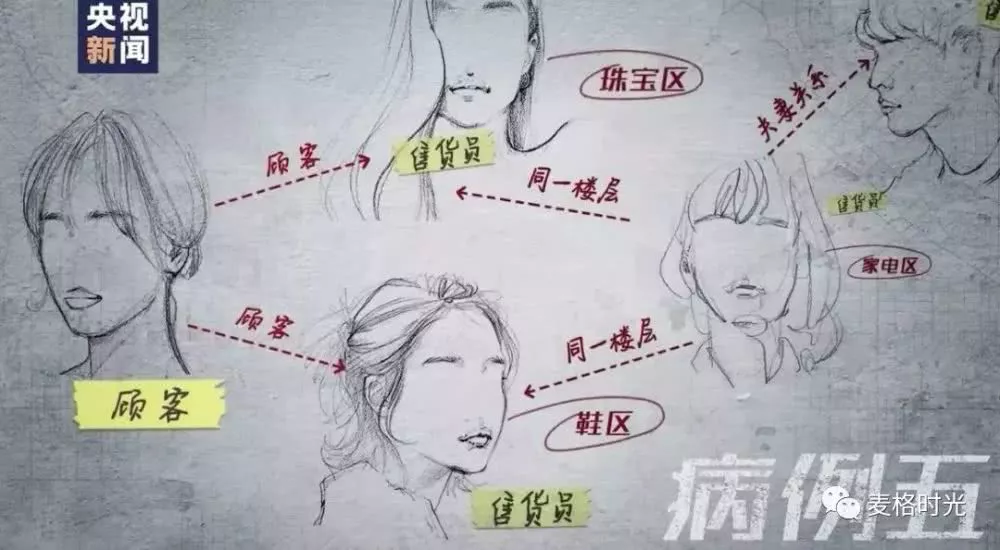
**#.新冠肺炎流行病学知识图谱**

**#.1简介以及潜在应用**

新冠肺炎流行病学知识图谱包含了流行病学知识图谱Schema和基于此Schema定义的新冠肺炎资源实例。其中流行病学知识图谱Schema重点刻画流行病学的基本概念、流行病学调查等内容，未包含“流行病学研究方法”、“预防与控制策略”、“临床治疗与预后”等内容，需要在后续版本中与其他新冠肺炎知识图谱集成或连接。新冠肺炎资源实例包含了发生在2019年~2020年2月期间，COVID-19疫情中确诊患者的流行病学调查数据，这些数据来自中各省市卫健委公布的个案流调信息。例如：患者基本情况、患者活动事件（含时间和地点、干了什么等）、及患者社会关系信息、当前现状、疾病传播路径等。

流行病学调查是一个很复杂和繁琐的事情，但它对于帮助人们了解疾病，制定有效的防治与控制策略非常重要。正如上海市卫健委主任邬惊雷所说：“通过流行病学调查，明确每一个患者的感染路径，对于未来的疫情防控能起到很大的作用。对于公众而言，也可以知道用什么方式避免感染，感染之后通过什么路径去解决问题。”

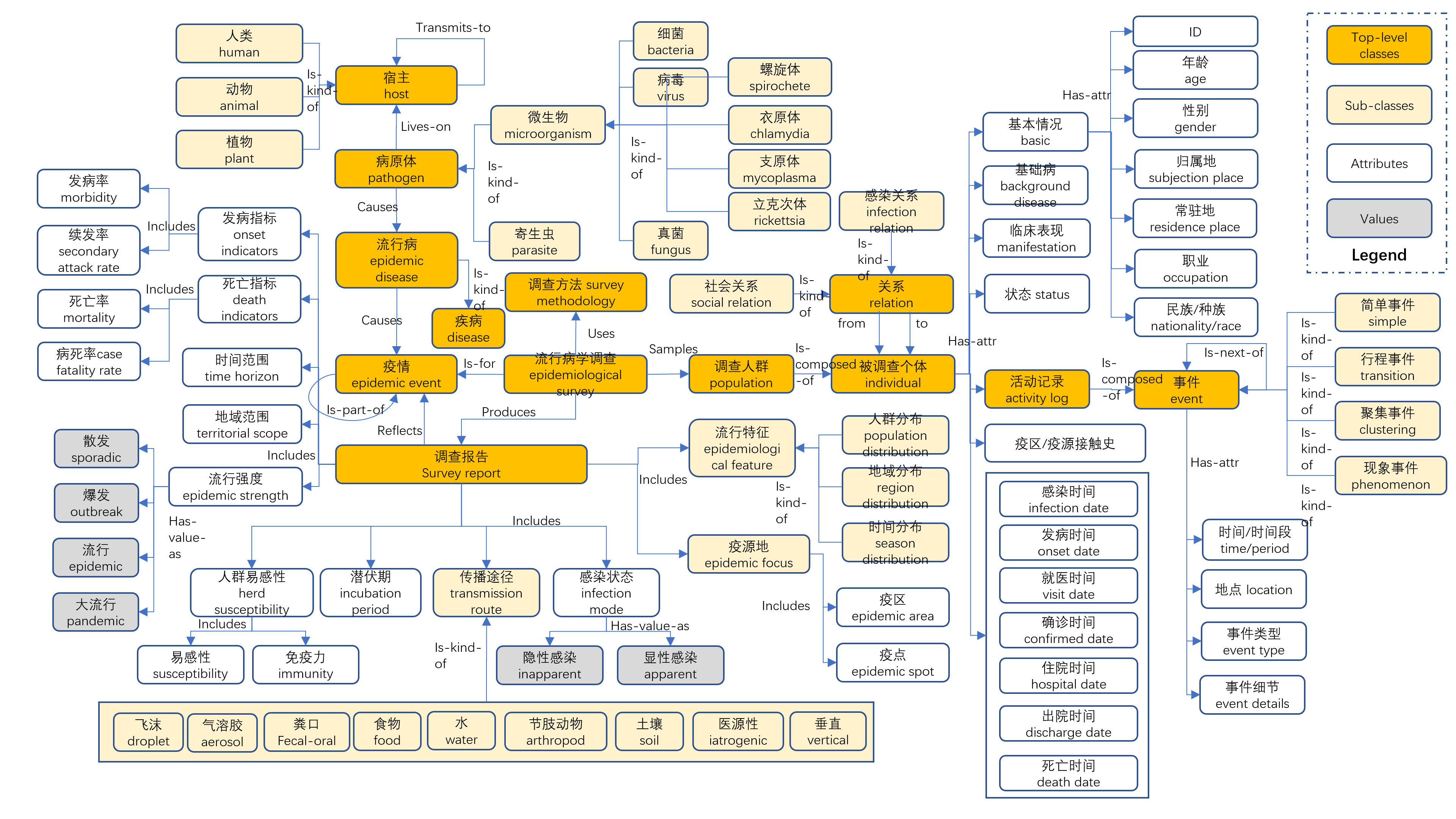
说到流行病学调查的复杂性，一则真实事件（天津宝坻百货大厦流行病学调查与溯源：<https://new.qq.com/omn/20200214/20200214A0TCXB00.html> ）可以给大家更多感性认识。调查过程堪比烧脑推理小说！新冠肺炎的传播链，远比你想象的危险……。



然而，针对数万病人进行流行病调查、溯源、密切接触者追踪将是非常耗时、耗力的工作。能否利用自然语言理解技术、基于流行病学知识图谱的分析推理技术为医护人员和疾病防控人员提供技术支撑，加速流调研究呢？答案是肯定的，这正是本知识图谱的一个典型应用场景。

**#.2数据schema**

流行病学知识图谱的schema定义了流行病学的基本概念，例如：流行病、病原体、宿主、疫情、流行病学调查、调查方法、调查人群、被调查个体、调查报告。其中主要对病原体、流行病学调查和调查报告进行了详细定义。详见下图。



**9.3数据来源与规模**

数据来源：

* 王建华，刘民，《流行病学-第7版》人民卫生出版社
* 陈清，徐德忠，《流行病学复习考试指导》人民军医出版社
* 各地卫健委公开信息

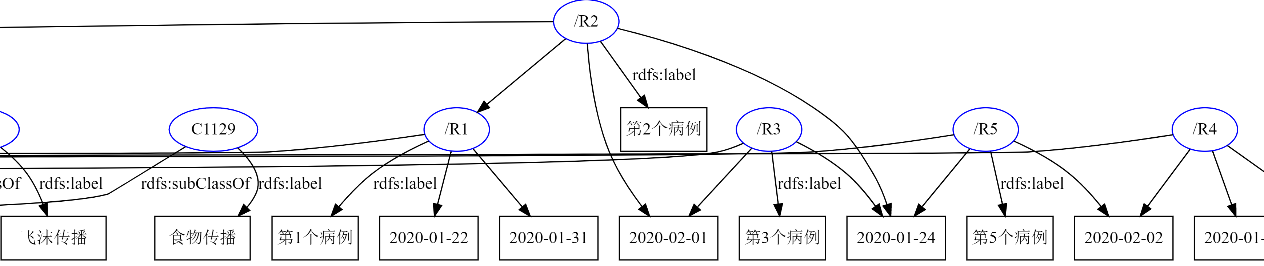
数据规模

* 概念：49 个
* 实例：6 个
* 数值属性：34个
* 对象属性：19个
* 三元组数量：374 个

**9.4规范**

* 命名空间：http://www.openkg.cn/COVID-19/epidemiology/
* 属性定义：http://www.openkg.cn/COVID-19/epidemiology/property/P+ID（Integer）
* 概念定义：http://www.openkg.cn/COVID-19/epidemiology /class/C+ID（Integer）
* 实体定义：http://www.openkg.cn/COVID-19/epidemiology/resource/R+ID（Integer）

**6.5可视化的图谱样例**



**6.6作者介绍**

* 李静（[jingli@cn.ibm.com](mailto:jingli@cn.ibm.com)） IBM中国研究院
* 王健（[wangwj@cn.ibm.com](mailto:wangwj@cn.ibm.com)） IBM中国研究院
* 蒋建民（jiangjm@cn.ibm.com） IBM中国研究院
* 王轲（wkebj@cn.ibm.com）IBM中国研究院
* 王彦菲（[yanfeiw@cn.ibm.com](mailto:yanfeiw@cn.ibm.com)）IBM中国研究院
* 梅婧（[meijing@cn.ibm.com](mailto:meijing@cn.ibm.com)） IBM中国研究院