

# 材料测试表征和制备工艺数据库无代码生成系统<sup>①</sup>



杨小渝<sup>1,2</sup>, 张 蕾<sup>1,2</sup>

<sup>1</sup>(中国科学院 计算机网络信息中心, 北京 100190)

<sup>2</sup>(中国科学院大学 计算机科学与技术学院, 北京 100049)

通讯作者: 杨小渝, E-mail: kxy@cnic.cn

**摘 要:** 许多新材料研发企业的研发设计、制备以及测试表征分处于不同部门, 导致材料制备工艺数据、测试表征数据以及研发设计数据呈离散化、碎片化状态, 形成信息孤岛, 且不能与材料计算有效地关联. 开发一个通用的, 能将材料测试表征、制备和服役数据深度融合的数据库系统, 进而基于数据驱动方法, 构建“成分-组织-工艺-性质”预测模型, 有着非常重要的意义. 然而, 材料种类和体系众多, 不同材料其测试表征方法以及制备方法多样性, 给开发一种通用的融合材料测试表征和制备工艺的数据库系统带来很大挑战. 基于此, 我们首先参照 CSTM 材料基因组数据通则, 提出了一个材料测试表征和制备工艺数据库框架, 进而基于无代码编程理念和语义 UI 模板, 研发了一个通用的材料测试表征和制备工艺数据库无代码构建系统 MatFusion. MatFusion 能适用于不同的材料体系, 帮助企业快速构建材料专用数据库, 并在液态金属测试表征和制备工艺专用数据库的构建中得到了应用.

**关键词:** 材料研发; 材料专用数据库; 语义 UI 模板; 无代码编程

引用格式: 杨小渝, 张蕾. 材料测试表征和制备工艺数据库无代码生成系统. 计算机系统应用, 2021, 30(12): 84-94. <http://www.c-s-a.org.cn/1003-3254/8236.html>

## Zero-Code Development System for Building Specific Material Characterization and Manufacturing Database

YANG Xiao-Yu<sup>1,2</sup>, ZHANG Lei<sup>1,2</sup>

<sup>1</sup>(Computer Network Information Center, Chinese Academy of Sciences, Beijing 100190, China)

<sup>2</sup>(School of Computer Science and Technology, University of Chinese Academy of Sciences, Beijing 100049, China)

**Abstract:** For many material companies, material design, manufacturing, and characterization are usually completed in different departments, and then the life-cycle data of materials are separately located. Also, material characterization data is presented in data files with their specific format of characterization instruments. The data of material properties and manufacturing, which is discretely located, is not appropriate for using a novel data-driven approach for material research and development. This initiates a need for a data infrastructure that can help to integrate these separately located material data to formulate a database system. However, it is challenging to develop such a data infrastructure due to diverse material characterization instruments and manufacturing processes. To address this need, we develop a general data infrastructure that can help to build a specific material characterization and manufacturing database by following the CSTM material genome engineering data principle and employing the zero-code development idea. In this study, we illustrate the design and development of the data infrastructure and use it to build a characterization and manufacturing database for liquid metal. It is demonstrated that the idea and data infrastructure for building a specific material characterization and manufacturing database are appropriate and useful.

**Key words:** material research and development; material-specific database; semantic UI template; zero-code development

① 基金项目: 云南科技厅稀贵金属材料基因工程专项 (202002AB080001-5)

Foundation item: Special Project for Rare Metal Material Genetic Engineering of Science and Technology Bureau, Yunnan Province (202002AB080001-5)

收稿时间: 2021-03-10; 修改时间: 2021-04-07; 采用时间: 2021-04-20

我们调研发现,对于很多新材料研发企业而言,其材料研发设计、制备以及测试表征分处于不同的部门,导致材料制备工艺数据、测试表征数据以及材料研发设计数据,呈离散化、碎片化状态,形成信息孤岛.测试表征人员要用到多种实验仪器来测量材料不同性能,但是不同表征设备输出的数据格式各不相同,大多保存在测试表征部门的电脑中.对于同样的物性,还会涉及到用不同测试设备进行表征.比如电导率的测试表征方法就有涡流法、U型管和平管3种方法;热导率的测试表征方法就有激光导热仪和导热系数测试两种方法.对材料制备工艺数据而言,或没有得到保存,或仅保存在生产部门的电脑中.而测试表征设备和制备工艺的多样性,导致数据记录方式也不尽相同,有的数据需要人工手动记录,有的数据以电子文档格式呈现.材料实验数据的上述特点,概括起来,就是“多源、异构”.这种测试表征和制备工艺分散的数据存储方式不便于材料研发数据共享,更不便于借助AI方法开展新材料设计<sup>[1,2]</sup>.

为了解决多源、异构的材料数据集中存储和共享,国内外已有一些材料数据库平台,旨在解决这个问题.其中最知名是国外的 Citrination<sup>[3,4]</sup>.它是一个开放的材料数据和分析云平台,允许用户上传自己的材料数据,开放共享,并基于此开展材料数据挖掘. Citrination 支持十多种文件格式(如 csv、cif、dat、fid、txt),文件上传后能自动被转换为 Citrination 定义的标准格式 PIF (Product Information File) 进行标准化存储.但是它不能很好地支持针对各类测试表征设备所获取数据的录入和集中管理存储,尤其是离散型数据的录入.国内主要有材料基因数据库平台 MgeData (<https://mgedata.cn/>), 它的数据录入要针对材料基因工程数据汇交,并不主要针对材料研发企业的制备和测试表征数据录入.

因此为了实现一个能帮助不同材料研发企业构建他们专用材料数据库的一个统一、通用的系统,从而便于企业在一个集成的环境下录入材料测试表征和制备工艺数据,本文结合材料基因工程数据通则标准,以及无代码编程理念,研发了一个材料测试表征和制备工艺数据库无代码生成系统,称之为 MatFusion. MatFusion 主要包括基于 CSTM 材料数据通则<sup>[5]</sup>的一个材料测试表征和制备工艺数据库框架,以及基于此材料数据库框架而开发的一个材料数据库无代码生成

系统. MatFusion 设计为一款无代码生成系统,在代码上提供抽象层,以代码为基础,允许用户以视觉方式构建数据录入的 UI 界面.这不仅有助于提高开发速度,也给用户带来了更大的自主权,使用户的需求很大程度地“自定义化”.

CSTM 指中关村材料实验数据联盟<sup>[6]</sup>,它于2019年8月颁布了材料基因工程数据通则,用于指导我国的材料基因工程.无代码编程是指一种无需代码基础,适合业务人员、IT开发及其他各类人员来组装和配置程序.无代码并不是指软件没有代码,软件的底层依旧是由很多代码组成,但是把传统需要写大量代码才能实现的功能组件化了<sup>[7,8]</sup>.用户可以在图形界面中,使用可视化建模的方式,来组装和配置应用程序,跳过所有的基础架构,只关注于使用应用模块来实现业务逻辑.

MatFusion 的最大创新在于,将 CSTM 材料基因工程数据通则和无代码开发理念相结合,开发了一个通用的材料专用数据库的无代码生成系统,用户仅需自定义语义 UI 模板,便能动态生成针对不同材料测试表征设备的数据录入页面. MatFusion 能够适用于不同的材料研发企业,解决了新材料研发企业存在的材料研发设计、制备以及测试表征分处于不同的部门而形成的“数据孤岛”问题.并将分散独立的测试表征和制备工艺数据融合,方便数据共享且有助于基于数据驱动的方法,帮助新材料研发.

本文介绍了 MatFusion 的设计和开发理念,并以液态金属测试表征和制备工艺离散型数据的录入为案例,介绍了 MatFusion 的使用.

## 1 CSTM 材料数据通则

要做到一种通用的解决方案,首先要遵循一个标准.中关村材料实验数据联盟 CSTM,于2019年8月颁布了材料基因工程数据通则,建立了适合材料基因工程需求的数据标准,规范数据产生过程中要收集的信息和遵循的格式,满足数据 FAIR 原则 (Findable, Accessible, Interoperable, Reusable)<sup>[9]</sup>. CSTM 是在中国工程院战略研究下,结合国家深化标准化工作改革方案,在国家标准委、工信部、中国工程院、中关村管委会等部门的支持下成立的.目前中关村材料实验数据联盟负责中国材料基因工程相关标准制定,以及其它材料研发和试验标准的制定.

CSTM 材料基因工程数据通则基于材料科学在数

据驱动模式下对数据的需求,将数据分为样本信息、源数据(未经处理的数据)与衍生数据(经分析处理得到的数据)三类,以操作(样品制备/表征/计算/数据处理)为条目单位,对每次操作分别赋予独立资源标识(根据国标 GB/T 32843 或 DOI)。每条数据收集与操作相关的元数据。元数据主要包括:方法、条件、结果、和科技资源标识。样品可以是实验产生的实物,也可以是经计算产生的虚拟物。同理,源数据可以来自于表征

或是直接测量,也可以通过模拟计算产生。

参照 CSTM 材料基因工程数据通则,本文提出了一个基于 CSTM 材料基因工程数据通则的材料数据库架构体系,如图 1 所示。从图中可以看到对于不同来源的数据,数据库要存储数据的样本信息、源数据以及它的衍生数据,同时对于样品的制备或者测试方法等都需要给出它的测试方法、条件、结果以及科技资源标识等信息。

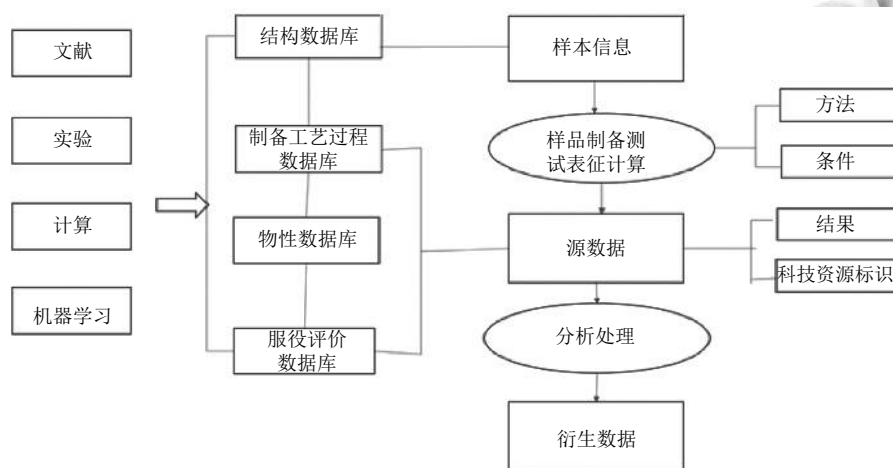


图 1 基于 CSTM 材料基因工程数据通则的材料数据库架构体系

## 2 材料数据库框架设计

基于图 1,本文提出了一个测试表征和制备工艺专用数据库的架构。从材料生命周期看,包括如下数据库表。由于不同材料或样品,采用的设备、工艺、方法等均不相同,本文提出了采用语义模板的方法,动态生成数据的采集页面。

- (1) 样本信息: 录入化学式、结构、成分等信息。
- (2) 制备工艺: 录入制备工艺相关信息。
- (3) 物相组成: 录入该样本物相组成等信息。
- (4) 微观组织: 录入该样本的微观组织信息。
- (5) 基本物性: 录入该样本的基本物性数据。
- (6) 服役性能: 录入该样本测试的服役性能数据。

从实验角度(制备和表征)来看,一般有组成元素,组成物相,制备工艺以及物性等。材料属性又有牌号和名称等。因此针对金属和无机非金属材料的一个制备表征的数据库架构如图 2 所示。主要的数据库表结构如下:

- 材料 (ID, 名称, 牌号, 化学表达式, ...)
- 组成元素 (ID, 材料 ID, 元素序号, 含量, ...)
- 元素属性 (ID, 组成元素 ID, 电负性, ...)

- 制备工艺 (ID, 材料 ID, 原料, 工艺, ...)
- 组成物相 (ID, 材料 ID, 名称, 含量, 类型, ...)
- 晶体结构 (ID, 物相 ID, 空间群, 晶格常数, 原子占位, ...)
- 物性 (ID, 材料 ID, 表征 ID, 温度条件, 尺寸条件, 其它条件, ...)

表征方法 (ID, 仪器, 操作, ...)

通过晶体结构可以与计算数据库建立起关联,从而实现制备表征数据库、晶体结构数据库与计算数据库的关联,从而真正实现材料基因工程所倡导的计算、制备和表征的一体化集成。

## 3 材料数据库框架无代码构建系统实现

### 3.1 材料测试表征和制备工艺数据库无代码生成系统架构

材料测试表征和制备工艺数据库无代码生成系统如图 3 所示。主要包括材料测试表征和制备工艺数据库,离散型数据录入,以及电子数据录入。材料测试表征和制备工艺数据库,分为成分信息,制备工艺,物相

组成, 组织结构、基本物性, 以及服役性能几个层面. 数据录入分为: 手工录入 (离散型) 和电子数据导入. 由

于手工录入数据和电子数据导入的特点不同, 因而数据录入方法不同.

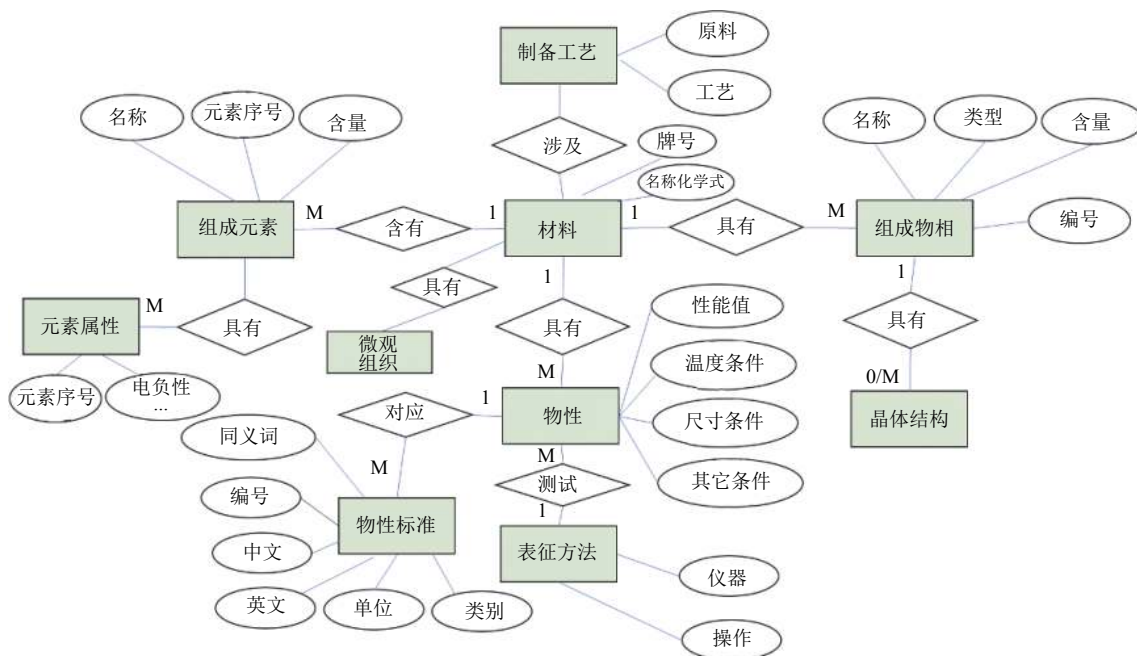


图2 材料数据库架构

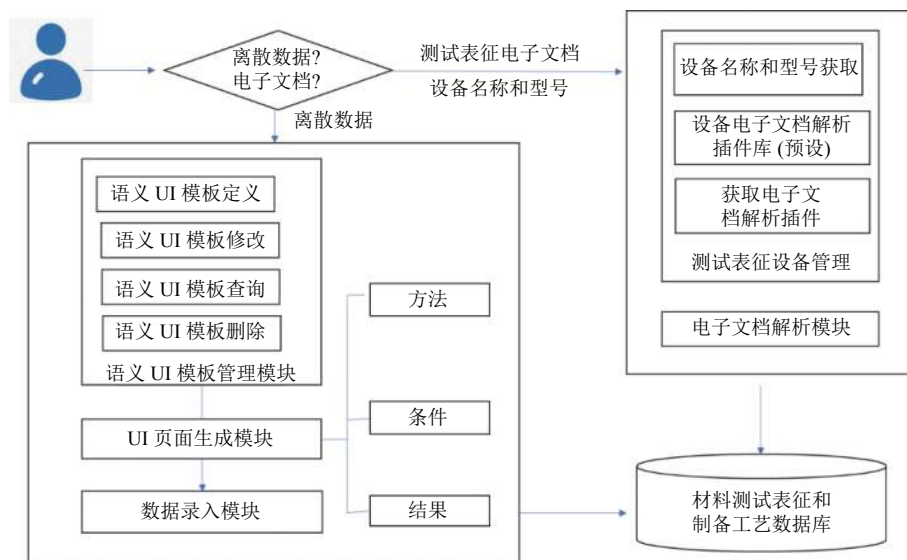


图3 系统架构图

### 3.1.1 手工录入数据的流程

(1) 根据材料体系, 通过语义 UI 模板, 定义该材料体系的测试方法, 测试条件和结果, 以及制备方法, 制备条件和结果.

(2) UI 页面生成模块, 调用上述定义的语义 UI 模

板, 生成相应的成分信息录入页面, 制备数据录入页面, 物相组成录入页面, 组织结构录入页面, 基本物性录入页面, 以及服役性能录入页面.

(3) 基于上述的各录入页面, 用户调用数据录入模块, 进行成分信息, 制备方法, 物相组成, 组织结构、基

本物性,以及服役性能的录入,进入材料测试表征和制备工艺数据库中。

### 3.1.2 电子数据导入的流程

(1) 用户输入设备名称和型号。

(2) 测试表征设备管理模块基于上述的设备名称和型号,通过设备名称和型号获取模块,获取输入设备名称和型号,从而在预设的设备电子文档解析插件库中,获取相匹配的电子文档解析插件。

(3) 用户输入该材料测试表征的电子文档结果,电子文档解析模块通过调用上述的电子文档解析插件,解析出该材料的测试表征数据,进入测试表征和制备工艺数据库中。

## 3.2 无代码生成系统设计

前文可知本系统数据库的设计遵循了CSTM规范,无论对于制备和表征,都只能通过方法、条件和结果去定义,这个“方法、条件和结果”如何与材料的成分信息、制备工艺、物相组成、微观组织、基本物化属性以及服役性能关联?以测试表征为例,不同材料的测试表征方法不同,即便是对于相同的材料的同一属性的测试表征方法都有所不同。例如对于密度这一物理性质来说,就有比重瓶法和密度天平法两种不同的测试方法。

基于上述考虑,本系统引入了语义UI模板的设计。即对材料的测试表征方法和制备工艺,预先定义一个通用的语义UI模板,定义测试或制备所涉及的方法、条件和结果。一旦语义UI模板定义好后,在用户进行材料表征和制备数据录入时,便可以选定义好的模板进行数据的录入,解决了数据录入时“通用”和“个性化”的问题。

### 3.2.1 语义UI模板构成元素及元数据

语义UI模板元素构成,根据上述讨论,遵循CSTM规范,语义UI模板的构成元素为:方法、条件和结果。条件又分为一级条件1、一级条件2等。每个一级条件下又可含有二级条件1、二级条件2等。

模板元素的元数据描述,如何定义方法、条件和结果呢?为了解决这一问题,本文提出了如下的元数据规范来描述它们,如图4所示。

从图4可以看出,方法由方法名和方法两个元数据定义。测试条件由条件名、条件值和条件值单位3个元数据描述;结果同样由结果名、结果值和结果值单位3个元数据描述。条件又分为一级条件、二级条件等。例如对测试条件温度来说,在温度这一条件下又分为外焰温度和内焰温度,此时温度就是一级条件,外焰温度和内焰温度则是二级条件。

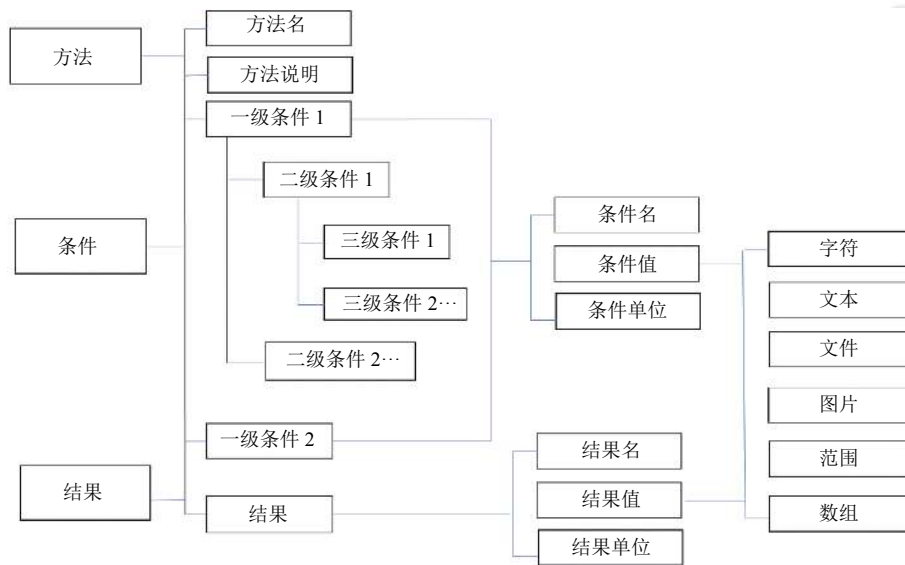


图4 UI模板的构成

无论是条件值或者结果值,都可以是字符类型、文本类型、文件类型、图片类型、范围值以及数组和矩阵类型。因此用户可以通过UI模板对层级关系以及

数据的类型进行自定义。

### 3.2.2 语义UI模板的实现

根据语义UI模板的构成要素和元数据规范,如何

描述该语义 UI 模板呢?

在实现技术上,可以使用文本或者 XML 或者 JSON 或者知识图谱等来描述对象,本系统中使用的是轻量级的 JSON.对于测试条件来说,用户只需定义测试条件名、测试条件值类型、测试条件单位,系统会自动将用户的输入以 JSON 的格式描述(目前用户暂只能以 JSON 格式输入).

例如用户要录入测试条件名为“温度”,其下又包含“内焰温度”和“外焰温度”两个二级条件,数值类型均为字符类型,单位均为“摄氏度”,并且外焰温度值为100.相关的 UI 模板语句如下所示:

```
{
  "name": "温度",
  "valueType": "object",
  "value": [{
    "name": "内焰温度",
    "unit": "摄氏度",
    "valueType": "string"
  }, {
    "name": "外焰温度",
    "unit": "摄氏度",
    "value": "100",
    "valueType": "string"
  }
]
```

本系统也对多种的录入数据格式进行了定义.例如在录入温度时可能需要录入最高温度、最低温度,此时在进行语义 UI 模板的填写时将 valueType 写为“range”即可.当前本系统支持的数据类型有字符类型、文本类型、矩阵、文件/图片、范围值,分别为“string”“text”“array”“file”“range”.之后也将会根据实际的使用情况进行数据类型的扩充.

### 3.2.3 模板数据库存储设计

由前文可知语义 UI 模板的构成元素为方法、条件和结果,以及相应的元数据描述.因此本系统设计了如下的数据库,用于存储语义 UI 模板的构成元素和元数据描述,如图5所示,数据库中“MethodTemplate”存储表用于存储模板类型、方法名和方法描述.其中模板类型指明是用于描述“结果”或者“条件”.“Condition”表用于存储条件的元数据描述.“Result”表用于存储结果的元数据描述.

### 3.2.4 语义 UI 模板管理

为了实现用户对语义 UI 模板的管理,本系统不仅提供了模板定义和查看功能,还提供了模板的修改和删除功能,这样用户不仅可以录入方法模板,也可以对已经录入的方法模板存在的不恰当的地方进行修改和删除,这让语义 UI 模板的使用更加灵活.

修改模板是用户对已添加好的方法模板进行的管理操作.当用户选定了某一个具体的测试方法后,点击“修改模板”按钮将会跳转到修改模板页面,页面的内容则是用户选择的该方法模板的具体内容.用户可以对方法名、方法描述、测试条件和测试结果进行修改,测试条件和测试结果的录入语句遵循前文提到过的语义 UI 模板的规范.当用户对需要修改的内容修改完毕后点击“保存”按钮将会触发相应的方法将用户修改后的内容更新到数据库中.

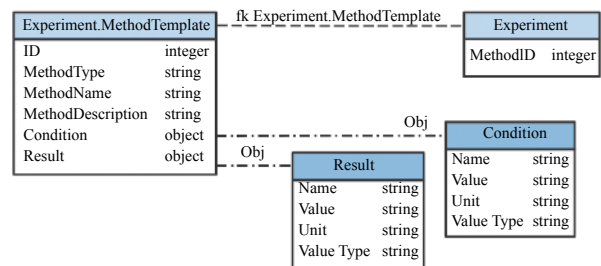


图5 模板数据库设计

删除模板是用户对已添加好的方法模板进行的管理操作.当用户选定了某一个具体的测试方法后,点击“删除模板”按钮将会触发相应的方法将用户选择的该方法模板从数据库中删除.

### 3.2.5 UI 页面生成

语义 UI 模板实现了用户在录入数据时对录入页面的自定义功能.当用户录入模板语言并保存时,本系统会调用模板解析模块对用户录入的模板信息进行解析;基于对模板的解析,调用页面生成模块最终生成数据录入页面.

用户录入数据时,当用户选择模板后,系统便会依次读取语义模板数据库中存储的信息,最终根据存储的“valueType”的值来生成相应的页面.系统会根据数据值的类型,字符类型、文本类型、矩阵、文件\图片、范围值,分别生成相应的页面.

- (1) 当数据类型为字符类型时页面如图6所示.
- (2) 当数据类型为文本类型时页面如图7所示.



图6 字符类型页面显示

(3) 当数据类型为矩阵、数组类型时页面如图8所示。

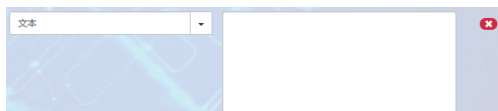


图7 文本类型页面显示



图8 矩阵、数组类型页面显示

(4) 当数据类型为文件、图片类型时页面如图9所示。

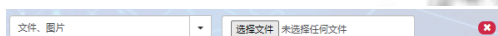


图9 文件、图片类型页面显示

(5) 当数据类型为范围值时页面如图10所示。

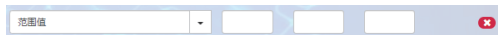


图10 范围值页面显示

### 3.2.6 语义UI模板的使用案例

接下来本文将举两个简单的例子来说明语义UI模板的使用。分别是制备工艺的“化学合成法”和磁滞回线的测试表征方法“振动样品磁强计”。

#### (1) 化学合成法制备工艺模板

“化学合成法”的制备条件为“温度”和“时间”，单位分别是“摄氏度”和“小时”，输入定义为字符型，则该方法的制备条件的模板语句为：

```
[{
  "name": "温度",
  "unit": "摄氏度",
  "valueType": "string"
}, {
  name": "时间",
  "unit": "小时",
  "valueType": "string"
}]
```

制备结果包括“颜色”和“状态”，同样按字符定义，定义的语义UI模板语句如下：

```
[{
  "name": "颜色",
```

```
"unit": "",
"valueType": "string"
}, {
  "name": "状态",
  "unit": "",
  "valueType": "string"
}]
```

一旦定义好后，该语义模板相关信息会存于数据库中。最终生成的基于化学合成法制备工艺的数据录入页面如图11所示。



图11 化学合成法UI页面

#### (2) 振动样品磁强计语义UI模板

磁滞回线的测试表征方法“振动样品磁强计”，其测试条件为“温度”，单位为“摄氏度”，输入定义为字符类型，则该测试表征方法的表征条件模板语句为：

```
[{
  "name": "温度",
  "unit": "摄氏度",
  "valueType": "string"
}]
```

表征结果为饱和磁化强度的一个图片文件，因此定义的语义UI模板如下：

```
[{
  "name": "结果文件",
  "valueType": "file"
}]
```

基于振动样品磁强计生成的数据录入界面如图12。

### 3.3 电子文档数据的录入

电子文档数据的录入，相对比较容易。电子文档数据的录入和测试表征设备密切相关。MatFusion针对常见的测试表征设备所产生的电子文档数据，预先开发了相应的解析模块。用户在录入数据时，只需输入测试表征设备、型号，以及关注的物性数据，然后上传该电子文档

文件,系统会自动解析电子文档,提取关键数据,存入数据库中.对于没有定义的测试表征设备,根据该电子文档格式,需要进行个性化的开发,集成到 MatFusion 中,如图 13.

#### 4 案例

本章以录入 Ga 基合金 Ga61In25Sn13Zn1 液态金属的离散型实验数据为例,说明 MatFusion 的使用.要录入的数据包括:成分信息、制备工艺、物相组成、组织结构、基本物性以及服役性能几个层面.要录入的样本数据如表 1 所示.

从表 1 分析可以看出,不同的物化性能,测试表征所用的仪器不同,测试表征条件,数据类型也不尽相同.而且,即便是同一种物化性质(如热导率),测试表征所用仪器也不相同.



图 12 振动样品磁强计 UI 页面



图 13 MatFusion 系统主页面

表 1 样本数据表

类别	待录入的信息	测试表征设备	条件	备注	
成分信息	化学成分占比	X 射线频谱分析(EDS)	室温	离散数据	
		X射线光电子能谱分析(XPS)			
制备工艺	真空冶炼法		50 °C, 1 h		
物相组成	XRD 图	X射线衍射法(XRD)	室温		
组织结构	SEM 图	扫描电镜(SEM)			
物理性质	密度	比重瓶法 密度天平法	室温	离散数据	
	粘度	旋转流变仪MCR502	室温	离散数据	
	接触角	高温高真空接触角测量仪 光学接触角测量仪	抽真空加热 室温	电子文档 电子文档	
电学性质	电导率	涡流法		离散数据	
		U型管		离散数据	
		平管		离散数据	
基本物性	力学性质	表面张力	接触角测量仪(悬滴法)		
		布氏硬度	布氏硬度计		
		抗拉强度	抗拉强度测试仪		
	热学性质	比热容	DSC差式扫描量热法(200F3)		离散数据
		熔点	DSC差式扫描量热法(200F3)		文档/离散
		熔化焓	DSC差式扫描量热法(200F3)		文档/离散
		热导率	激光导热仪(LFA467) 导热系数测试(DRL-III)		文档/离散 文档/离散
热膨胀系数	热机械分析仪(TMA 402F3)		文档/离散		
磁学性质	饱和磁化强度	振动样品磁强计		电子文档	
	磁滞回线	振动样品磁强计		电子文档	
服役性能				暂无	

采用本文提出的方法,录入的方法如 4.1, 4.2 节.

##### 4.1 录入页面

根据材料测试表征和制备工艺数据库,分为成分信息、制备工艺、物相组成、组织结构、基本物性以

及服役性能几个层面,开发出如图 14 的录入页面.

##### 4.2 离散数据的手工录入

###### (1) 成分信息录入

成分信息的测试方法,包括 X 射线频谱分析 (EDS),



以及 X 射线光电子能谱分析 (XPS) 等. 这里选择用 X 射线频谱分析 (EDS), 测试条件是在室温下, 结果为图像. 则对于 X 射线频谱分析方法录入测试表征数据所定义的语义 UI 模板关键语句如下 (JSON 语句格式).

测试方法: X 射线频谱分析.

测试条件主要是温度. 定义的语句如下:

```
[{
  "name": "温度",
  "unit": "摄氏度",
  "valueType": "string"
}]
```

测试结果为 EDS 样本, 以文件的形式呈现, 定义的语义 UI 模板如下:

```
[{
  "name": "EDS 样本",
  "unit": "",
  "valueType": "file"
}]
```

基于上述语义 UI 模板, 产生的数据录入 UI 页面如图 15 所示. 这时用户就可上传 X 射线频谱分析的图像文件.



图 14 电子数据录入界面



图 15 X 射线频谱分析 UI 页面生成

(2) 制备表征

制备表征的数据录入, 分为两个部分: 制备原料和制备方法. 制备原料一般包括: 原料、纯度、用量、厂家和批号, 而制备方法主要是文本描述. 制备原料的数据录入是基于静态页面. 制备方法和制备结果则是使

用语义 UI 模板进行定义. 对于真空冶炼法的数据录入, 所定义的语义 UI 模板关键语句如下.

制备方法: 真空冶炼法

制备条件为温度和时间, 定义的语句如下:

```
[{
  "name": "温度",
  "unit": "摄氏度",
  "valueType": "string"
}, {
  "name": "时间",
  "unit": "小时",
  "valueType": "string"
}]
```

制备结果为颜色和状态, 定义的语句如下:

```
[{
  "name": "颜色",
  "unit": "",
  "valueType": "string"
}, {
  "name": "状态",
  "unit": "",
  "valueType": "string"
}]
```

生成的制备工艺数据录入页面如图 16 所示.



图 16 制备工艺数据录入页面

(3) 物相组成

材料的物相组成可通过 X 射线衍射法 (XRD), 测试条件为室温, 其结果也是图像, 因此生成物相组成数据录入页面的 UI 定义语句如下:

测试方法: X 射线衍射法 (XRD)

测试条件:

```
[{
  "name": "温度",
```

```
"unit": "摄氏度",
"valueType": "string" ]}]
```

测试结果:

```
[{
  "name": "结果文件",
  "unit": "",
  "valueType": "file" }]
```

生成的物相组成数据录入页面如图 17 所示。



图 17 物相组成数据录入页面

#### (4) 微观组织

该类 Ga 基合金的微观组织用扫描电镜 (SEM), 在室温下表征, 表征结果为图像. 要支持该扫描电镜表征结果的录入, 则定义的针对扫描电镜测试表征数据录入的语义 UI 模板关键语句如下.

测试方法: 扫描电子显微镜

测试条件:

```
[{
  "name": "温度",
  "unit": "摄氏度",
  "valueType": "string"
}]
```

测试结果:

```
[{
  "name": "结果文件",
  "unit": "",
  "valueType": "file"
}]
```

通过上述语句定义的 UI 模板, 产生的扫描电镜图像录入页面如图 18 所示, 这时用户就可录入微观组织数据.



图 18 扫描电镜图像录入页面

#### (5) 基本物化属性数据的录入

这里以输入该类材料体系的比热容、熔点和熔化焓热学性质为例, 说明如何定义语义模板. 这些数据都是离散型. 定义生成比热容、熔点和熔化焓热学数据的 UI 录入页面语句如下:

测试方法: DSC 差式扫描量热仪 (200F3)

测试条件

```
[{
  "name": "温度",
  "unit": "摄氏度",
  "valueType": "string"
}]
```

测试结果:

```
[{
  "name": "比热熔",
  "unit": "",
  "valueType": "string"
}, {
  "name": "熔点",
  "unit": "摄氏度",
  "valueType": "file"
}, {
  "name": "熔化焓",
  "unit": "",
  "valueType": "file"
}]
```

生成的录入界面如图 19 所示。



图 19 基本物化属性数据录入界面

## 5 讨论

相较于其它材料基因数据库平台或系统, MatFusion 的最大特点在于提出了一种通用的方法, 用于整合材料研发企业分散的测试表征和制备工艺数据. 其主要创新点, 体现在以下几个方面.

(1) 提供了一种通用的材料测试表征和专用数据库架构, 解决了由于测试表设备和方法的不同带来的

数据存储多样性问题. 将新材料研发所需要的各种数据整合在一起, 方便实验人员之间数据的传递以及保证数据的安全性.

(2) 基于上述架构, 提出了一种材料专用数据库的无代码生成系统. 用户仅需自定义语义模板, 便能生成数据动态录入页面. 用户通过自定义语义模板, 动态的生成数据录入界面, 这为离散的数据的录入带来很大的便利. 同时模板的设计还包含了多种形式数据的录入, 这可以支持更多方法的录入, 使用户有更多的选择.

(3) 能便捷地帮助实现材料测试表征和制备工艺数据的融合. 该方案能将材料的分散独立的数据整合到一起方便了实验人员之前的数据共享, 而且有助于对新材料开展机器学习, 构建 QSPR 模型, 预测新型材料的性质, 加快新材料的研发.

目前的语义 UI 模板的定义, 主要通过轻量级的 JSON. 实际上, 语义 UI 模板的定义, 还可通过知识图谱的方式, 使得数据的录入变得更加智能化.

## 6 结束语

由于测试表征的多样性以及各种数据的独立存储, 导致新材料在研发时的各种数据等变得离散化、碎片化, 形成信息孤岛. 如何将这些离散且多样性的数据以一种统一的方式存储起来, 是新材料研发时亟待解决的问题. 因此为了实现在一个统一的、通用的环境下实现对不同材料种类和体系数据的录入, 本文结合材料基因工程数据通则, 以及无代码生成理念, 研发了要一个材料测试表征和制备工艺数据库无代码生成系统 MatFusion. 该系统提出了语义 UI 模板语言, 用户能够通过自己定义的语义 UI 模板实现对不同测试表征和制备工艺的数据录入界面的自定义.

我们将会对语义 UI 模板语言进一步优化, 通过更加友好的页面让用户定义描述方法、条件和结果的元数据, 使其拥有更加良好的交互界面和更加全面的功能,

满足更多“多源、异构”的材料数据融合、集成和共享.

## 参考文献

- 1 Horton MK, Dwaraknath S, Persson KA. Promises and perils of computational materials databases. *Nature Computational Science*, 2021, 1(1): 3–5. [doi: [10.1038/s43588-020-00016-5](https://doi.org/10.1038/s43588-020-00016-5)]
- 2 Yang XY, Wang ZG, Zhao XS, *et al.* MatCloud: A high-throughput computational infrastructure for integrated management of materials simulation, data and resources. *Computational Materials Science*, 2018, 146: 319–333. [doi: [10.1016/j.commatsci.2018.01.039](https://doi.org/10.1016/j.commatsci.2018.01.039)]
- 3 O'Mara J, Meredig B, Michel K. Materials data infrastructure: A case study of the citrination platform to examine data import, storage, and access. *JOM*, 2016, 68(8): 2031–2034. [doi: [10.1007/s11837-016-1984-0](https://doi.org/10.1007/s11837-016-1984-0)]
- 4 Saal JE, Oliynyk AO, Meredig B. Machine learning in materials discovery: Confirmed predictions and their underlying approaches. *Annual Review of Materials Research*, 2020, 50: 49–69. [doi: [10.1146/annurev-matsci-090319-010954](https://doi.org/10.1146/annurev-matsci-090319-010954)]
- 5 中关村材料试验技术联盟.T/CSTM 00120-2019 材料基因工程数据通则. <http://www.cstm.com.cn/article/details/ef49a444-80ca-4e71-99eb-e1e76c039d9f>. (2019-11-5)[2020-12-04].
- 6 中关村材料试验技术联盟. 中国材料实验与标准平台 CSTM. <http://www.cstm.com.cn/channel/details/home>. [2020-12-04].
- 7 Kiciman E, Melloul L, Fox A. Position summary: Towards zero-code service composition. *Proceedings Eighth Workshop on Hot Topics in Operating Systems*. Elmau: IEEE, 2001. 172. [doi: [10.1109/HOTOS.2001.990088](https://doi.org/10.1109/HOTOS.2001.990088)]
- 8 Sanchis R, García-Perales Ó, Fraile F, *et al.* Low-code as enabler of digital transformation in manufacturing industry. *Applied Sciences*, 2020, 10(1): 12.
- 9 Wilkinson MD, Dumontier M, Aalbersberg IJ, *et al.* The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 2016, 3: 160018. [doi: [10.1038/sdata.2016.18](https://doi.org/10.1038/sdata.2016.18)]