

ICEES+ KG API Endpoints

- **Exposures Provider** (Hao Xu, Kara Fecho, Steve Cox):
 - related_to: association between features
 - provided_by: the data source providing the node or edge
 - has_attribute: to describe attribute not in the Biolink model (e.g., P value)
 - has_count: counts of patients in the contingency table (see below)
 - has_chi_square, has_p_value, ...: Chi Square statistics and P values [other statistical models (e.g., GLM) and machine learning algorithms (e.g., random forest) are in development]
 - mapped_icees_feature: the icees feature that is mapped to this node or edge
 - in_icees_cohort: concepts of cohort, latitude/longitude [i.e., space], year [approaches to address other aspects of temporality are in development], table, context
 - APIs:
 - 2 * N associations KG:
 - input parameters:
 - query_options:
 - table : ICEES table
 - year : ICEES year
 - cohort_features: features for defining the cohort

[Will make table, year, cohort optional, with default being all available years of data for all available patients]
 - feature: a feature and operator and value for stratifying a cohort or examining the cohort with respect to relevant outcome measures (dependent variables)
 - maximum_p_value: ICEES maximum P value. The P value is calculated for each ICEES feature in the table, using 2 * N contingency table where the rows are stratified cohort levels and the columns are individual values of that feature. Any feature with P value greater than maximum p value is filtered out.
 - regex: filter target node name by regex
 - output:
 - a knowledge graph starting from cohort to features with P value less than the maximum_p_value

- url:
 - https://icees.renci.org:16340/apidocs/#/default/post_knowledge_graph
- One-hop KG:
 - input parameters:
 - query_options:
 - table : ICEES table
 - year : ICEES year
 - cohort_features: features for defining the cohort

[Will make table, year, cohort optional, with default being all available years of data for all available patients]
 - maximum_p_value: ICEES maximum P value. The P value is calculated for each ICEES feature in the table with the node type, using the pairwise contingency table where the rows are values of one feature and the columns are individual values of the other feature. Any feature with P value greater than maximum_p_value is filtered out.
 - regex: filter target node name by regex
 - query : node curie -[edge type]-> node type
 - output:
 - a knowledge_graph containing all one hop nodes with P value less than the maximum_p_value
 - https://icees.renci.org:16340/apidocs/#/default/post_knowledge_graph_one_hop
- Overlay KG:
 - input parameters:
 - query_options:
 - table: ICEES table
 - year: ICEES year
 - cohort_features: features for defining the cohort

[Will make table, year, cohort optional, with default being all available years of data for all available patients]
 - maximum_p_value: ICEES maximum P value. The P value is calculated for each ICEES feature in the table that maps to a node in the knowledge_graph, using the pairwise contingency table where the rows are values of one feature and the columns are individual values of that feature. Any feature with P

- value greater than maximum_p_value is filtered out.
 - regex: filter target node name by regex
 - knowledge_graph
- output:
 - a knowledge_graph with additional edges for P value less than the maximum_p_value
- url:
 - https://icees.renci.org:16340/apidocs/#/default/post_knowledge_graph_overlay

Example 1: [eventually, endpoint will accept SRI-normalized identifiers]

```
{
  "message": {
    "knowledge_graph": {
      "nodes": [
        {
          "node_id": "n00",
          "curie": "PUBCHEM:2083",
          "type": "drug"
        },
        {
          "node_id": "n01",
          "curie": "PUBCHEM:281",
          "type": "chemical_substance"
        }
      ],
      "edges": [
        {
          "id": "e00",
          "type": "association",
          "source_id": "n00",
          "target_id": "n01"
        }
      ]
    }
  }
}
```

Example 2:

```
{
  "message": {
    "knowledge_graph": {
      "nodes": [
        {
          "node_id": "n00",
          "curie": "PUBCHEM:5865",
          "type": "drug"
        },

```

```

        {
          "node_id": "n01",
          "curie": "MESH:D052638",
          "type": "chemical_substance"
        }
      ],
      "edges": [
        {
          "id": "e00",
          "type": "association",
          "source_id": "n00",
          "target_id": "n01"
        }
      ]
    }
  }
}

```

Example 3:

```

{
  "message": {
    "knowledge_graph": {
      "nodes": [
        {
          "node_id": "n00",
          "curie": "PUBCHEM:5865",
          "type": "drug"
        },
        {
          "node_id": "n01",
          "curie": "MONDO:0004766",
          "type": "disease"
        }
      ],
      "edges": [
        {
          "id": "e00",
          "type": "association",
          "source_id": "n00",
          "target_id": "n01"
        }
      ]
    }
  }
}

```