

2017 “FME 的魔力” 博客征集大赛

——FME 在不动产数据整合中的应用

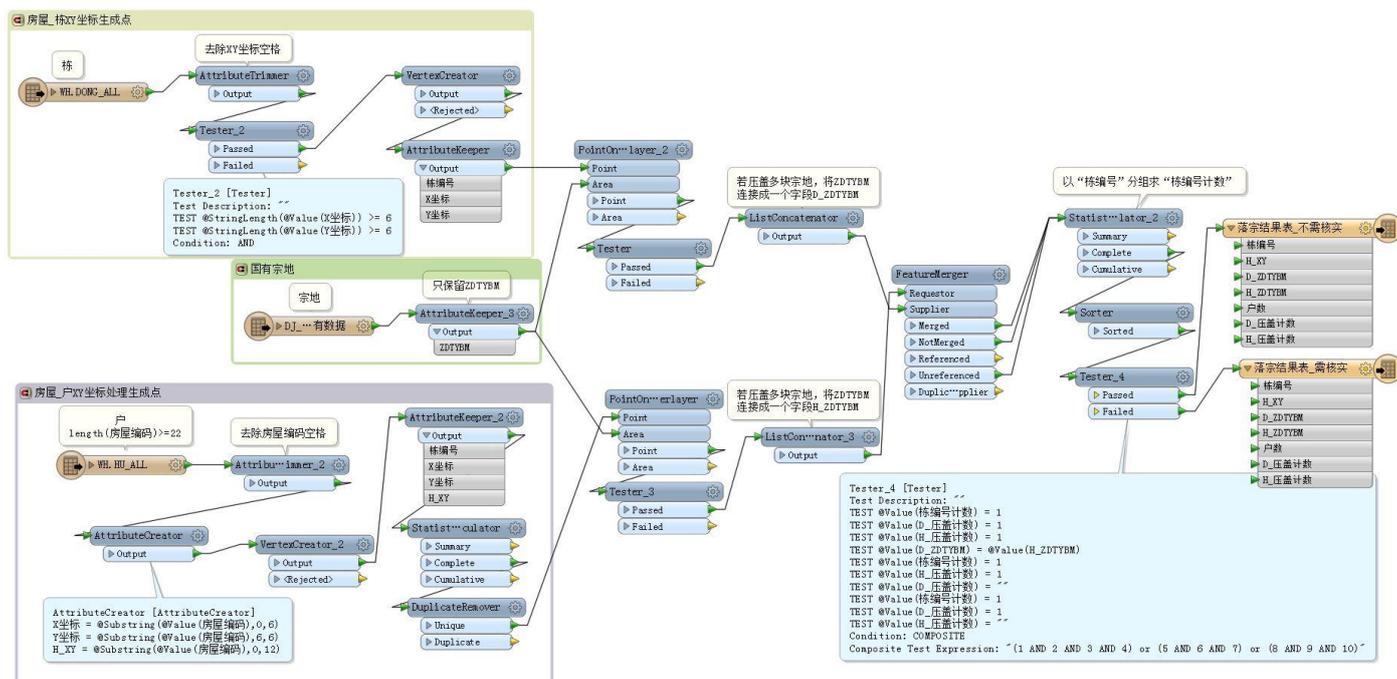
(by 风~~~~涟漪)

0 前言

目前全国正全力推动不动产统一登记的各项工作中，其中不动产登记数据整合建库作为信息系统建设的基础，为不动产登记各项业务办理提供数据支撑。FME以其强大的数据转换、处理分析、入库功能，贯穿于不动产数据整合的整个流程，包括国有宗地大小证分析、房屋落宗、房地挂接、栋基底和分层分户图处理、生成标准数据库GDB、数据检查及入库等，整个过程非常复杂且繁琐，并需要大量人工核实。鉴于篇幅有限，本文着重介绍几个应用实例。

1 房屋落宗

房屋户落宗采用四种方法：栋坐标落宗、户坐标落宗、栋图落宗、外业核实采集栋XY落宗。通过栋XY、户编码、栋图及外业采集XY生成坐标点，与宗地叠加，获取户所在宗地的宗地代码，用于后续不动产单元编码。由于成都市中心城区房屋栋和户数据大多包含坐标信息，因此主要采用栋和户坐标落宗，获取**落宗关系表**，其余两种方式作为补充这里不作说明。



实现过程：①栋表的X、Y坐标生成点与宗地叠加；②户（length(房屋编码)>=22）截取其房屋编码的前12位分别作为X、Y生成坐标点，按照**栋编号**和**H_XY**，**分组统计户数**、**再去重**，与宗地叠加；③两类落宗结果通过栋编号关联，结果以“栋编号”分组求“栋编号计数”，按“栋编号计数、D_压盖计数、H_压盖计数、压盖的两者宗地代码（D_ZDITYBM、H_ZDITYBM）”字段组合条件，输出**不需核实**和**需核实**的**落宗结果表**（前者抽查，后者逐条核实）。

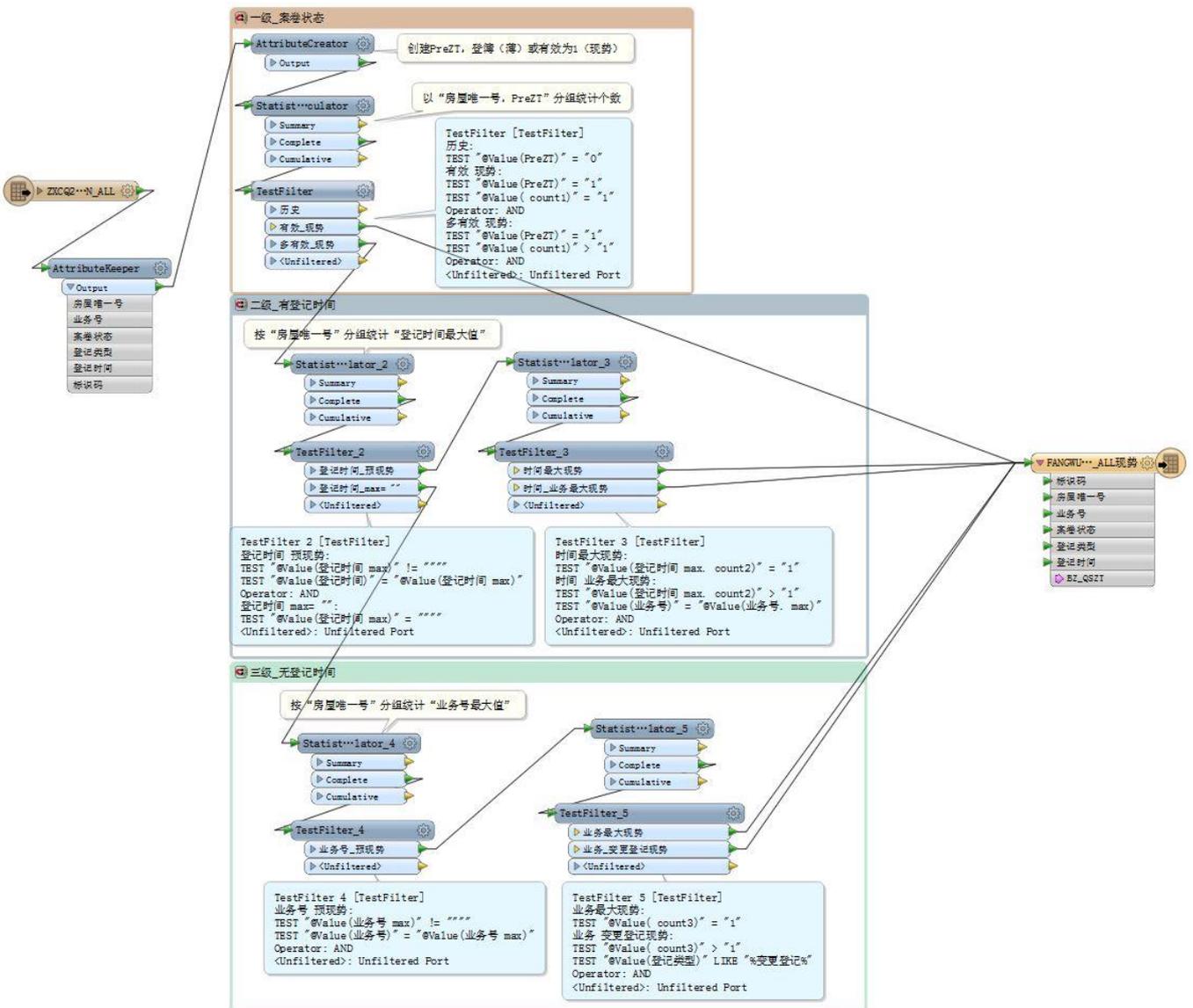
需核实的情况主要包括：栋（或户）坐标压盖多宗地、两者压盖宗地代码不一致、一个栋编号对应多个H_XY等。

2 房地关联

不动产数据整合数据库标准中，房地产权表是国土分户证与房屋产权现势数据相关信息的集中体现，需将二者正确关联。房地关联主要采用字段信息匹配、分户关系表、人工核查方式。本文主要说明字段信息匹配关联方式。

(1) 房屋产权现势判断

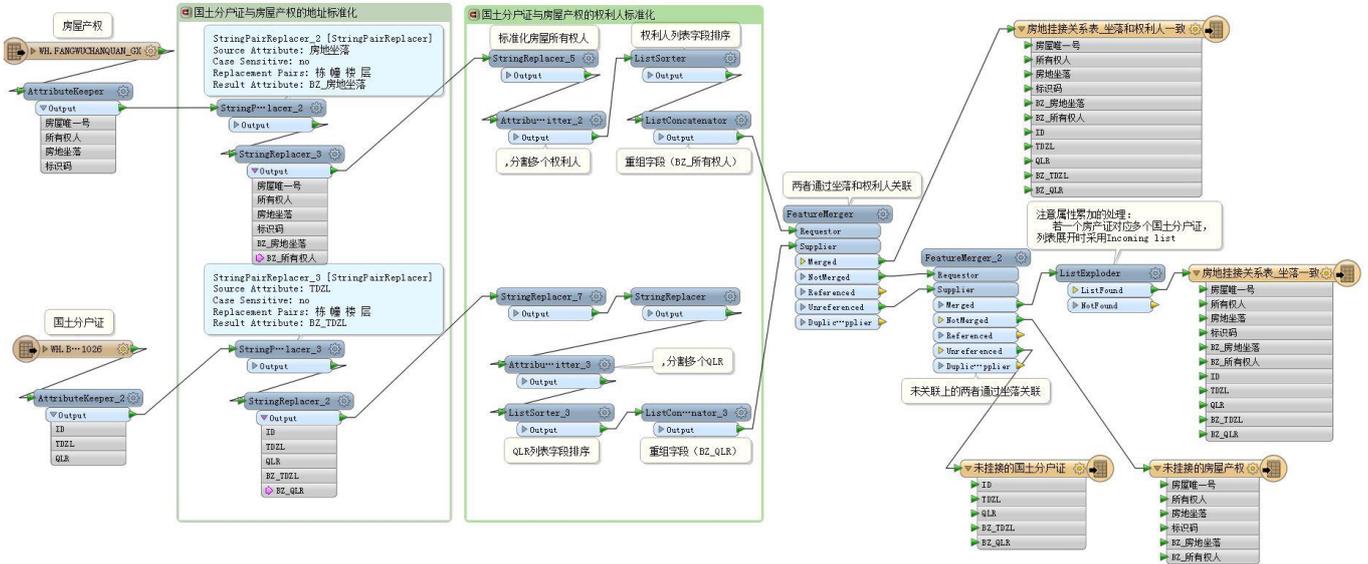
房屋产权数据（一房屋唯一号包括一条现势和多条历史），与国土分户证关联，需判断房屋产权现势，其判断规则：案卷状态为“登簿（簿）或有效”为现势，其余为历史，其中①若有多条现势，有登记时间的情况，最新时间为现势，仍有多条现势按业务号最大为现势；②若有多条现势，无登记时间的情况，最大业务号为现势，仍有多条现势按登记类型为变更登记为现势。



关键点：根据规则选择相应的字段作分组统计个数、最大值，再根据条件分类进行下一步处理。

(2) 国土分户证书与房屋产权现势数据关联

关联方式主要采用关键字段匹配。将国土与房产数据的坐落和权利人作标准化处理，然后标准字段关联匹配，获取**房地关联表**。



实现过程：①StringPairReplacer和StringReplacer将国土分户证与房屋产权的地址标准化；②两者权利人标准化；③通过标准化后的坐落和权利人关联；④未关联上的两类数据再通过标准化后的坐落匹配（若一房产对应多国土证，注意列表展开属性累加的设置）；⑤输出**房地关联表**（坐落和权利人一致、坐落一致），未关联的房屋产权表和国土分户证表。

关键点：对房屋产权的“所有权人”标准化，StringReplacer（正则表达式“(份额.*?)”将所有权人字段中含（份额*****）都去掉）、AttributeSplitter（“，”分割）、ListSorter和ListConcatenator（重组字段“BZ_所有权人”）。国土分户数据的权利人标准化类似，根据具体情况设置StringReplacer中参数。由于两类权利人字段存在“**相同的多个权利人，但顺序不一致**”的情况，因此对分割后的两类权利人列表字段皆排序，再重组权利人字段。

```
权利人 (份额: 1/2) 权利人 (份额: 1/2)
权利人 (份额: 99%) 权利人 (份额: 1%)
权利人 (份额: 1%) 权利人 (份额: 49%) 权利人 (份额: 50%)
权利人 (份额: ) 权利人 (份额: 50%)
权利人 (份额: 1平方米) 权利人 (份额: 1平方米) 权利人 (份额: 70.47平方米)
```

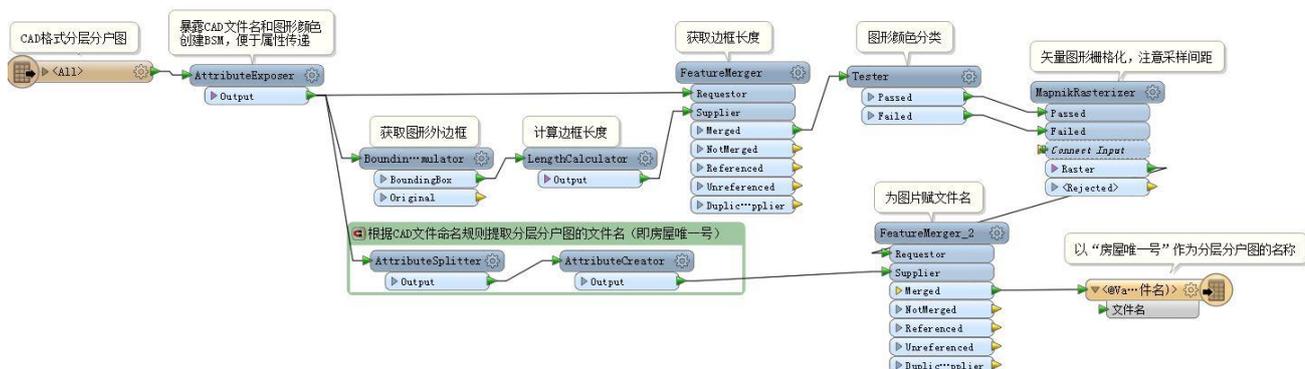
3 分层分户图转换与入库

房屋分层分户图数据转换与入库（数据库→CAD→JPG→数据库），处理简单，但流程较多、数据量较大。从数据库→CAD，采用AttributeFileWriter将存储图片的blob字段写出为CAD文件即可；从JPG→数据库，采用RasterExtractor将获取序列化栅格要素字段写入对应数据库表的blob字段即可。从CAD→JPG的处理过程稍微复杂一些，下面作具体说明。

(1) 单幅转换

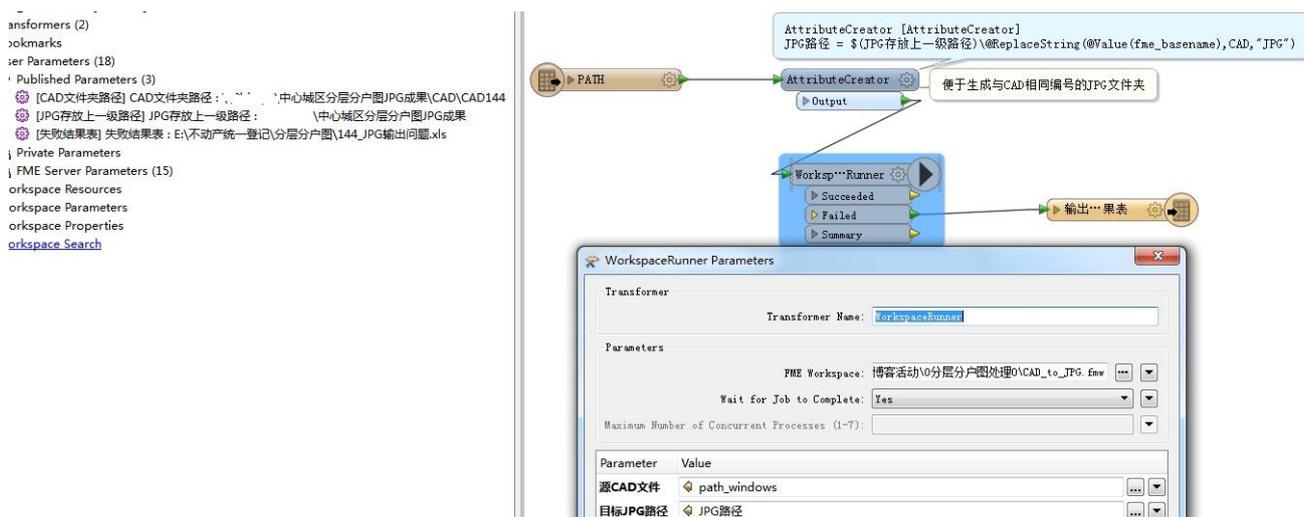
实现过程：①暴露CAD文件名和图形颜色，创建BSM；②获取图形外边框，计算长度；③矢量图形栅格化，根据图形颜色分类设置样式，注意采样间距（外边框长度/数值）；④转换器

BoundingBoxAccumulator和MapnikRasterizer处理皆会丢失属性，需创建BSM关联，便于属性传递；⑤获取CAD文件名的“房屋唯一号”作为文件名，写出JPG。



(2) 批量转换

成都市中心城区的房屋分层分户图多达三百多万，处理方式：先将CAD文件按一定数据量分为多个文件夹，然后采用FME的WorkspaceRunner转换器，调用单幅转换模板，批量处理灵活。一个关键点是JPG路径创建，可动态生成与CAD相同编号的JPG文件夹。



4 结语

不动产数据整合建库工作，当然远远不止如此。个人认为房屋落宗关系表和房地关联表，是整合数据的基础和关键，后续的工作基于两个关系表展开，过程较繁琐，这里不便详述。

FME的魔力就是实践中使用越多，越离不开它。现在数据分析处理、入库，首先想到它，省时省力、方便快捷，工作有了它，很多问题so easy! ^-^

感谢safe，感谢安图，带给我们这么好用的工具！