图纸识别逻辑层方案

在前期关系层输出信息的基础上，提取电气属性信息，形成XML文件。

端子排图中电气属性有关的信息主要有以下几类：

1. 单元格内的文字信息：单元格按行识别文字信息，可能包含的信息：端子排名标识，端子号、内部接线标识、回路功能描述、回路编号、缆芯功能描述、对侧端子号。
2. 电缆:表格外竖线，表示不同屏柜间的电气连接。
3. 电缆缆芯：表格外横线，与电缆竖线相连。
4. 电缆文字描述：电缆竖线上的文字描述，可能包含的信息：电缆编号、电缆型号、电缆规格、电缆备用芯数、已使用芯数、电缆去向。
5. 缆芯文字描述：缆芯横线上的文字描述，可能包含的信息：回路编号、缆芯功能描述、电缆芯号。
6. 接地符号：表格外接地符号，接地符号与端子连接关系识别。
7. 短连片：识别多个连续端子电气连接关系。
8. 短接线：表格外横竖线组合，用于柜内端子间连接，识别柜内端子间连接关系。

逻辑层识别分为表格信息识别和连接线信息识别两个部分，先识别连接线信息，再识别表格信息

**连接线信息识别**：

按照先识别电缆和电缆所属的缆芯，后面再识别电缆文字描述和缆芯文字描述，在电缆文字描述识别环节，按顺序识别电缆去向、电缆编号、电缆规格、电缆型号、电缆备用芯数、电缆已使用芯数，在缆芯文字描述识别环节按顺序识别回路编号、缆芯功能描述、电缆芯号，识别规则如下：

1. 电缆识别：竖线与一对多横线相连，且竖线左侧或者右侧有“至\*\*\*\*”字样，则确定是电缆。
2. 电缆缆芯识别：与电缆竖线相连的横线，确定为该电缆的缆芯。
3. 电缆去向识别：“至\*\*\*\*\*\*\*\*”字样，确定为电缆去向，“至”后面的字符为电缆终点。
4. 电缆编号识别：电缆文字描述中有一串字母数字或者特殊连字符“-”组合，判断为电缆编号。
5. 电缆规格识别：电缆文字描述中以A\*B为特征的一串字符代表电缆规格，其中A为电缆芯数（整数），B（整数或者实数，小数点后面最多两位）为缆芯截面积，控制电缆一般2~61芯，截面积有0.75、1.5、2.5。
6. 电缆型号识别：电缆文字描述中，电缆型号也是一串字母数字组合，在识别时与电缆编号容易混淆，电缆型号一般会与电缆规格组合在一起，当一串字母数字组合后面有电缆规格字符A\*B如7\*2.5,将此字母数字组合判断为电缆型号，否则判断为电缆编号。
7. 电缆芯号识别：横线与竖线交汇点附近的整数数字为电缆芯号。
8. 电缆备用芯数识别：电缆文字描述中，在电缆规格后面括号中的整数数字，如14×2.5(2)，括号中“2”为整数，代表备用芯数。
9. 电缆已使用芯数识别：电缆文字描述中，电缆文字描述中单独一个整数数字为电缆已使用芯数，电缆芯数=电缆已使用芯数+电缆备用芯数
10. 回路编号识别：与单元格相连有横线，且是缆芯，判断横线上是否有字母数字组合，如果有当成回路编号，如果没有则取与缆芯相连单元格内字母数字组合为回路编号。
11. 对侧端子号识别：与缆芯相连的单元格中，根据规则（\*D\*,\*D:\*、\*D-\*,\*D代表端子标识，最后一个\*代表端子号，可能是数字，数字+字母，数字+特殊字符“'”）识别出来是端子号，表示对侧屏柜接线端子号。
12. 缆芯功能描述识别：缆芯文字描述中有汉字、数字、字母组合，判断为缆芯功能描述,如缆芯文字描述中没有汉字、数字、字母组合，则查找与缆芯相连的表格中有汉字、数字、字母组合，判断为缆芯功能描述。
13. 短接线识别：竖线与多个横线相连，横线与单元格相连，竖线上无“至\*\*\*\*”的字符，竖线和多个横线一起判断为柜内短接线，各横线连接关系为上下顺序连接，结合表格信息识别出来的单元格中的端子号信息确定端子间的短接线连接关系，。

**表格信息识别**：

先识别端子排名和端子号，再按行识别单元格中的其他信息如内部接线标识，短连片，回路功能描述等，识别规则如下：

1. 端子排名识别：

按端子排区域part首先识别图纸中的分段标识

1)筛选单元格下方有一对多关系的单元格

2)筛选单元格中字母数字组合来辅助判别

3)单元格下方形成多列，筛选单元格下方某一列有连续的数字字母组合来最终判别，确定为端子排名标识。

1. 端子号识别：按列判断数字或者字母连续性（3个及以上端子连续），连续性分为以下几种情形：
2. 1,2,3,4,5。
3. 1,2,3,3a,3b,或者1,2,3,3A,3B,4
4. 1,2,3,4,4’4”
5. 在端子排名单元格下方按行形成单元格集合，考虑合并单元格的情况，每一行按顺序识别内部接线标识，回路编号，回路功能描述、缆芯功能描述、对侧端子号。
6. 内部接线标识：分类判别：装置类\*n，端子类(本图纸端子排标识识别+按端子通配关系判别，\*D\*,\*D:\*、\*D-\*,\*D代表端子标识，最后一个\*代表端子号，可能是数字，数字+字母，数字+特殊字符“'”)，压板类LP，按钮类FA、YA、QA、FGA，空开类\*K,\*KK,\*JK，把手类QK,WK,SK，指示灯类，继电器类\*ZJ、K\*。
7. 回路编号：与单元格相连有横线，且是缆芯，判断横线上是否有字母数字组合，如果有当成回路编号，如果没有则取单元格内字母数字组合为回路编号
8. 回路功能描述：单元格中有汉字、数字、字母组合，且没有与电缆缆芯横线相连，判断为回路功能描述。
9. 缆芯功能描述：缆芯横线上有汉字、数字、字母组合，判断为缆芯功能描述，如果没有则取与缆芯横线相连单元格中汉字、数字、字母组合为缆芯功能描述
10. 对侧端子号：与电缆缆芯横线相连的单元格中，单元格内可能有两部分内容，一部分是回路编号，一部分是对侧端子号，结合对侧屏柜端子排名识别带有\*D\*,\*D:\*、\*D-\*特征的,\*D代表端子标识，最后一个\*代表端子号，可能是数字，数字+字母，数字+特殊字符“'”，判断为对侧端子号。

一个part中按行形成的单元格信息全部识别之后再识别短连片、接地符号、试验端子。

1. 短连片识别：根据关系层输出加上逻辑层端子排名和端子号识别结果，确定短连片连接位置
2. 接地符号识别：根据关系层输出接地符号与单元格连接关系，加上逻辑层端子排名和端子号识别结果，确定接地符号连接位置
3. 试验端子识别：关系层识别出来试验端子及试验端子与单元格关系，加上逻辑层端子排名和端子号识别结果，输出哪些端子是试验端子。