/reference/readwrite/generated/networkx.readwrite.json_graph.node_link_data.html) function. It can likewise be loaded to a networkx object using the corresponding [node\_link\_graph()](https://networkx.org/documentation/stable/reference/readwrite/generated/networkx.readwrite.json_graph.node_link_graph.html#networkx.readwrite.json_graph.node_link_graph) function. The root-level JSON object consists of graph-level properties specifying that it is directed and a multigraph, a “nodes” key which holds the list of nodes, and a “links” key which holds the list of edges.

**Node attributes:**

Person nodes (Entity.Person or Entity.Person.CEO)

* **type** – The type of node.
* **id** – The unique identifier of the node and the name of the person
* **dob** – The person’s date of birth
* **country** – The country associated with the entity.

Organization nodes (Entity.Organization, Entity.Organization.Company, Entity.Organization.FishingCompany, Entity.Organization.LogisticsCompany, Entity.Organization.NewsCompany, Entity.Organization.FinancialCompany, Entity.Organization.NGO)

* **type** – The type of node.
* **id** – The unique identifier of the node and the name of the organization
* **country** – The country associated with the entity.
* **HeadOfOrg –** The head of the organization. Optional.
* **ProductServices –** A list of products and services that the organization provides. Optional.
* **TradeDescription –** A short description of what the organization does. Optional.
* **PointOfContact –** The person the company has listed to contact for questions. Optional.
* **revenue –** The last reported annual revenue for the company in local currency; *all empty values have been set to 0.* Optional.

**Edge attributes:**

Event.Owns.BeneficialOwnership, Event.Owns.Shareholdership edges, Event.WorksFor edges:

* **type** – The type of edge.
* **start\_date** – The date at which the event began.
* **end\_date –** The date at which the event ended. This attribute is optional.

FamilyRelationship edges only have **type** attributes.

**Metadata attributes:**

Every single node and edge should have each of these fields.

* **\_last\_edited\_by** – The name of the user who last modified this node/edge
* **\_last\_edited\_date** – The last time this node/edge was modified
* **\_date\_added** – The day on which this node/edge was first added to the graph
* **\_raw\_source** – The source from which the information was originally obtained
* **\_algorithm** – The method in which the information was obtained (for this mini-challenge, either automatically imported from pre-existing databases or manually updated by FishEye analysts)