

基于 FME 进行的电子档案抽取与重挂接

涂倩云

1. 项目背景

实施不动产统一登记制度以来，各地持续推进不动产存量数据整合工作。不动产存量数据整合主要是将不动产统一登记制度实施以前，土地房屋分开登记的相关信息整合，实现房地一体化。在数据整合过程中就包含了从原房产系统抽取数据库中的电子档案，并将电子档案挂接至新不动产登记系统上的工作。

2. 解决方案

2.1 数据情况

原房产系统的电子档案存放于 Oracle 数据库的 FJ 字段，FJ 字段是 BLOB 类型的，电子档案与业务的关联字段为 LCSLID，FJCLGS 字段值为 PDF 的是 PDF 格式电子档案，其余为 JPG 格式电子档案。

SQL 输出 统计表

```
select * from cadastral_clfj t for update
```

	CLFJID	LCSLID	SJDBH	XH	CLFJMC	FJCLGS	FJ
4	d4d3f7ad-c0d8-4849-8448-980c8660048b ...	20150000074	1 ...	20150000074档案 ...	PDF ...	<BLOB> ...
5	1cc21ced-1bae-4803-85b6-1c6a0e04144f ...	20151170083	2 ...	商品房买卖合同	<BLOB> ...
3	ad449896-99c5-4883-b6c7-34dddab96305 ...	20151170083	1 ...	完税证明	<BLOB> ...
1	2bfb2f23-2d71-462d-adb6-0f9a13e94b6e ...	20151170083	1 ...	身份证复印件	<BLOB> ...
2	c43fa582-1226-4964-b385-82d861f7863c ...	20151170083	1 ...	商品房买卖合同	<BLOB> ...

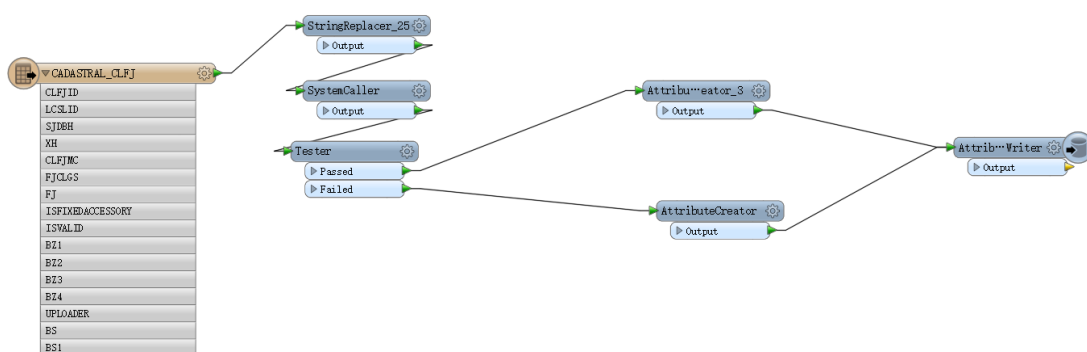
2.2 处理思路

从原房产库中导出 JPG 和 PDF 格式的电子档案，以 LCSLID 为单位，将同一 LCSLID 的档案存放到一个文件夹中。将档案路径写入新登记系统的数据库中，通过 FTP 访问电子档案。

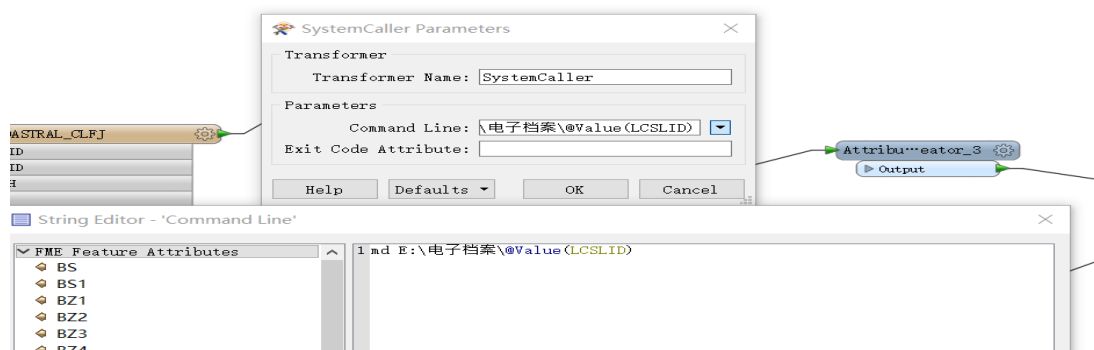
2.3 FME 模板实现

2.3.1 导出电子档案

主要用到 SystemCaller 和 AttributeFileWriter 转换器

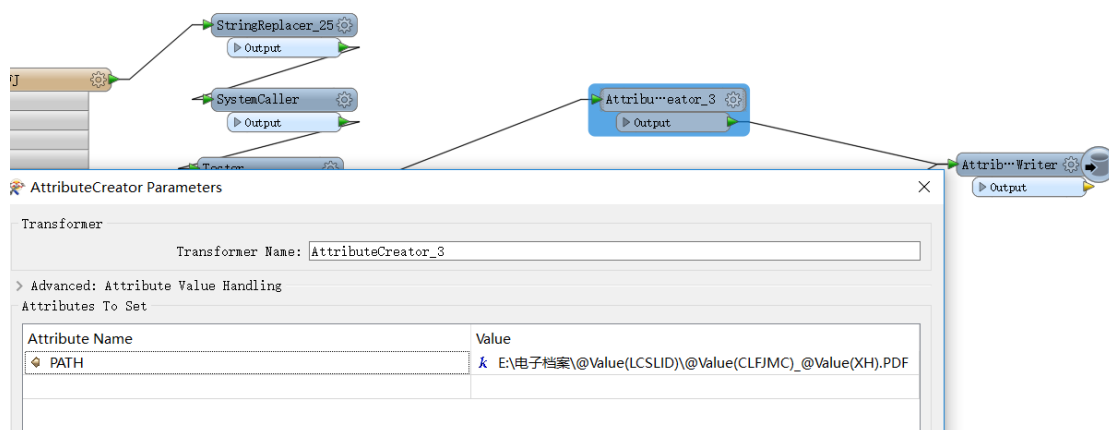


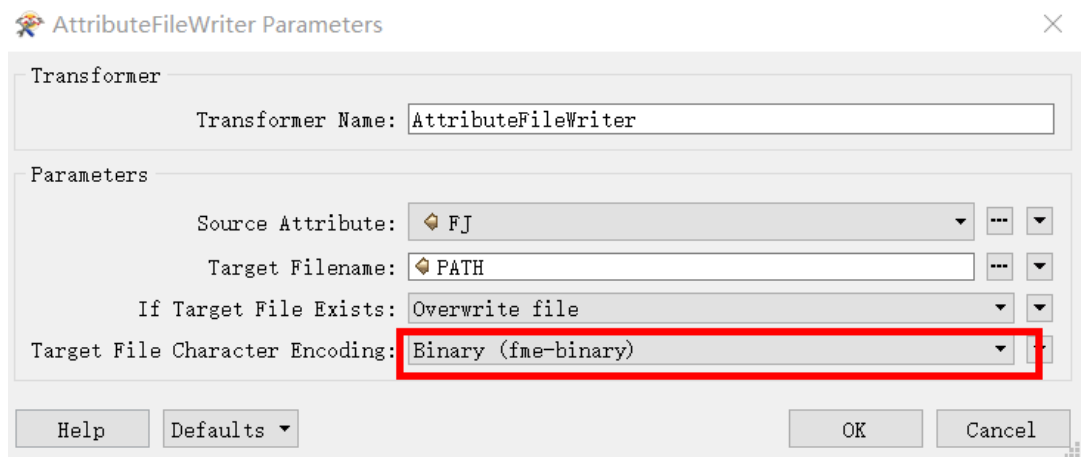
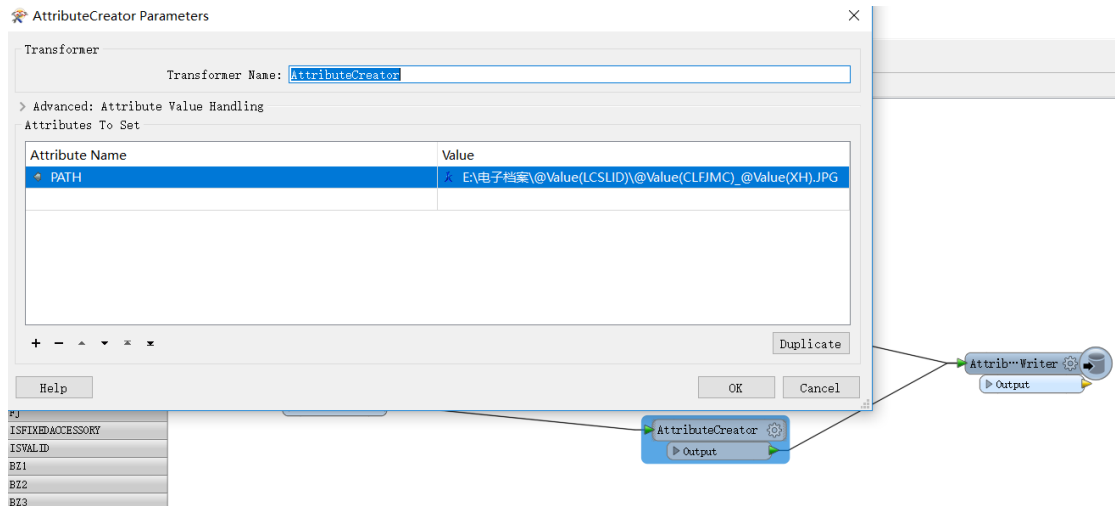
SystemCaller 转换器：用于运行程序或操作系统命令。
用系统命令 md 创建以 LCSLID 为文件名的文件夹



AttributeFileWriter 转换器：写出文件。

判断文件格式是 PDF 还是 JPG，根据文件格式及文件存放路径定义一个属性
“PATH”作为目标文件名。目标文件字符编码选择二进制编码。





导出结果



2.3.2 挂接电子档案

挂接电子档案主要是将电子档案的路径与其关联字段的关系对应好。通过 FME 获取电子档案路径的技术，最关键是使用到了 Directory and File Pathnames 这个读模块格式，将读模块路径定位到要读取的文件夹，并勾选此文件夹下面的所有子项，如图预览得到下图

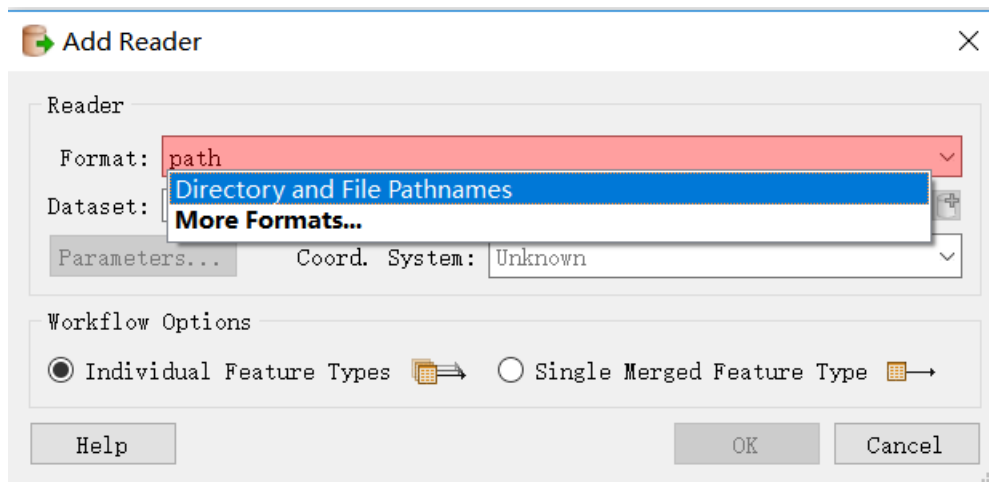
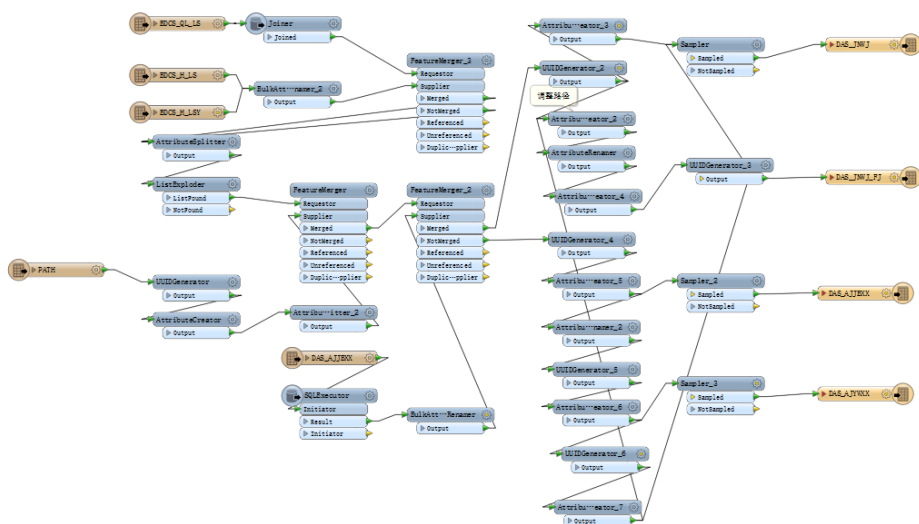


Table View

	path_unix	path_windows	path_rootname	path_filename	path_extension	path_filesize	path_type	path_ownership
1	E:\电子档案\20150000074	E:\电子档案\20150000074	20150000074	20150000074		0	directory	
2	E:\电子档案\20151170083	E:\电子档案\20151170083	20151170083	20151170083		0	directory	
3	E:\电子档案\20150000074\20150000074档案_1.pdf	E:\电子档案\20150000074\20150000074档案_1.pdf	20150000074档案_1	20150000074档案_1.pdf	pdf	0	file	
4	E:\电子档案\20151170083\商品买卖合同_1.jpg	E:\电子档案\20151170083\商品买卖合同_1.jpg	商品买卖合同_1	商品买卖合同_1.jpg	jpg	0	file	
5	E:\电子档案\20151170083\商品买卖合同_2.jpg	E:\电子档案\20151170083\商品买卖合同_2.jpg	商品买卖合同_2	商品买卖合同_2.jpg	jpg	0	file	
6	E:\电子档案\20151170083\完税证明_1.jpg	E:\电子档案\20151170083\完税证明_1.jpg	完税证明_1	完税证明_1.jpg	jpg	0	file	
7	E:\电子档案\20151170083\身份证复印件_1.jpg	E:\电子档案\20151170083\身份证复印件_1.jpg	身份证复印件_1	身份证复印件_1.jpg	jpg	0	file	

可根据 path_type 字段区分电子档案读取的层次，path_filename 为文件名，根据具体需求将 path_windows 内容加工填写到数据库中。



3.结语

FME 是一个非常强大的数据处理转换工具，在这个案例中，只是用到了他非常少的功能。在日常工作中，FME 给我提供了非常大的便利。对于批量处理数据，我首先想到的是利用 FME 进行操作。FME 的操作运用，值得我不断探索学习！