**元数据梳理完成度指标说明**

# 一、元数据梳理考核指标：

功能覆盖度A1：填报的系统功能数/系统实际功能数×100%

数据表覆盖度A2：填报的数据表数/系统中实际数据表数×100%

数据字典覆盖度A3：填报的数据字段数/系统中有效表的字段数×100%

总体覆盖度A：（A1+A2+A3）/3×100%

字段总数N：M01～M06的字段数×记录数之和

总体规范度M：M01～M06中符合规范要求的字段总数/N×100%

要求总体覆盖度A达到100%，总体规范度M达到90%。

# 二、元数据梳理模板填写规范：

# 1、总体要求

* 梳理模板各项内容按照生产系统和生产数据库状态填写
* “主页”应当填写“填报依据文件清单”，并将各项填报依据文件一同上报，填报依据文件包括生产系统细化到最末级功能菜单的截图
* “文档修订记录”中的日期作为元数据梳理时间，表明其它各表中的数据的填写时间
* 下文所提《EA架构》为《中国南方电网有限公司企业架构（2018修订版）》
* 单元格不可合并
* 文档中填写的部门、岗位名称应当与人力资源系统中公司组织架构的信息保持一致。

# 2、M01元数据-应用域和系统，每个系统一条记录

* 单位：各系统填报单位全称，如广西电网有限责任公司
* 一级应用域编号：根据《EA架构》中应用架构的应用域编号填写
* 一级应用域名称：根据《EA架构》中应用架构的应用域名称填写
* 应用编号：根据《EA架构》中应用架构的应用编号填写，没有应用编号的留空不填写。
* 应用名称或系统名称：根据《EA架构》中应用架构的应用名称或应用系统名称填写（资产管理系统填写子系统名称）
* 系统版本：系统当前版本，格式为“VN.n”，如“V2.1”
* 系统建设时的业务负责部门：该系统建设时，负责的业务部门；通常为提出主要业务需求的业务部门。应当填写人力资源部门规定的部门全称。
* 主要业务使用部门：主要使用该系统的业务部门；通常一个综合性业务系统为多个业务部门提供服务。应当填写人力资源部门规定的部门全称。
* 承载网络：系统主要运行的网络区域，即安全I区、安全II区、安全III区等
* 数据库类型：系统使用的数据库产品名称，如Oracle、greenplum等
* 数据库版本：系统使用的数据库产品版本，如果为Oracle，则填写形如“11.2.0.4”
* 系统概述：系统主要功能概述
* 系统设计厂商：系统功能、技术、界面的设计厂商，全称、简称皆可。
* 设计商联系人：系统设计厂商的联系人和联系方式，格式为“联系人姓名：联系方式（手机号码或座机号码）”。
* 系统开发厂商：系统的功能开发、技术实现的厂商，全称、简称皆可。
* 开发商联系人：系统开发厂商的联系人和联系方式，格式为“联系人姓名：联系方式（手机号码或座机号码）”。
* 系统运维厂商：系统上线运行后维护正常使用的厂商，全称、简称皆可。
* 运维商联系人：系统运维厂商的联系人和联系方式，格式为“联系人姓名：联系方式（手机号码或座机号码）”。
* 首次上线时间：系统首次上线时间，精确到月份，格式为“YYYY-MM”，如“2017-03”
* 最近一次上线时间：系统最近一次上线时间，精确到月份，格式为“YYYY-MM”，如“2017-03”

# 3、M02元数据-系统数据资产概要，每个生产数据库一条记录

* 应用编号：与M01的对应项一致
* 应用名称或系统名称：与M01的对应项一致
* 数据库名称：系统使用的数据库实例名或服务名，如果该实例下有多个schema，需要标明对应系统使用的schema，实例名和schema以“:”（英文半角冒号）分隔，schema之间以“/”分隔，如“nwoa:cfg/oa/trs”
* 数据库字符集：系统使用的数据库字符集，统一大写英文，如“ZHS16GBK”。
* 数据总量（GB)/记录数：填报时的生产系统数据总量（生产库大小和记录数，以“/”分隔），填写不带分隔符的阿拉伯数字。
* 日增量(MB)/记录数：填报时的生产数据库日增量（生产库大小和记录数，以“/”分隔），可根据多个备份时点的大小差异估算，无需精确，填写不带分隔符的阿拉伯数字。
* 数据保留时间：系统数据保留时间，未设定保留时间的填“永久”。
* 主要业务数据对象：根据《EA架构》中数据架构的概念实体填写，没有对应概念实体的，按数据架构的数据主题填写。
* 备份方式：数据库采用的备份方式，如磁带、虚拟磁带、磁盘。
* 备份周期：数据库的备份周期和策略，填写格式例如“每周六全量”、“每天增量”等。
* 历史数据保存时间：历史数据库中已有历史数据的保存时间，如“1年”、“6个月”等；没有历史数据库，所有数据均在生产库的，则根据备份数据的保存时间填写。
* 是否可恢复：备份数据是否可恢复
* 恢复数据的系统要求：恢复数据时需要的系统、数据库、工具等环境，格式为“系统名称（版本号）/数据库名称（版本号）/工具名称（版本号）”
* 历史数据和生产数据格式是否一致：历史数据的数据结构是否和当前生产库一致；没有历史数据库的填写“是”。
* 备注：需要说明的其它内容

# 4、M03元数据-系统数据接口，每个接口一条记录

* 应用编号：与M01的对应项一致
* 应用名称或系统名称：与M01的对应项一致
* 接口编号：设计文档中规定的接口编号，设计文档中未规定时，不填写
* 接口名称：设计文档中规定的接口名称，设计文档中未规定时，不填写
* 接口类型：填写“接收”、“发送”或“双向”。
* 对方系统编号：根据《EA架构》中应用架构的应用编号填写，对于未纳入EA架构的系统，留空不填写。
* 对方系统名称：根据《EA架构》中应用架构的应用名称或应用系统名称填写（资产管理系统填写子系统名称），对于未纳入EA架构的系统，按系统实际名称填写
* 对方接口编号：本方系统设计文档中对对方接口编号有明确规定的，填写该编号；本方系统设计文档中对对方接口编号未规定的，按照对方系统梳理结果清单中的对应接口编号填写；如对方系统尚未进行梳理时，留空不填写。
* 接口协议：仅指实现数据交换的接口实现方式，包括SOA、FTP、数据库直连、消息队列、数据库复制、其它，当填写其它时，在备注内予以说明。
* 主要的接口数据对象：可以明确数据表的，填写表英文名称，否则根据《EA架构》中数据架构的概念实体填写，没有对应概念实体的，按数据架构的数据主题填写。如果存在多个数据对象，以“/”分隔。
* 数据包格式：数据包的实际构成格式，而非从设计文档中获得，如xml、json、txt、csv等，当无法明确数据包格式时，填写“无”。
* 实时/批量：数据的同步时效是实时的或者是批量的，不是接口协议执行的时效，填写实时或批量
* 传输时间窗口要求：存在时间窗口要求则填写传输的起止时间要求，形如“每周六-每周日”、“每月10日”、“每月最后一天”、“每天18:00-次日8:00”等，当填写时间时，精确到分钟，格式为“hh24:mm”；不存在时间窗口的，例如由实时操作引发的数据传输，没有固定频率的，填写“无”。
* 传输频率：实时接口填写“无”；如果是批量接口，需要填写传输频率或周期，如日、8小时、周、月等，非周期性传输的，填“无”。
* 是否具备数据校验机制：接口传输数据时是否具备数据校验的机制，填“是”或“否”。
* 数据包大小限制：数据包大小限制，字节为单位，KB、MB、GB等大写
* 备注：需要说明的其它内容

# 5、M04系统数据表，每个物理表一条记录。

生产系统具有，但数据库设计文档中不具备的表，用橙色背景表示：此类表需要通过理解业务流程和功能，并和开发商确认才能判定是否有效，验证为有效后，需填写表中文名称和其他相关信息，数据表验证状态变更为“有效”；验证后为无效的表，按无效表标注，数据表验证状态为“无效”，无需填写其他信息；处于待验证状态的表，数据表验证状态为“待验证”。最终填报时不应当存在“待验证”的表。

* 应用编号：与M01的对应项一致
* 应用名称或系统名称：与M01的对应项一致
* 表中文名称：表的中文名称，而非英文名称拷贝
* 表英文名称：表的英文名称
* 数据表验证状态：有效（数据表在用，可以明确与业务流程和功能的关联，或支撑系统运行）、无效（数据表确认无效，是临时表或备份表，下线或删除后对系统正常生产不产生任何影响）
* 表描述：表的业务含义，而非表中文名拷贝
* 数据域：根据《EA架构》中数据架构的数据域名称填写，用于支撑系统运行的表填写“系统支撑表”
* 一级数据主题：根据《EA架构》中数据架构的数据主题名称填写，系统支撑表填写“无”。
* 二级数据主题：根据《EA架构》中数据架构的概念实体名称填写，概念实体无定义时，填写“无”，系统支撑表填写“无”。
* 创建时间：数据库中标识的表的创建时间，格式为“YYYY-MM-DD”
* 创建者：数据库中标识的表的创建者
* 管理员：对该数据负有管理职责的岗位和人员名称，通常为甲方DBA的人员名称和岗位，格式为“人员名称/岗位名称”。
* 现有记录数：截止到填报日期时数据库中该表的记录数，填写不带分隔符的阿拉伯数字。
* 存储周期：数据表的存储周期，未设定的填写“永久”。
* 初始记录创建时间：如果该表有时间戳或能表明记录创建时间的字段，以该字段的最早时间为准，格式为“YYYY-MM-DD”；如果表内记录数为“0”，则填写“无”；如果无法准确获得该表初始记录创建时间，则留空不填写。
* 最新记录创建时间：如果该表有时间戳或能表明记录创建时间的字段，以该字段的最新时间为准，格式为“YYYY-MM-DD”；如果表内记录数为“0”，则填写“无”；如果无法准确获得该表最新记录创建时间，则留空不填写。
* 数据标准责任部门：负责该数据的数据标准及标准符合性的部门名称，从一级部门开始填，可包含下级部门及路径。
* 数据标准责任岗位：负责该数据的数据标准及标准符合性的岗位名称
* 数据质量责任部门：负责该数据的数据质量的部门名称，从一级部门开始填，可包含下级部门及路径。
* 数据质量责任岗位：负责该数据的数据质量的岗位名称

以上四项的说明：

已经建立数据管控体系的，按体系中规定的标准和质量责任部门和岗位填写，未建立数据管控体系的，则：

数据标准责任部门：填写该表所对应功能需求的发起部门。

数据标准责任岗位：填写该表所对应功能需求的发起岗位。

数据质量责任部门：可以不填写。

数据质量责任岗位：可以不填写。

* 归属的一级菜单名称：该表归属的应用系统的一级菜单名称，与M05里面的一级功能对应。
* 备注：需要说明的其它内容

# 6、M05元数据-功能、业务流程和数据表。梳理每个菜单功能的业务流程、环节、功能点，一个大流程由多个菜单功能实现时，需填写每个功能实现的流程环节；需求编号和需求名称仅为帮助理解业务功能，可不填；一个功能可能需要多个数据表，需详细、准确填写功能使用的每张数据表；

* 应用编号：与M01的对应项一致
* 应用名称或系统名称：与M01的对应项一致
* 一（二、三、四）级功能名称：按每个最末级菜单功能项梳理，无三、四级菜单时，三、四级菜单填“无”；最末级菜单超出四级菜单时，按实际菜单级别数量扩展。其中，一级功能要与M04“归属的一级菜单名称”对应。
* 业务流程编号：按需求规格说明书中定义的业务流程编号填写，并标明该流程所处页码，格式为“编号/文档名/页码（格式：Pxx）”；需求规格说明书未定义时，可以不填写
* 业务流程名称：按需求规格说明书中定义的业务流程名称填写，并标明该流程所处页码，格式为“编号/文档名/页码（格式：Pxx）”；需求规格说明书未定义时，可以不填写
* 功能需求编号：可以明确需求编号则填写，否则不填写
* 功能需求名称：可以明确需求名称则填写，否则不填写
* 业务环节编号：如果业务流程的所有环节都在此最末级菜单内实现，可不填，否则应当明确本级功能内实现的业务环节编号
* 业务环节名称：如果业务流程的所有环节都在此最末级菜单内实现，可不填，否则应当明确本级功能内实现的业务环节名称
* 数据表1（2,3……N）中文名/表英文名：填写数据表的中文名和英文名，以“/”分隔
* 备注：需要说明的其它内容

# 7、M06元数据-数据字典,每个字段一条记录

底色标记为红色的，表示从生产系统采集出来的字段定义与数据库设计文档中数据字典里该字段定义不一致；底色标记为黄色的，表示生产库中存在该字段，但数据库设计分册中的数据字典里不存在该字段。

* 表英文名称：数据表的英文名称
* 表中文名称：数据表的中文名称
* 字段名称：字段的英文名称
* 字段中文名称：字段的中文名称
* 数据域：根据《EA架构》中数据架构的数据域名称填写，当字段所属数据域与表所属数据域一致时，填写“一致”。
* 一级数据主题：根据《EA架构》中数据架构的数据主题名称填写，当字段所属一级数据主题与表所属一级数据主题一致时，填写“一致”。
* 二级数据主题：根据《EA架构》中数据架构的概念实体名称填写，当字段所属二级数据主题与表所属二级数据主题一致时，填写“一致”。
* 业务规则和含义：字段的业务含义描述，而非字段中文名拷贝
* 缺省值：字段的缺省值，无缺省值时填写“无”。
* 字段类型：数据库中定义的字段类型
* 常用单位：字段是度量值时，填写常用单位；没有单位的填写“无”。
* 取值范围：字段是度量值时，填写字段的取值范围；字段是枚举值的，填写枚举值，格式为“枚举值1/枚举值2/……”；字段是枚举代码时，如果代码存储在统一代码表内，则填写代码的编号、名称、代码值和代码含义，格式为“（编号/名称/代码值/代码含义）（编号/名称/代码值/代码含义）……”；如果代码表为独立表，则在M05中对应功能应填写该代码表，此处填写代码值和代码含义，格式为“（代码值/代码含义）（代码值/代码含义）……”。
* 数据长度：数据库中定义的字段长度
* 数据精度：当字段为数值类型时，小数点右侧的位数；非数值类型的填写“无”。
* 数据格式：填写该字段的格式要求，如字段为编码，则填写编码规则。应当填写形如“YYYY-MM-DD”、“hh24:mm:ss”、“型如ABBCCC，其中A为1～9的自然数，BB为00～19，CCC为001～999”等
* 是否主键：是或否
* 是否可空：是或否
* 是否外键：是或否
* 外键关联信息：该外键字段关联的对象表内的相应字段，格式为“表英文名.表英文字段名”。
* 是否加密：是否需要加密，是或否
* 是否共享：是否可以内部共享，是或否
* 共享条件描述：当“是否共享”填“是”时，需要填写共享条件，若无条件共享，填写“无”。

以上三项的说明：是否共享、加密等可以在数据目录构建过程中由业务部门确认，元数据梳理过程可以不填写。