

产业链图谱构建指南

Guidelines for construction of industry chain graph

2023 - 10 - 10 发布

2023 - 11 - 10 实施

前 言

本标准按照GB/T 1.1—2020《标准化工作导则 第1部分：标准化文件的结构和起草规则》的规定起草。

请注意本标准中的某些内容可能涉及专利。本标准的发布机构不承担识别专利的责任。

本标准由浙江省经济和信息化厅提出、归口并组织实施。

本标准主要起草单位：浙江省工业和信息化研究院、浙江省统计研究与信息发布中心、杭州量知数据科技有限公司、杭州道生数据科技有限责任公司

本标准主要起草人：毛媛媛、张志宁、王国勇、凌云、罗延发、詹佳祥、娄万总、刘兵、庄越挺、邵健、张啸天、顾兴正、厉慧婷、郭慧敏、孙舟海、章哲。

产业链图谱构建指南

1 范围

本标准给出了产业链图谱构建的图谱要素、构建原则和构建程序。
本标准适用于制造业领域产业链图谱的构建，其他领域可参照执行。

2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本标准必不可少的条款。其中，注日期的引用文件，仅该日期对应的版本适用于本标准；不注日期的引用文件，其最新版本（包括所有修改单）适用于本标准。

GB/T 42131—2022 人工智能 知识图谱技术框架
YD/T 4044—2022 基于人工智能的知识图谱构建技术要求

3 术语和定义

GB/T 42131—2022 界定的以及下列术语和定义适用于本标准。

3.1

产业链 industry chain

一种或几种资源通过若干产业层次不断向下游产业转移直至到达消费者的路径。

[来源：GB/T 25469—2010，2.1]

3.2

产业链图谱 industry chain graph

以结构化形式描述的产业链知识元素及其联系的集合。

[来源：GB/T 42131—2022，3.6，有修改]

3.3

产业链图谱节点 industry chain node

产业链中在各个生产环节涉及的产品及服务。

4 要素与关系

4.1 基本要素

4.1.1 要素概述

产业链图谱由产业链节点和产业链关系构成。产业链节点是产业链图谱的基本组成单位；产业链关系由产业链上下位关系和产业链上下游关系组成，是串联产业链节点以构成产业链图谱的基本逻辑架构。

产业链图谱的要素与关系概念应符合GB/T 42131—2022中5.1~5.2的描述。

4.1.2 节点与边

4.1.2.1 节点

节点是产品与服务的抽象表达，对产业链节点的要素说明如下：

- a) 节点的实体类型是产品或服务；
- b) 节点应在产业链划分范围内；
- c) 节点支持不同层级的抽象；
- d) 节点具备关联企业、技术、人才等属性。

4.1.2.2 边

边是产业链关系的表示，对产业链关系的要素说明如下：

- a) 关系类型包括上下位关系和上下游关系；
- b) 上下位关系属于双向关系类型；
- c) 上下游关系属于单向关系类型。

4.2 上下位层级

上下位层级是纵向层级分类关系，指实体与其概念词，以及概念词和其上位概念词之间的联系。

上下位层级分类应细于国民经济行业分类的小类节点，层级数量一般不低于三层，具体层级数量视具体业务需求而定。

上下位层级涵盖的节点应涵盖基础产业链环节，且无常识性缺漏。

产业链上、中、下游属于基础分类框架，不单独作为一层级计入分类体系。

4.3 上下游关系

上下游关系是基于产业链工序的横向实体间关系。

上下游关系应提前明确关系类型和关系定义。本标准归纳形成了生产原料、生产配件、辅助材料、生产设备、辅助设备、加工工艺、技术服务和产品业务等上下游关系类型，具体释义见表1所示。

在产业链图谱实际应用中，上下游关系类型宜根据业务情况作整合或拆分调整，关系类型间不应重复。

表1 产业链图谱的上下游关系类型

上下游关系类型	上下游关系类型释义
生产原料	<ul style="list-style-type: none"> • 生产产品或提供服务时需要的最基本材料，构成主体的成分
生产配件	<ul style="list-style-type: none"> • 生产配件主要包括自制件和外购件，自制件分为主要零件和其它零件，外购件又分为外协外购件和标准件 • 主要配件指技术含量较高、高附加值，制造企业通常自行生产的各类非标准件，如齿轮类零件、轴类零件，箱体类零件等 • 其它配件指技术含量较低、低附加值，制造企业不一定自行生产的各类非标准件，如端盖类零件、视孔盖类零件等 • 外协外购件指制造企业不组织自行生产的其它零部件 • 标准件包括轴承、螺栓等

表1 产业链图谱的上下游关系类型（续）

上下游关系类型	上下游关系类型释义
辅助材料	<ul style="list-style-type: none"> 主体应用时的配套材料，比如：空调-制冷剂、干燥剂、墨盒-油墨、润滑油、机动车-原油 用于让主要材料发生变化，使主体具有区别于替代品的特性，比如：添加剂、助剂、催化剂 主体完工前应使用的其他非生产原料类材料，比如：包装材料
生产设备	<ul style="list-style-type: none"> 生产设备，包括加工设备、组装设备、必要的运输设备、包装设备等 用于运输或传输数据、水、电、货物的电缆、管道、运输工具等
辅助设备	<ul style="list-style-type: none"> 不参与主体的生产，仅用来维护生产环境或用来检测、记录生产状态的设备，比如：仪器仪表，检测设备，加湿器，干燥机，食品灭菌设备 参与生产过程，但不具有加工性能的设备，比如：容器、模型、模具、掩膜版 生产完成后用来保护主体性能，或使主体恢复原有性能的设备，比如：汽保设备 主体应用时的配套设备，比如：注射器、遥控器、充电桩、一次性电池 非物流、贸易类主体用到的运输工具，比如：远洋渔业产品-船舶及船舶装备 主体生产过程中借助的，并非用来生产主体的其他辅助设备，比如软件-服务器
加工工艺	<ul style="list-style-type: none"> 加工过程中的具体操作流程和步骤方法，比如：机械加工工艺、橡胶加工工艺等
技术服务	<ul style="list-style-type: none"> 为主体提供的技术服务，包括加工、物流、安装、检测、维修、咨询、设计、工程、集成等
产品业务	<ul style="list-style-type: none"> 主体作为中间产品，可直接供下游产品或服务集成应用。比如：锂离子电池作为中间产品，可供汽车电池系统所集成应用，因此锂离子电池和汽车电池系统之间为产品业务关系

上下游关系结构是根据产业链节点的上下游关联关系构建的图谱结构。产业链节点上下游关系示意图按图1。

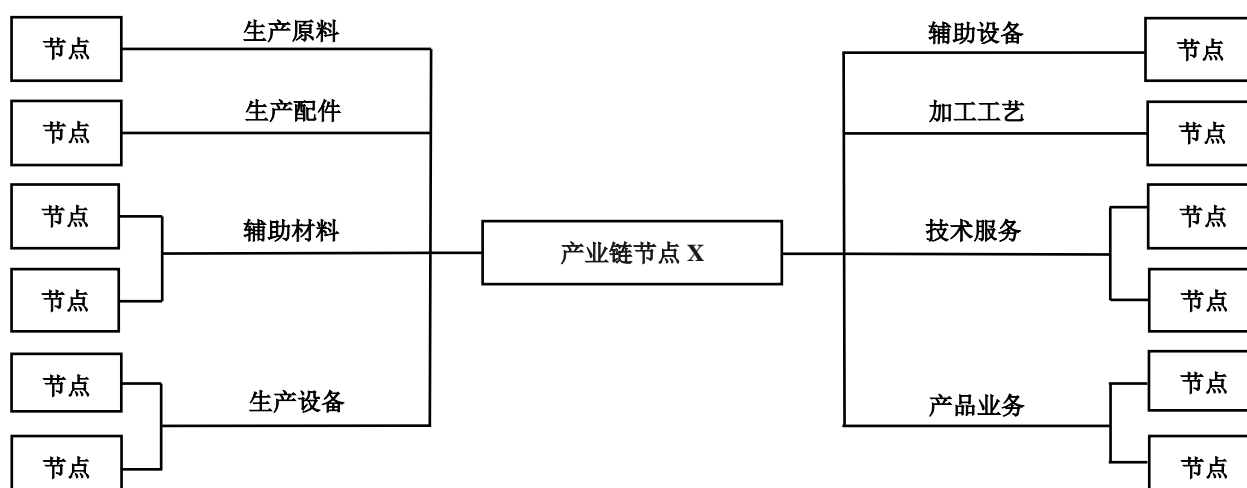


图1 产业链节点上下游关系示意图

5 构建原则

5.1 工艺路线清晰

产业链图谱构建前应根据业务需求明确技术路线，特殊产业链应基于技术路线差异分别构建，避免因技术路线不同导致的图谱结构变化引起歧义。

5.2 节点颗粒明确

本标准采用经济活动的同质性原则划分产业链，宜按照同一种经济活动的性质划分产业链。

本标准采用层级分类原则确定产业链节点等级，即按照符合业务需求且兼具科学性及逻辑性的分类方法，进一步细化大类节点，形成中类节点，以此类推实现节点的层层细分下钻，形成不重不漏的产业链上下位层级体系。

产业链图谱服务于具体业务场景，产业链节点的颗粒度设计应与实际业务需求相匹配，明确节点颗粒度目标，细化节点语义库语料词，使节点数据颗粒度细于国民经济行业分类的小类节点，保障语义识别与数据融合的精度。

5.3 节点层次分明

产业链图谱应有清晰的上下位层次结构：

- a) 产业链图谱中的所有节点应包含于图谱体系中的对应层次；
- b) 产业链图谱的层次分布应与产业链节点颗粒度相匹配；
- c) 产业链图谱层次应按照节点的上下位关系进行设定，以节点间的包含关系为判断准则，对于不同类别、型号、应用领域的产业链节点，应提取通用化节点放入对应上级层次，且同一类型的节点不应同时列入两个或两个以上的层次体系中；
- d) 产业链图谱的层次数量视具体情况而定，产业链图谱的层次不宜过多；
- e) 产业链图谱的层次设定应基于业务体系，建立相匹配的机器训练模型，以便于后续所抽取节点的上下位挂接能够满足业务需求。

5.4 关系类型完整

产业链图谱的上下游关系关联，应覆盖所有产品间的上下游关系，保障产业链图谱上下游关系的完整性。上下游关系具体可分为生产原料、生产配件、辅助材料、生产设备、辅助设备、加工工艺、技术服务和产品业务等上下游关联关系。

5.5 关系边界清晰

产业链图谱的上下游关联关系应边界明确，关系类型间无重复内容。

6 构建程序

6.1 基本原则

产业链图谱构建程序应符合GB/T 42131—2022中6.1~6.3、YD/T 4044—2022 6.1~6.7的流程。

6.2 数据采集

确定产业链节点和产业链关系的抽取数据源，整合上市公司披露信息、互联网信息、政府数据、招聘信息数据、采购信息数据、专利数据等多源信息渠道。

6.3 知识建模

完成产业链知识到图谱结构转化的模型积累，形成产业链知识加工逻辑规则，初步完成产业链图谱基础内容的构建。知识构建包括节点知识建模和关系知识建模。

知识建模流程通常包括：

- a) 训练产业链实体（关系）识别模型，基于节点（关系）抽取信息源中的节点（关系）概念，积累节点（关系）概念候选词；
- b) 建立产业链节点（关系）同义词语料库和节点（关系）同义词判断规则库，对节点（关系）候选词进行同义消歧和归类，得到节点（关系）概念与同义词的集合。

6.4 知识获取

汇聚非结构化或半结构化的产业原始数据获取图谱标准知识，实现产业链知识自动化构建，知识获取主要包括产业链实体获取。

知识获取流程通常包括：

- a) 应支持通过自然语言处理模型、光学字符识别模型等方式，从文档、图片、音频、视频、格式性文件等半结构化数据和非结构化数据获取结构化产业链实体数据，更新或连接到产业链图谱中；
- b) 应支持特定实体属性采集，从不同信息源中抽取数据。

6.5 知识融合

筛选多源异构的产业链实体知识，将产品抽取与归类得到的产业链上下位关系知识和产业链上下游关系知识相融合，实现实体自身的多层级分类和实体之间的关系连接。

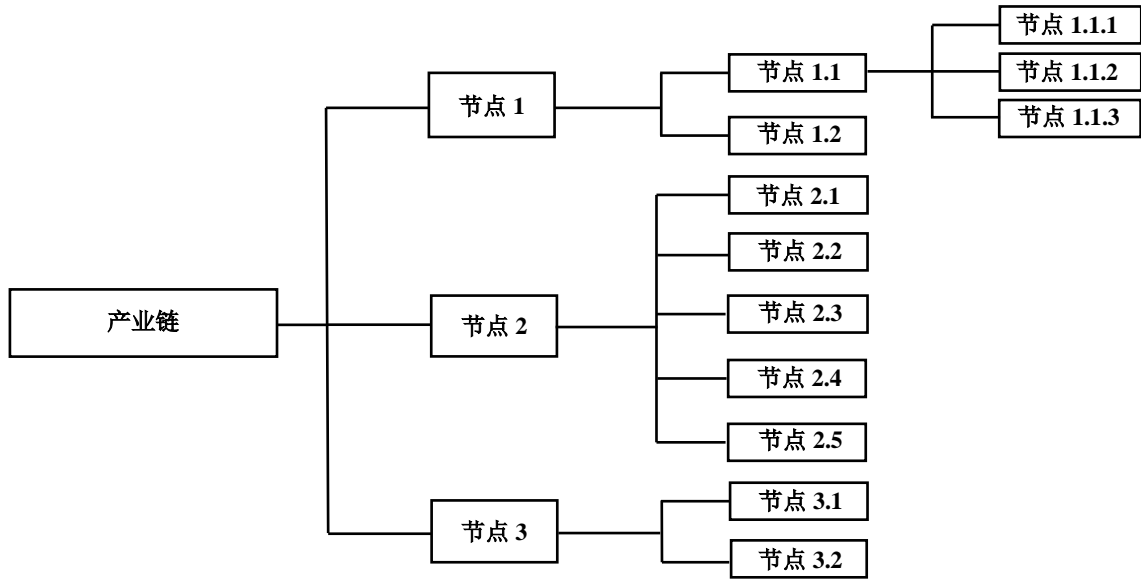
知识融合流程通常包括：

- a) 根据节点上下位关系词的语料和判断模型，识别节点概念间的上位节点与下位节点，并进行体系挂接；
- b) 建立关系表达同义语料库和规则库，识别节点之间的上下游关联关系类型，建立节点间的上下游关联关系。

6.6 图谱生成

融合上下游关系和上下位关系后，形成最终产业链图谱形态。产业链图谱不仅支持上下位层级和上下游关系的单独展示，也应支持纵横交错的产业链图谱融合展示，具体产业链图谱实例图见附录A。

产业链图谱上下位层级示意图按图2。



注：图中“节点1.1”、“节点1.2”表示节点1下一层级的第1种和第2种细分产品。

图2 产业链图谱上下位层级示意图

产业链图谱上下游关系示意图按图3。

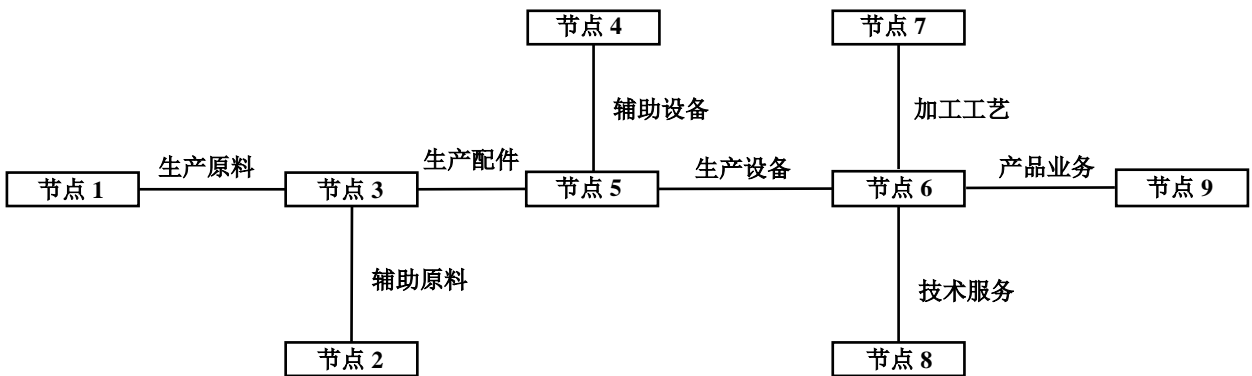


图3 产业链图谱上下游关系示意图

产业链图谱融合两种关系后应支持纵横交错的结构展示，其示意图按图4。

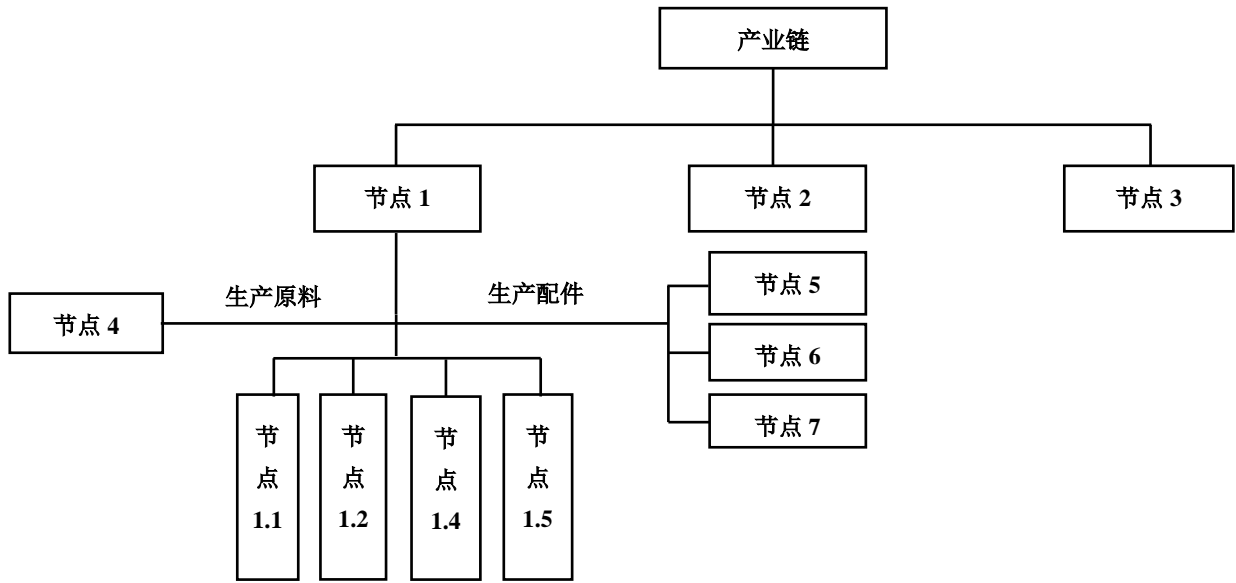


图4 产业链图谱融合关系示意图

产业链图谱在实际构建过程中，常以圆形节点和多属性边的力导图结构进行可视化呈现，其示意图按图5。

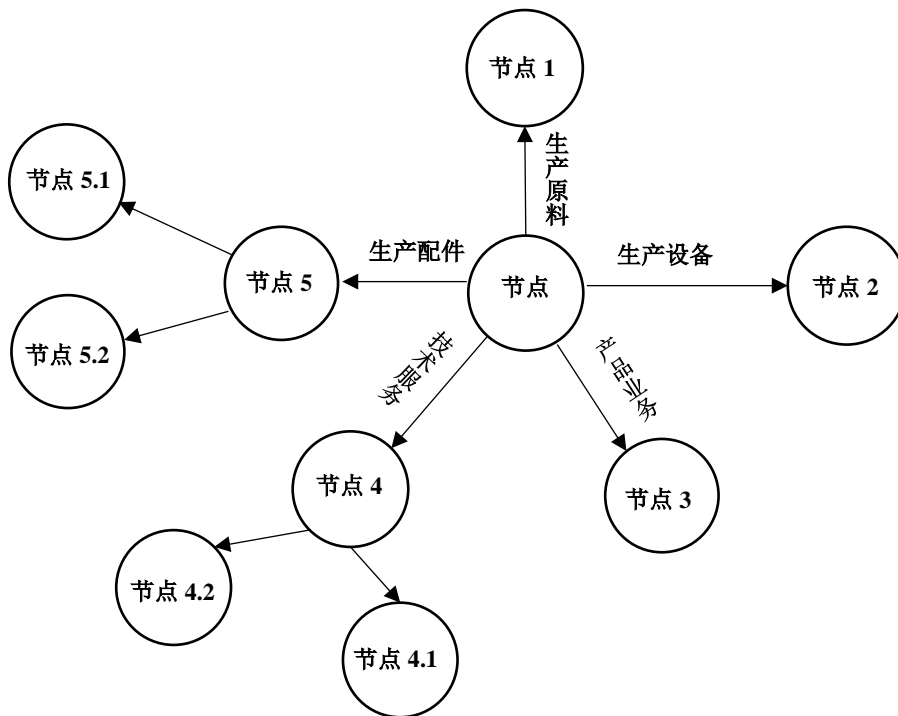


图5 产业链图谱融合关系力导图

6.7 人工校验

人工校验是为了保障产业链图谱的数据质量，主要包括专家评估和群智校验两种方式。

人工校验流程通常包括：

- a) 对节点抽取与归类结果进行人工审核和校验，沉淀更新节点知识语料库，提高产业链图谱上下位体系的精确度；
- b) 引入专家意见与群体的人工校验，对产业链图谱上下游关系的抽取与归类结果进行推理和重组；
- c) 结合人工智能校验算法，针对产业链图谱结构质量进行人机协同校验。

6.8 应用复核

根据产业链图谱的实际应用效果，对图谱产出的知识结果进行人工复核，对失准的知识概念进行链路式回溯，调整相应的实体与关系，完善产业链图谱的实体颗粒度、属性丰富度、关系丰富度、知识准确度。

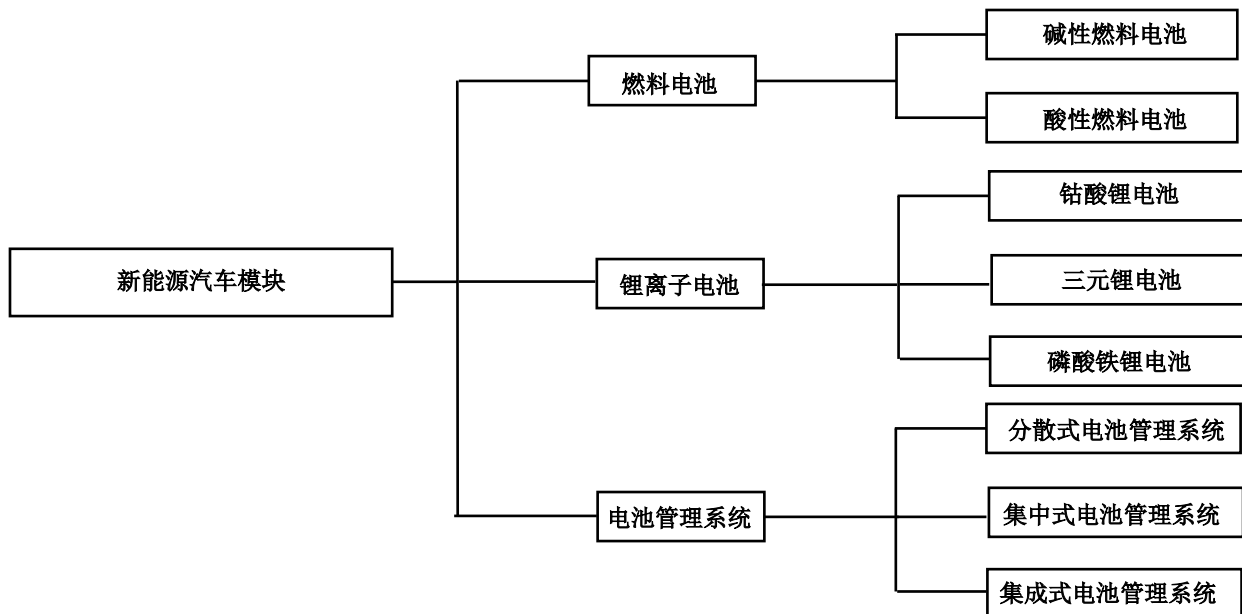
6.9 动态更新

依据现有产业链图谱框架，通过机器学习，结合人工审查机制，实现产业链图谱的节点及关系迭代升级。

附录 A
(资料性)
产业链图谱实例

A.1 产业链图谱上下位层级实例图

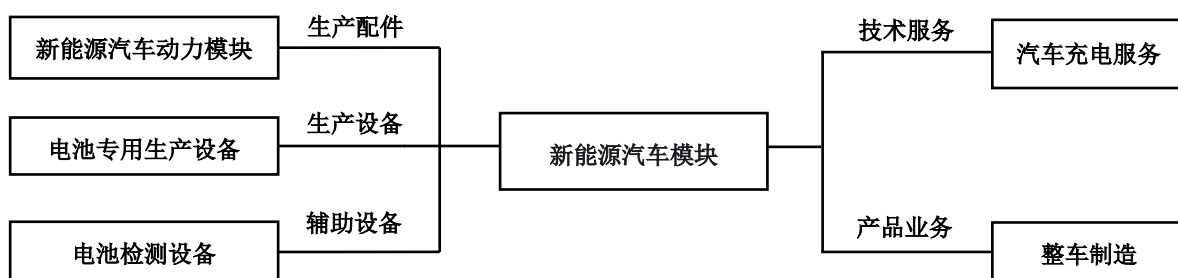
以新能源产业链为例的产业链图谱上下位层级实例见图A.1。



图A.1 产业链图谱上下位层级实例图

A.2 产业链图谱上下游关系实例图

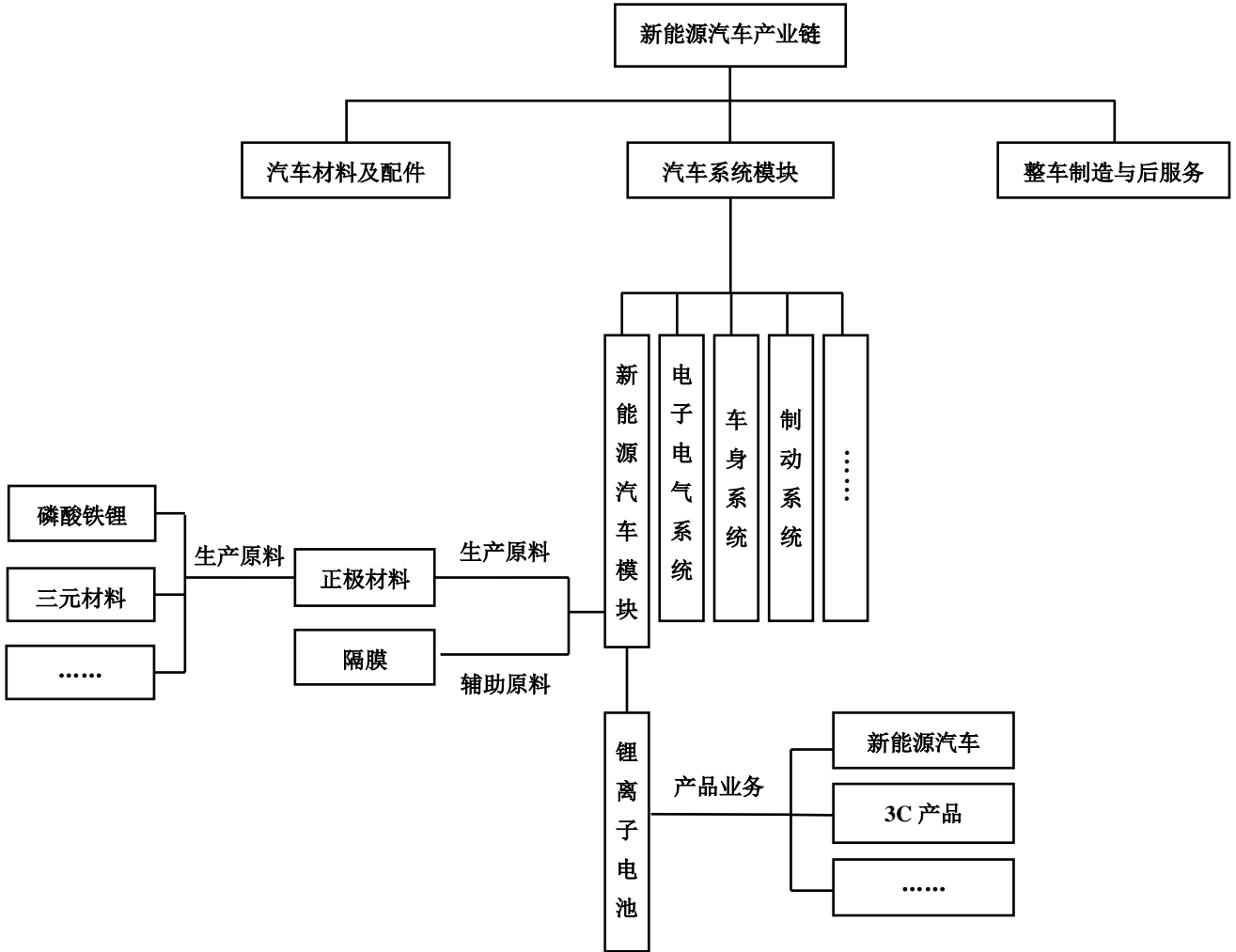
以新能源产业链为例的产业链图谱上下游关系实例见图A.2。



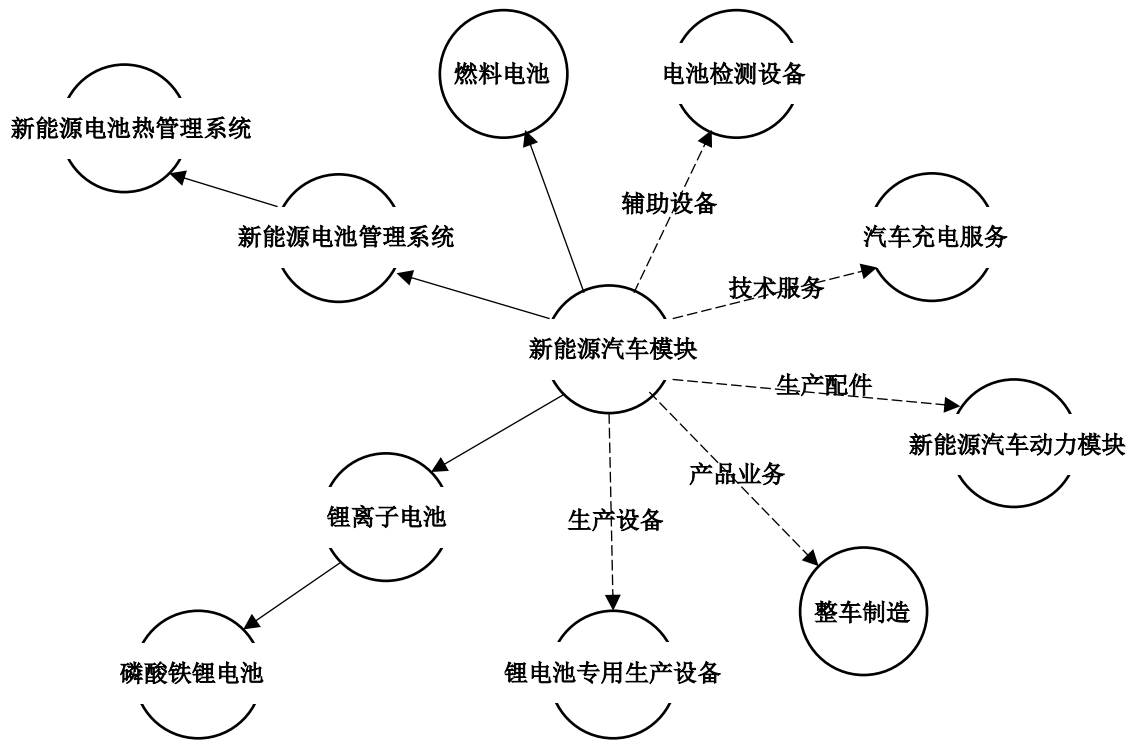
图A.2 产业链图谱上下游关系实例图

A.3 产业链图谱融合关系实例图

以新能源产业链为例的产业链图谱融合关系实例见图A.3.1，力导图结构实例见图A.3.2。



图A.3.1 产业链图谱融合关系实例图



图A. 3.2 产业链图谱融合关系力导图