Metadata of MOBINS

The **MOBINS** comprises six datasets: *Transportation-[Seoul, Busan, Daegu, NYC]* and *Epidemic-[Korea, NYC]*. Figures 1 and 3 illustrate how the OD movements and time series are organized within the spatial networks of South Korea and New York. Additionally, we provide detailed descriptions of the metadata for each dataset. On our repository, we share the three components of each dataset in CSV format, along with metadata that is created by referring to the simple Croissant format [11].

Metadata of Transportation Datasets

Listing 1 presents the metadata for the *Transportation-Seoul* dataset, including CSV format files. Each file contains information about a single node or a node pair, which is abstracted for simplicity by describing only the i-th node. We omit the detailed description in Listings 2 and 3 because the CSV file structures are identical to Listing 1, differing only in the number of nodes, which is unique to each dataset. Listing 4 follows a similar structure, with the exception of the variable for node time-series features, where ridership replaces inflow and outflow.

```
 1  {
 2    "@type": "sc:Dataset",
 3    "name": "Transportation-Seoul",
 4    "description": "This dataset contains transportation data for Seoul, including spatial
        network data, node time series features, and origin-destination movements.",
 5    "license": "https://creativecommons.org/licenses/by-nc/4.0",
 6    "url": "https://drive.google.com/file/d/1MPkqpoeysY6tF-LUuya-bai5V0Fygx77/view?usp=
        sharing",
 7    "distribution": [
 8      {
 9        "@type": "cr:FileObject",
10        "@id": "SPATIAL_NETWORK.csv",
11        "name": "SPATIAL_NETWORK.csv",
12        "contentUrl": "Transportation-Seoul/SPATIAL_NETWORK.csv",
13        "encodingFormat": "text/csv"
14      },
15      {
16        "@type": "cr:FileObject",
17        "@id": "NODE_TIME_SERIES_FEATURES.csv",
18        "name": "NODE_TIME_SERIES_FEATURES.csv",
19        "contentUrl": "Transportation-Seoul/NODE_TIME_SERIES_FEATURES.csv",
20        "encodingFormat": "text/csv"
21      },
22      {
23        "@type": "cr:FileObject",
24        "@id": "OD_MOVEMENTS.csv",
25        "name": "OD_MOVEMENTS.csv",
26        "contentUrl": "Transportation-Seoul/OD_MOVEMENTS.csv",
27        "encodingFormat": "text/csv"
28      }
29    ],
30    "recordSet": [
31      {
32        "@type": "cr:RecordSet",
33        "name": "SPATIAL_NETWORK",
34        "description": "This table represents the spatial network of nodes in Seoul.",
35        "field": [
36          {
37            "@type": "cr:Field",
38            "name": "INDEX",
39            "description": "The index column containing node identifiers. ",
40            "dataType": "sc:Integer",
41            "references": {
42              "fileObject": { "@id": "SPATIAL_NETWORK.csv" },
43              "extract": {
44                "column": "INDEX"
45              }
46            }
47          },
```

---

[11] https://github.com/mlcommons/croissant

```json
 48          {
 49            "@type": "cr:Field",
 50            "name": "Ni",
 51            "description": "Columns representing spatial connections between all nodes and i
     -th node in the spatial network to describe an adjacency matrix ",
 52            "dataType": "sc:Integer",
 53            "references": {
 54              "fileObject": { "@id": "SPATIAL_NETWORK.csv" },
 55              "extract": {
 56                "column": "Ni"
 57              }
 58            }
 59          }
 60        ]
 61      },
 62      {
 63        "@type": "cr:RecordSet",
 64        "name": "NODE_TIME_SERIES_FEATURES",
 65        "description": "This table contains time series features for each node over multiple
     timestamps.",
 66        "field": [
 67          {
 68            "@type": "cr:Field",
 69            "name": "datetime",
 70            "description": "The datetime column representing timestamps (year, month, day,
     hour).",
 71            "dataType": "sc:DateTime",
 72            "references": {
 73              "fileObject": { "@id": "NODE_TIME_SERIES_FEATURES.csv" },
 74              "extract": {
 75                "column": "datetime"
 76              }
 77            }
 78          },
 79          {
 80            "@type": "cr:Field",
 81            "name": "Ni_INFLOW",
 82            "description": "Columns representing inflow time series for i-th node.",
 83            "dataType": "sc:Integer",
 84            "references": {
 85              "fileObject": { "@id": "NODE_TIME_SERIES_FEATURES.csv" },
 86              "extract": {
 87                "column": "Ni_INFLOW"
 88              }
 89            }
 90          },
 91          {
 92            "@type": "cr:Field",
 93            "name": "Ni_OUTFLOW",
 94            "description": "Columns representing outflow time series for i-th node.",
 95            "dataType": "sc:Integer",
 96            "references": {
 97              "fileObject": { "@id": "NODE_TIME_SERIES_FEATURES.csv" },
 98              "extract": {
 99                "column": "Ni_OUTFLOW"
100              }
101            }
102          }
103        ]
104      },
105      {
106        "@type": "cr:RecordSet",
107        "name": "OD_MOVEMENTS",
108        "description": "This table represents origin-destination movements between nodes.",
109        "field": [
```

```
110            {
111              "@type": "cr:Field",
112              "name": "Ni_Nj",
113              "description": "Columns representing movements from i-th node to j-th node.",
114              "dataType": "sc:Integer",
115              "references": {
116                "fileObject": { "@id": "OD_MOVEMENTS.csv" },
117                "extract": {
118                  "column": "Ni_Nj"
119                }
120              }
121            }
122          ]
123        }
124    ]
125 }
```

Listing 1: Metadata of Transportation-Seoul.

```
 1 {
 2   "@type": "sc:Dataset",
 3   "name": "Transportation-Busan",
 4   "description": "This dataset contains transportation data for Busan, including spatial
      network data, node time series features, and origin-destination movements.",
 5   "license": "https://creativecommons.org/licenses/by-nc/4.0",
 6   "url": "https://drive.google.com/file/d/1PU9aqp7kc4UzuFOAqLPkraM5MbmtyodM/view?usp=
      sharing",
 7   "distribution": [
 8     {
 9       "@type": "cr:FileObject",
10       "@id": "SPATIAL_NETWORK.csv",
11       "name": "SPATIAL_NETWORK.csv",
12       "contentUrl": "Transportation-Busan/SPATIAL_NETWORK.csv",
13       "encodingFormat": "text/csv"
14     },
15     {
16       "@type": "cr:FileObject",
17       "@id": "NODE_TIME_SERIES_FEATURES.csv",
18       "name": "NODE_TIME_SERIES_FEATURES.csv",
19       "contentUrl": "Transportation-Busan/NODE_TIME_SERIES_FEATURES.csv",
20       "encodingFormat": "text/csv"
21     },
22     {
23       "@type": "cr:FileObject",
24       "@id": "OD_MOVEMENTS.csv",
25       "name": "OD_MOVEMENTS.csv",
26       "contentUrl": "Transportation-Busan/OD_MOVEMENTS.csv",
27       "encodingFormat": "text/csv"
28     }
29   ]
30 }
```

Listing 2: Metadata of Transportation-Busan.

```
 1 {
 2   "@type": "sc:Dataset",
 3   "name": "Transportation-Daegu",
 4   "description": "This dataset contains transportation data for Daegu, including spatial
      network data, node time series features, and origin-destination movements.",
 5   "license": "https://creativecommons.org/licenses/by-nc/4.0",
 6   "url": "https://drive.google.com/file/d/1ggcUHlpT7QCYB1jwlouhZLgM2FzDnuKU/view?usp=
      sharing",
 7   "distribution": [
 8     {
 9       "@type": "cr:FileObject",
10       "@id": "SPATIAL_NETWORK.csv",
11       "name": "SPATIAL_NETWORK.csv",
```

```
12        "contentUrl": "Transportation-Daegu/SPATIAL_NETWORK.csv",
13        "encodingFormat": "text/csv"
14      },
15      {
16        "@type": "cr:FileObject",
17        "@id": "NODE_TIME_SERIES_FEATURES.csv",
18        "name": "NODE_TIME_SERIES_FEATURES.csv",
19        "contentUrl": "Transportation-Daegu/NODE_TIME_SERIES_FEATURES.csv",
20        "encodingFormat": "text/csv"
21      },
22      {
23        "@type": "cr:FileObject",
24        "@id": "OD_MOVEMENTS.csv",
25        "name": "OD_MOVEMENTS.csv",
26        "contentUrl": "Transportation-Daegu/OD_MOVEMENTS.csv",
27        "encodingFormat": "text/csv"
28      }
29    ]
30 }
```

Listing 3: Metadata of Transportation-Daegu.

```
1  {
2    "@type": "sc:Dataset",
3    "name": "Transportation-NYC",
4    "description": "This dataset contains transportation data for NYC, including spatial
       network data, node time series features, and origin-destination movements.",
5    "license": "https://creativecommons.org/licenses/by-nc/4.0",
6    "url": "https://drive.google.com/file/d/1R97wE9_v2WVAS1MAWKUbZWOhPQNQYCnV/view?usp=
       sharing",
7    "distribution": [
8      {
9        "@type": "cr:FileObject",
10       "@id": "SPATIAL_NETWORK.csv",
11       "name": "SPATIAL_NETWORK.csv",
12       "contentUrl": "Transportation-NYC/SPATIAL_NETWORK.csv",
13       "encodingFormat": "text/csv"
14     },
15     {
16       "@type": "cr:FileObject",
17       "@id": "NODE_TIME_SERIES_FEATURES.csv",
18       "name": "NODE_TIME_SERIES_FEATURES.csv",
19       "contentUrl": "Transportation-NYC/NODE_TIME_SERIES_FEATURES.csv",
20       "encodingFormat": "text/csv"
21     },
22     {
23       "@type": "cr:FileObject",
24       "@id": "OD_MOVEMENTS.csv",
25       "name": "OD_MOVEMENTS.csv",
26       "contentUrl": "Transportation-NYC/OD_MOVEMENTS.csv",
27       "encodingFormat": "text/csv"
28     }
29   ],
30   "recordSet": [
31     {
32       "@type": "cr:RecordSet",
33       "name": "SPATIAL_NETWORK",
34       "description": "This table represents the spatial network of nodes in Seoul.",
35       "field": [
36         {
37           "@type": "cr:Field",
38           "name": "INDEX",
39           "description": "The index column containing node identifiers.",
40           "dataType": "sc:Integer",
41           "references": {
42             "fileObject": { "@id": "SPATIAL_NETWORK.csv" },
```

```
43          "extract": {
44            "column": "INDEX"
45          }
46        }
47      },
48      {
49        "@type": "cr:Field",
50        "name": "Ni",
51        "description": "Columns representing spatial connections between all nodes and i
      -th node in the spatial network to describe an adjacency matrix ",
52        "dataType": "sc:Integer",
53        "references": {
54          "fileObject": { "@id": "SPATIAL_NETWORK.csv" },
55          "extract": {
56            "column": "Ni"
57          }
58        }
59      }
60      ]
61    },
62    {
63      "@type": "cr:RecordSet",
64      "name": "NODE_TIME_SERIES_FEATURES",
65      "description": "This table contains time series features for each node over multiple
      timestamps.",
66      "field": [
67        {
68          "@type": "cr:Field",
69          "name": "datetime",
70          "description": "The datetime column representing timestamps (year, month, day,
      hour).",
71          "dataType": "sc:DateTime",
72          "references": {
73            "fileObject": { "@id": "NODE_TIME_SERIES_FEATURES.csv" },
74            "extract": {
75              "column": "datetime"
76            }
77          }
78        },
79        {
80          "@type": "cr:Field",
81          "name": "Ni_RIDERSHIP",
82          "description": "Columns representing ridership time series for i-th node.",
83          "dataType": "sc:Integer",
84          "references": {
85            "fileObject": { "@id": "NODE_TIME_SERIES_FEATURES.csv" },
86            "extract": {
87              "column": "Ni_RIDERSHIP"
88            }
89          }
90        }
91      ]
92    },
93    {
94      "@type": "cr:RecordSet",
95      "name": "OD_MOVEMENTS",
96      "description": "This table represents origin-destination movements between nodes.",
97      "field": [
98        {
99          "@type": "cr:Field",
100         "name": "Ni_Nj",
101         "description": "Columns representing movements from i-th node to j-th node.",
102         "dataType": "sc:Integer",
103         "references": {
104           "fileObject": { "@id": "OD_MOVEMENTS.csv" },
```

```
105        "extract": {
106          "column": "Ni_Nj"
107        }
108      }
109    }
110   ]
111  }
112  ]
113 }
```

<div align="center">Listing 4: Metadata of Transportation-NYC.</div>

Metadata of Epidemic Datasets

The Listings 5 and 6 present the metadata for the *Epidemic-[Korea, NYC]* datasets, respectively. Both datasets share a consistent structure in terms of node time-series features, OD movements, and spatial networks. However, *Epidemic-Korea* consists of 16 nodes while *Epidemic-NYC* consists of 5 nodes, resulting in a different number of columns in their respective CSV files. Nevertheless, they are datasets in the same context. Additionally, they differ in their licensing conditions.

```
1  {
2    "@type": "sc:Dataset",
3    "name": "Epidemic-Korea",
4    "description": "This dataset contains epidemic data for South Korea, including spatial
        network data, node time series features, and origin-destination movements.",
5    "license": "https://creativecommons.org/licenses/by-nc-nd/4.0/",
6    "url": "https://drive.google.com/file/d/1G0P4-HdRoU6X2p18VmOaEMJVPIUEEhiE/view?usp=
        sharing",
7    "distribution": [
8      {
9        "@type": "cr:FileObject",
10       "@id": "SPATIAL_NETWORK.csv",
11       "name": "SPATIAL_NETWORK.csv",
12       "contentUrl": "Epidemic-Korea/SPATIAL_NETWORK.csv",
13       "encodingFormat": "text/csv"
14     },
15     {
16       "@type": "cr:FileObject",
17       "@id": "NODE_TIME_SERIES_FEATURES.csv",
18       "name": "NODE_TIME_SERIES_FEATURES.csv",
19       "contentUrl": "Epidemic-Korea/NODE_TIME_SERIES_FEATURES.csv",
20       "encodingFormat": "text/csv"
21     },
22     {
23       "@type": "cr:FileObject",
24       "@id": "OD_MOVEMENTS.csv",
25       "name": "OD_MOVEMENTS.csv",
26       "contentUrl": "Epidemic-Korea/OD_MOVEMENTS.csv",
27       "encodingFormat": "text/csv"
28     }
29   ],
30   "recordSet": [
31     {
32       "@type": "cr:RecordSet",
33       "name": "SPATIAL_NETWORK",
34       "description": "This table represents the spatial network of nodes in Seoul.",
35       "field": [
36         {
37           "@type": "cr:Field",
38           "name": "INDEX",
39           "description": "The index column containing node identifiers.",
40           "dataType": "sc:Integer",
41           "references": {
42             "fileObject": { "@id": "SPATIAL_NETWORK.csv" },
43             "extract": {
44               "column": "INDEX"
45             }
```

```
 46                    }
 47                 },
 48                 {
 49                    "@type": "cr:Field",
 50                    "name": "Ni",
 51                    "description": "Columns representing spatial connections between all nodes and i
       -th node in the spatial network to describe an adjacency matrix ",
 52                    "dataType": "sc:Integer",
 53                    "references": {
 54                       "fileObject": { "@id": "SPATIAL_NETWORK.csv" },
 55                       "extract": {
 56                          "column": "Ni"
 57                       }
 58                    }
 59                 }
 60              ]
 61           },
 62           {
 63              "@type": "cr:RecordSet",
 64              "name": "NODE_TIME_SERIES_FEATURES",
 65              "description": "This table contains time series features for each node over multiple
       timestamps.",
 66              "field": [
 67                 {
 68                    "@type": "cr:Field",
 69                    "name": "datetime",
 70                    "description": "The datetime column representing timestamps (year, month, day)
       .",
 71                    "dataType": "sc:DateTime",
 72                    "references": {
 73                       "fileObject": { "@id": "NODE_TIME_SERIES_FEATURES.csv" },
 74                       "extract": {
 75                          "column": "datetime"
 76                       }
 77                    }
 78                 },
 79                 {
 80                    "@type": "cr:Field",
 81                    "name": "Ni_INFECTION",
 82                    "description": "Columns representing infection cases time series for i-th node
       .",
 83                    "dataType": "sc:Integer",
 84                    "references": {
 85                       "fileObject": { "@id": "NODE_TIME_SERIES_FEATURES.csv" },
 86                       "extract": {
 87                          "column": "Ni_INFECTION"
 88                       }
 89                    }
 90                 }
 91              ]
 92           },
 93           {
 94              "@type": "cr:RecordSet",
 95              "name": "OD_MOVEMENTS",
 96              "description": "This table represents origin-destination movements between nodes.",
 97              "field": [
 98                 {
 99                    "@type": "cr:Field",
100                    "name": "Ni_Nj",
101                    "description": "Columns representing movements from i-th node to j-th node.",
102                    "dataType": "sc:Integer",
103                    "references": {
104                       "fileObject": { "@id": "OD_MOVEMENTS.csv" },
105                       "extract": {
106                          "column": "Ni_Nj"
```

```
107              }
108            }
109          }
110        ]
111      }
112    ]
113  }
```

Listing 5: Metadata of Epidemic-Korea.

```
 1  {
 2    "@type": "sc:Dataset",
 3    "name": "Epidemic-NYC",
 4    "description": "This dataset contains epidemic data for NYC, including spatial network
          data, node time series features, and origin-destination movements.",
 5    "license": "https://creativecommons.org/licenses/by-nc/4.0/",
 6    "url": "https://drive.google.com/file/d/1G0P4-HdRoU6X2p18VmOaEMJVPIUEEhiE/view?usp=
          sharing",
 7    "distribution": [
 8      {
 9        "@type": "cr:FileObject",
10        "@id": "SPATIAL_NETWORK.csv",
11        "name": "SPATIAL_NETWORK.csv",
12        "contentUrl": "Epidemic-NYC/SPATIAL_NETWORK.csv",
13        "encodingFormat": "text/csv"
14      },
15      {
16        "@type": "cr:FileObject",
17        "@id": "NODE_TIME_SERIES_FEATURES.csv",
18        "name": "NODE_TIME_SERIES_FEATURES.csv",
19        "contentUrl": "Epidemic-NYC/NODE_TIME_SERIES_FEATURES.csv",
20        "encodingFormat": "text/csv"
21      },
22      {
23        "@type": "cr:FileObject",
24        "@id": "OD_MOVEMENTS.csv",
25        "name": "OD_MOVEMENTS.csv",
26        "contentUrl": "Epidemic-NYC/OD_MOVEMENTS.csv",
27        "encodingFormat": "text/csv"
28      }
29    ]
30  }
```

Listing 6: Metadata of Epidemic-NYC.