2.存储使用运维规范

随着预测与健康管理技术的高速发展，PHM模型的复杂性越来越高，对开发者的综合素质要求也越来越高，不仅是编程知识点，其它维度的知识点也会影响到模型的最终交付质量。为了保障风洞一体化平台模型的综合质量，需要对模型的开发流程、编程规范、异常处理、功能测试、安全规约、数据存储、工程结构、设计规约等方面做出相应的规范进行约束。

2.1建模规范

在为了保障模型质量应做出的各个方面规范中，最重要的就是要在开发模型时遵循规范的软件工程开发方法。从60年代末以来，出现了许多软件工程方法，其中最具影响的是面向对象方法、结构化方法和形式化方法，前两种方法是使用最广的两种方法，本部分主要介绍面向对象和结构化两种软件工程方法的特点及过程，对两种方法做出对比，并对模块化的概念做出介绍。

2.2存储规范

《基于文档型非关系型数据库的档案数据存储规范》中对不同类型、不同格式的档案数据存储制定了规范。为了适应不同的内容数据格式以及元数据需要，模型库中字段的类型应包括字符串、数值、日期、时间、文本、二进制等，其中文本字段用于存储文本中的句子和段落，包括从电子文件中抽取的文字信息（或称文本数据），二进制字段用于存储任何类型的字符，包括ASCII字符以及图像、视频、音频等二进制数据。文档型数据库应该允许创建不同类型的字段，存储任意格式的数据。模型库中的存储对象要包括内容数据等非结构化数据、XML文件等半结构化数据以及元数据等。对于数据的存储方式，可以按照数据的类型使用不同的存储方式。

当需要存储图数据时，由于图数据中最重要的数据是图中的顶点和边的信息，因此可以使用图数据库来存储这一类数据。图数据库是以图论为基础的一种非关系型数据库，其数据库的存储结构以及数据库的查询方式都是以图论为基础的，能够很好的存储并查询图数据。在具体的应用中，可以选用目前较流行的Neo4j图数据库。

当需要存储表数据时，可以使用经典的关系型数据库来存储。关系型数据库建立在关系模型之上，关系模型中数据的表示方式即为二维表格形式，完美地匹配了表数据的格式。在具体的应用中，可以使用轻便、快速的关系型数据库MySQL，也可以选用Microsoft SQL Server、Oracle等。

当内容数据全部载入模型库时，模型库会迅速变大。当模型库容量超过备份介质的容量时，就难于对模型库整体进行备份，需要将模型库拆分成若干个小容量的模型库才能进行备份。当一个模型库被分成若干个数据完整的子模型库时，可以按照模型库中的记录为单位进行拆分，通过模型库中的记录号、记录的容量或检索和统计分析从原模型库中提取相应的记录，按顺序将记录导入各子模型库。每一条记录是一组完整的相关信息，包括内容数据及元数据，保证子模型库中各记录的关联数据完整。模型库拆分可以部分拆分，模型库拆分生成的子模型库结构与原模型库结构可以不同。

子模型库的合并可以按照子模型库中的记录为单位进行合并，通过各子模型库中的记录号、记录容量或检索和统计分析从各子模型库中提取相应的记录，按顺序将记录导入一个模型库。模型库合并可以合成子模型库中的部分，记录模型库合并形成的模型库结构与子模型库结构可以不同。

在计算机系统中，由于计算机只能够处理数字，如果要处理文本就必须先把文本转换为数字才能进行处理，因此要对文本字符串进行编码。在字符的编码和校验过程中，为了能够支持中文，应该使用utf-8编码，使用utf-8-cgi校验。除了对数据进行编码之外，还需要为模型库本身以及其内部的表和字段进行合适的命名，并且要对数据进行备份处理，命名规范和数据备份将在下面进行说明。

2.2.1模型命名规范

模型库的命名要采用26个英文字母（区分大小写）和0-9的自然数加上下划线’\_’组成，命名简洁明确，多个单词用下划线’\_’分隔，一个项目一个模型库，多个项目慎用同一个模型库。

在对模型库中的表命名时，遵循以下规范：

①采用26个英文字母（区分大小写）和0-9的自然数（经常不需要）加上下划线’\_’组成，命名简洁明确，多个单词用下划线’\_’分隔。

②全部小写命名，禁止出现大写。

③禁止使用模型库关键字，如：name，time， datetime，password等。

④表名称不应该取得太长（一般不超过三个英文单词）。

⑤表的名称一般使用名词或者动宾短语。

⑥用单数形式表示名称，例如，使用employee，而不是employees。

⑦表必须填写描述信息（使用SQL语句建表时）。

对模型库中字段进行命名时，应该遵循以下规范：

①采用26个英文字母（区分大小写）和0-9个自然数（经常不需要）加下划线’\_’组成，命名简洁明确，多个单词用下划线’\_’分隔。

②全部小写命名，禁止出现大写。

③字段必须填写描述信息。

④禁止使用模型库关键字，如：name、time、datetime、password等。

⑤字段名称一般采用名词或动宾短语。

⑥采用字段的名称必须是易于理解，一般不超过三个英文单词。

⑦在命名表的列时，不要重复表的名称。

⑧不要在列的名称中包含数据类型。

⑨字段命名使用完整命名，禁止缩写。

在风洞项目中，拟使用一个总模型库存放多个模型的相关数据。此时，可对多个模型进行编码排序，如将多个模型使用ABCD等英文字符或其它编码进行统一规范，每一个模型有全局唯一的编码名称。每个模型中又有不同的表存储相关数据，对于模型中表的命名，可以约定使用规范的表名或使用表中文拼音的缩写替代，这样有助于模型库中命名的一致性。模型库中数据表及字段命名结构示例如下图x所示：



图x 模型库中数据表及字段命名示例

2.2.2模型数据备份与恢复

在使用模型库存储数据文件时，可能会出现以下场景导致数据丢失或损坏：

①人为操作失误造成某些数据被误操作；

②软件 BUG 造成部分数据或全部数据丢失；

③硬件故障造成数据库部分数据或全部数据丢失；

④全漏洞被入侵数据恶意破坏。

还可能会由于特殊应用场景下基于时间点的数据恢复、开发测试环境模型库搭建、相同模型库的新环境搭建、模型库或者数据迁移等情况需要对数据进行迁移处理，因此应该对模型库本身及其数据进行备份，以便各种事件发生时能够迅速恢复数据。模型库备份的内容应该包括模型库中的数据、模型库结构和模型库定义文件。应按照模型库结构（字段）备份模型库数据。数据的备份方式可以有整体备份、拆分备份、增量与差异备份、在线备份和离线备份以及自动备份几种方式。

整体备份是指当模型库的容量小于备份介质的容量时，无需对模型库做任何处理，直接对模型库整体进行复制备份。拆分备份是指当模型库的容量大于备份介质的容量时，应将模型库拆分成容量小于备份介质容量的若干个子模型库，然后将各子模型库分别备份到备份介质上，保留原有的访问控制策略，并保证原模型库的完整性。增量与差异备份是指对模型库中新增的数据进行增量备份，对被修改的数据进行差异备份。在线备份指将模型库中的数据、模型库结构和模型库定义文件备份到在线存储介质上。离线备份指是将模型库中的数据、模型库结构和模型库定义文件备份到离线存储介质上。自动备份通过软件的控制方式将模型库中的数据、模型库结构和模型库定义文件进行备份。

在风洞一体化平台数据的备份时可以采取增量与差异备份，这种方法不会产生重复的备份数据，备份数据需要时间较短且相较于其它方法更能节省磁盘空间。

在对数据进行还原时，又有整体还原和合并还原几种方式。整体还原将整体备份的模型库数据还原到原模型库系统。合并还原将拆分备份的多个子模型库数据还原到原模型库系统，采用的方法包括：

（1）在新建模型库系统合并全部子模型库数据，然后将合并形成的模型库数据还原到原模型库系统；

（2）在原模型库系统中合并还原全部子模型库数据。

在意外发生时的恢复策略也有正常恢复和异常恢复两种。正常恢复指使用备份的模型库数据直接覆盖原模型库数据。异常恢复指在异常情况下，用备份的模型库数据、模型库结构、模型库定义文件和日志文件进行恢复，并进行数据完整性校验，以确保数据的完整性。