



# 知识疫图

## — 新冠疫情地图可视化的技术实践

殷 达

清华大学计算机系研究生

主办： AI Time

智谱.AI

AI Time是一群关注人工智能发展，并有思想情怀的青年人创办的圈子。AI Time旨在发扬科学思辨精神，邀请各界人士对人工智能理论、算法、场景、应用的本质问题进行探索，加强思想碰撞，打造成为北京乃至全国知识分享的聚集地。



## 殷达

- 清华大学计算机系硕士二年级
- 清华大学AMiner团队，知识工程实验室，师从唐杰教授
- 研究领域包括学术知识图谱中的数据挖掘、自然语言处理、图结构与神经网络
- 曾赴CMU访问学习
- 主导研发AMiner平台中的论文溯源树（MRT）等功能



# 知识疫图专栏简介

## \* 项目内容



# 知识疫图专栏简介

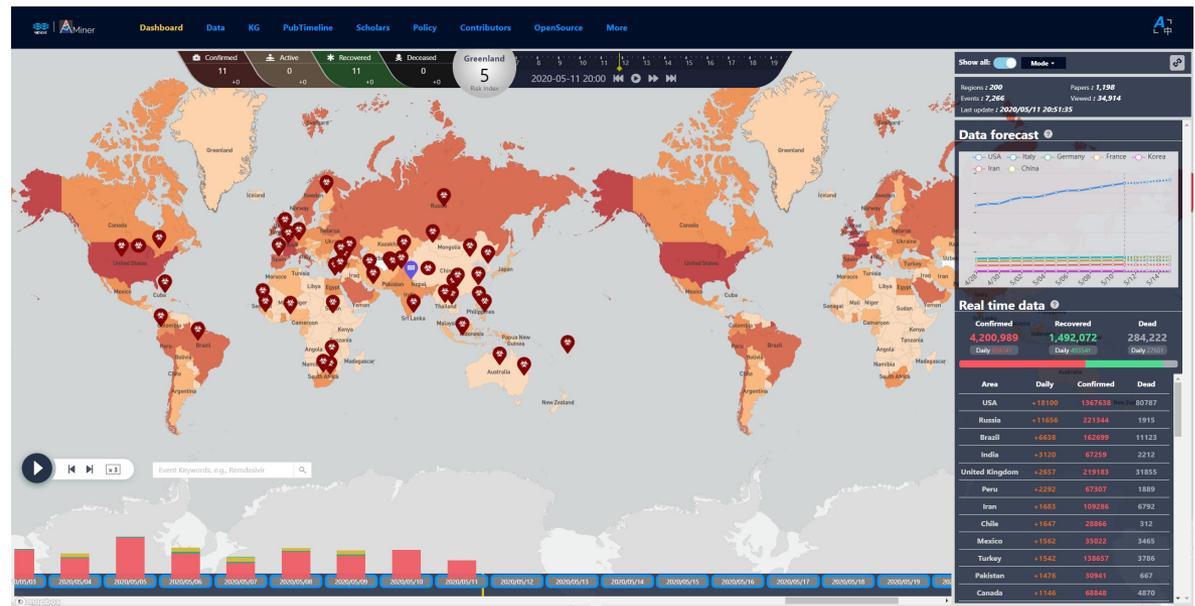


## \* COVID-19 Dashboard

- \* 疫情数据
- \* 趋势预测
- \* 风险评估
- \* 热点事件

## \* COVID-19 Map

- \* 数据处理
- \* 地图展示



# 疫情数据的采集与整合



爬虫采集



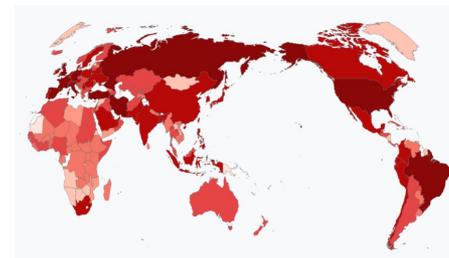
mongoDB

监控

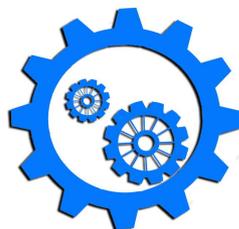


整合程序

整理对齐

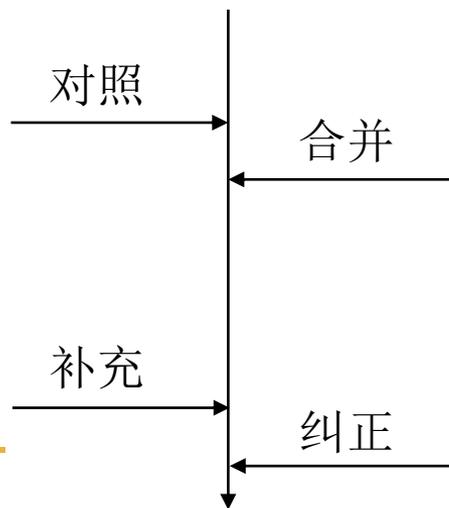


# 疫情数据的采集与整合

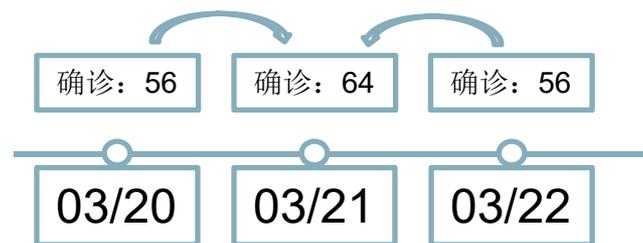


整合程序

原名	映射
Macau	Macao
British	United Kingdom
Lao People's Democratic Republic	Laos



源	优先级
Johns Hopkins	高
丁香园	低
.....	.....



# 疫情数据的采集与整合



```
{  
  "China": {  
    "2020-05-11": {  
      "Confirmed": 84450,  
      "Cured": 79539,  
      .....  
    }  
    .....  
  }  
}
```

→  
数据压缩

```
{  
  "China": {  
    "begin": "2020-01-20",  
    "data": [  
      .....  
      [84450, 267, 79539, ... ]  
    ]  
  }  
}
```

↓  
启用gzip传输

# 地理数据的使用

## \* 格式

- \* Geojson
- \* Topojson
- \* Shapefile
- \* ...

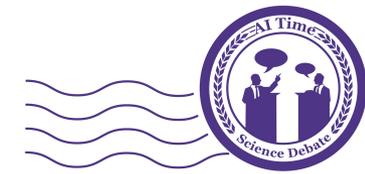
```
{ "type": "FeatureCollection",
  "features": [
    { "type": "Feature",
      "geometry": { "type": "Point", "coordinates": [102.0, 0.5] },
      "properties": { "prop0": "value0" }
    },
    { "type": "Feature",
      "geometry": {
        "type": "LineString",
        "coordinates": [
          [102.0, 0.0], [103.0, 1.0], [104.0, 0.0], [105.0, 1.0]
        ]
      },
      "properties": {
        "prop0": "value0",
        "prop1": 0.0
      }
    },
    { "type": "Feature",
      "geometry": {
        "type": "Polygon",
        "coordinates": [
          [ [100.0, 0.0], [101.0, 0.0], [101.0, 1.0],
            [100.0, 1.0], [100.0, 0.0] ]
        ]
      },
      "properties": {
        "prop0": "value0",
        "prop1": { "this": "that" }
      }
    }
  ]
}
```



# 地理数据的使用

- \* 格式
  - \* Geojson
  - \* Topojson
  - \* Shapefile
  - \* ...

```
{  
  "type": "Topology",  
  "transform": {  
    "scale": [  
      0.01671127518500191,  
      0.009729487698321453  
    ],  
    "translate": [  
      -179.942499,  
      -85.60903799999998  
    ]  
  },  
  "objects": {  
    "国家": {  
      "type": "GeometryCollection",  
      "geometries": [  
        {  
          "arcs": [  
            {  
              0,  
              1,  
              2,  
              3,  
              4,  
              5  
            }  
          ],  
          "type": "Polygon",  
          "properties": {"name": "Afghanistan"}  
        },  
        {  
          {
```



# 地理数据的使用

- \* 格式
  - \* Geojson
  - \* Topojson
  - \* Shapefile
  - \* ...



# 地理数据的使用



- \* 数据源
  - \* 阿里云
  - \* Natural Earth
  - \* .....
- \* 地区ID重组
  - \* Beijing – China|Beijing – 北京市
  - \* 繁体简体转换
  - \* 中英文对照



Free vector and raster map data @ [naturalearthdata.com](http://naturalearthdata.com)



# 地理数据的使用



## \* 数据修正

### \* 中国地区的修正

- \* 九段线
- \* 南海诸岛
- \* 台湾省
- \* 藏南地区等

### \* 其他世界地区的数据调整

- \* 意大利、法国等国家：国家 - 大区 - 省
- \* 英国：苏格兰、北爱尔兰、英格兰郡级



# 地理数据的使用



## \* 数据修正

### \* 中国地区的修正

- \* 九段线
- \* 南海诸岛
- \* 台湾省
- \* 藏南地区等

### \* 其他世界地区的数据调整

- \* 意大利、法国等国家：国家 - 大区 - 省
- \* 英国：苏格兰、北爱尔兰、英格兰郡级



# 地图绘制



## \* 前端渲染

- \* SVG绘制
- \* 国内的地图组件：百度地图、高德地图等
- \* 国外的地图组件：Google Map、Mapbox



**百度地图**  
开放平台  
lbsyun.baidu.com



**高德地图**  
amap.com

# 地图绘制



## \* 前端渲染

- \* SVG绘制
- \* 国内的地图组件：百度地图、高德地图等
- \* 国外的地图组件：Google Map、Mapbox



Google Maps for Work



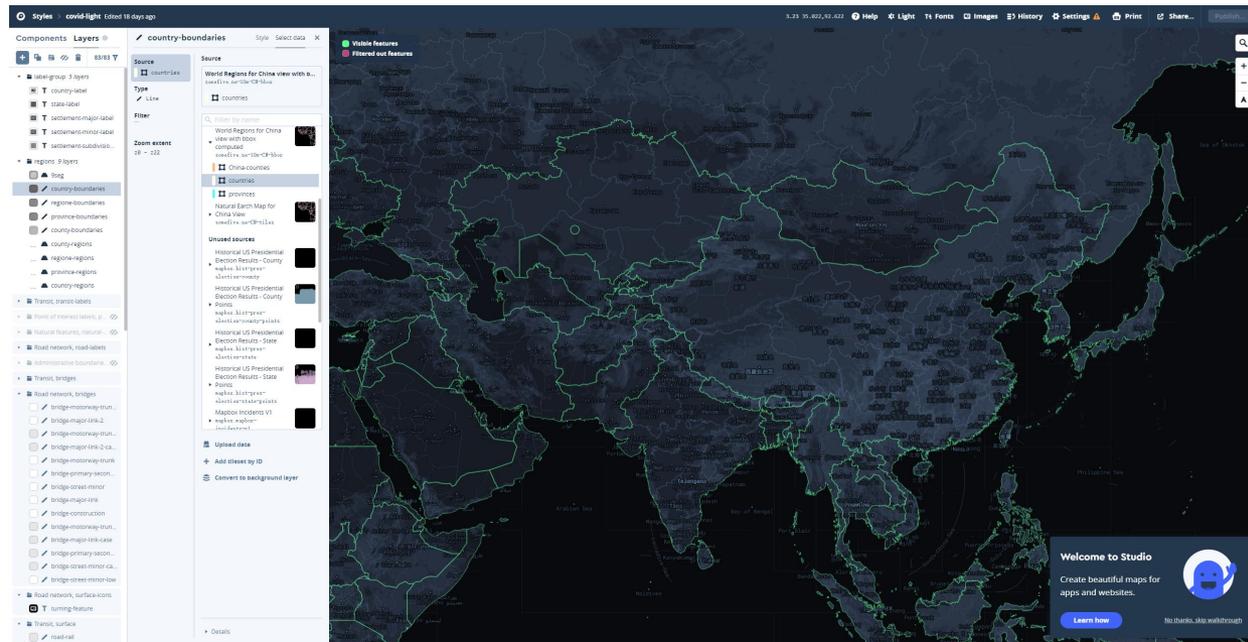


# 地图绘制



## \* 前端渲染

- \* Mapbox Studio
- \* 地图数据到区块数据





# 地图绘制

- \* 前端渲染
  - \* Source
    - \* Feature
  - \* Layer
    - \* Render Rules

中国地理数据

```
{ "type": "FeatureCollection", "features": [ { "type": "Feature", "properties": { "adcode": "110000", "name": "北京市", "center": [116.405285, 39.904989], "centroid": [116.419889, 40.189911], ...
```

边界层

渲染规则：根据Feature State中的stroke\_opacity决定透明度

计算颜色及透明度

设置中国的Feature State

中国疫情数据

填充层

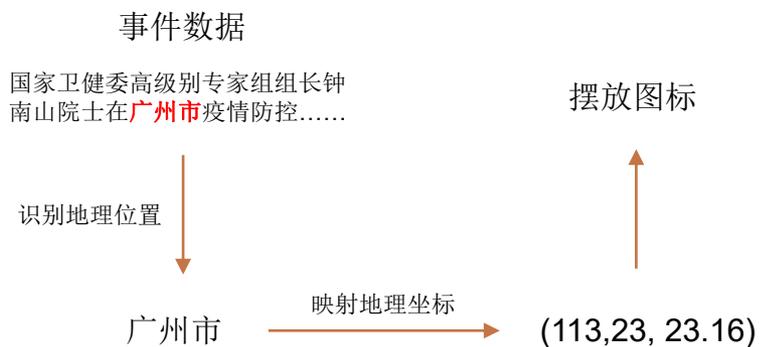
渲染规则：根据Feature State中的color决定颜色



# 地图绘制



- \* 前端渲染
  - \* 事件



# COVID Dashboard开源项目



\* [https://github.com/AMinerOpen/covid\\_dashboard](https://github.com/AMinerOpen/covid_dashboard)

The screenshot shows the GitHub repository page for 'AMinerOpen / covid\_dashboard'. The repository is currently on the 'master' branch. It has 8 commits, 2 branches, 0 packages, 0 releases, and 0 contributors. The repository is licensed under MIT. The file list includes:

File	Commit	Time
dujizhong api.md	7026a50	3 days ago
config	init	7 days ago
public	seo	6 days ago
scripts	init	7 days ago
src	header	3 days ago
.env	env	7 days ago
.env.development	env	7 days ago
.env.production	env	7 days ago

# 谢谢观看

后续分享更精彩，敬请期待！