



(12) 发明专利申请

(10) 申请公布号 CN 112966924 A

(43) 申请公布日 2021.06.15

(21) 申请号 202110228471.8

(22) 申请日 2021.03.02

(71) 申请人 杭州全视软件有限公司

地址 310011 浙江省杭州市拱墅区祥宏路5号5幢2层

(72) 发明人 裘灵 杨璐翠 张少原

(74) 专利代理机构 浙江千克知识产权代理有限公司 33246

代理人 周希良

(51) Int. Cl.

G06Q 10/06 (2012.01)

G06N 5/02 (2006.01)

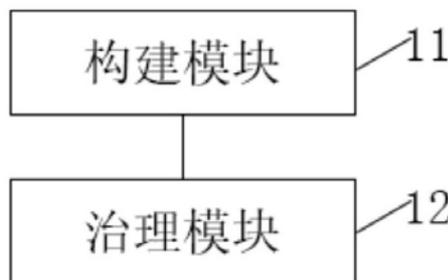
权利要求书1页 说明书7页 附图2页

(54) 发明名称

一种基于风险图谱的数据治理系统及方法

(57) 摘要

本发明公开一种基于风险图谱的数据治理系统及方法,其中涉及的一种基于风险图谱的数据治理系统,包括:构建模块,用于构建基于风险知识、风险信息、风险数据的风险图谱;治理模块,用于根据构建的风险图谱对待治理数据进行处理。本发明通过设计数据、信息、知识分层的图谱,通过知识层对信息层、信息层对数据层的图谱约束规则,结合数据(数据类及类规则)间、信息(信息类及类规则)间的图谱组合规则,形成数据治理的管理机制。形成了业务治理和数据治理的高度统一,可以为更好的治理认知提供方向,从而实现数据治理真正目标:业务创新。



1. 一种基于风险图谱的数据治理系统,其特征在于,包括:
构建模块,用于构建基于风险知识、风险信息、风险数据的风险图谱;
治理模块,用于根据构建的风险图谱对待治理数据进行处理。
2. 根据权利要求1所述的一种基于风险图谱的数据治理系统,其特征在于,所述构建模块具体包括:
风险数据模块,用于获取风险数据相对应的数据因子,并根据获取的数据因子构建风险数据图谱;
风险信息模块,用于根据获取的数据因子生成风险信息,并根据风险数据图谱以及生成的风险信息构建风险信息图谱;
风险知识模块,用于根据生成的风险信息确定风险知识,并根据风险信息图谱以及确定的风险知识构建风险知识图谱。
3. 根据权利要求2所述的一种基于风险图谱的数据治理系统,其特征在于,所述风险知识模块中确定的风险知识包括人的风险知识集、管理的风险知识集、物的风险知识集、环境的风险知识集。
4. 根据权利要求2所述的一种基于风险图谱的数据治理系统,其特征在于,所述风险信息模块中还包括对风险信息进行扩展处理。
5. 根据权利要求2所述的一种基于风险图谱的数据治理系统,其特征在于,所述治理模块中对待治理数据进行处理具体为通过风险知识对风险信息、风险信息对风险数据的图谱进行约束对待治理数据进行处理。
6. 一种基于风险图谱的数据治理方法,其特征在于,包括:
S1. 构建基于风险知识、风险信息、风险数据的风险图谱;
S2. 根据构建的风险图谱对待治理数据进行处理。
7. 根据权利要求6所述的一种基于风险图谱的数据治理方法,其特征在于,所述步骤S1具体包括:
S11. 获取风险数据相对应的数据因子,并根据获取的数据因子构建风险数据图谱;
S12. 根据获取的数据因子生成风险信息,并根据风险数据图谱以及生成的风险信息构建风险信息图谱;
S13. 根据生成的风险信息确定风险知识,并根据风险信息图谱以及确定的风险知识构建风险知识图谱。
8. 根据权利要求7所述的一种基于风险图谱的数据治理方法,其特征在于,所述步骤S13中确定的风险知识包括人的风险知识集、管理的风险知识集、物的风险知识集、环境的风险知识集。
9. 根据权利要求7所述的一种基于风险图谱的数据治理方法,其特征在于,所述步骤S12中还包括对风险信息进行扩展处理。
10. 根据权利要求7所述的一种基于风险图谱的数据治理方法,其特征在于,所述步骤S2中对待治理数据进行处理具体为通过风险知识对风险信息、风险信息对风险数据的图谱进行约束对待治理数据进行处理。

一种基于风险图谱的数据治理系统及方法

技术领域

[0001] 本发明涉及数据治理技术领域,尤其涉及一种基于风险图谱的数据治理系统及方法。

背景技术

[0002] 随着社会的发展,社会在从失范到规范、从混乱到有序的过程中社会问题和社会风险会集中迸发,这也会严重影响单位业务。

[0003] 数据已成为企事业(政府)的重要资产,也是企事业竞争力的重要武器。企业业务的成功,需要及时、完整和准确的数据提供支持。但经常能看到这样的现象:企业中几个部门的主管在汇报工作时,同一指标的数据不尽相同,甚至可能截然相反,到底哪个数据是正确的数据?无从确认。造成这种现象的原因很多,比如数据质量问题、统计口径问题,等等。如何确保能够为重要业务提供及时、相关和可信的数据?

[0004] 当前数据治理主要是遵循DAMA(国际数据管理协会)的规范,数据治理是对数据资产管理行使权力和控制的活动集合,也就说数据治理并不是一个简单的行为动作,而应该是一个形成体系的管理。包括所有核心业务系统、运营数据储存或数据仓库、ECIF(企业客户信息整合)、风险控制以及很多其他的系统,需要技术和“软内容”平台提供支撑,涵盖数据战略、元数据、数据质量、数据安全、主数据与参考数据、数据仓库与商业智能、数据集成、数据操作、文件和内容等领域,最终实现所有数据的全方位监管,实现端到端的数据梳理和管理,确保数据的有效性、可访问性、高质量、一致性、可审计和安全性。

[0005] 所以数据治理过程复杂、成本高,如果仅仅是实现数据资产,目前来看治理的效果这个数据资产都有不同时期的认知,所以“数据资产”只能一直提升,即该资产“折旧”速度极快。

[0006] 高投入需要有高的回报,按现有的机制难以有足够大的回报,所以数据治理需要融入到业务中,其治理水平是为了更好的业务风险管理,以及实现业务创新,以谋求更大的利益回报。

[0007] 业务过程会涉及到各种风险问题,如在消防火警中要解决的风险问题包括火警点位重复、火警点位位置错误、火警有规律的误报等等,这些问题是随着业务认知提升的,而不仅仅是数据管理部门能解决的问题,但是这些问题需要从数据治理的角度来规范并责任化,目前数据治理更多是从数据采集、清洗、转换、质量控制等来进行治理,没有形成有效的业务信息和业务知识。从而无法从业务的认知层面实现对数据级、信息级、知识级的层层规范。所以难以形成业务的认知体系构建和认知创新的扩展。

[0008] 因此,针对现有技术的缺陷,提出一种基于风险图谱的数据治理系统及方法。

发明内容

[0009] 本发明的目的是针对现有技术的缺陷,提供了一种基于风险图谱的数据治理系统及方法。

- [0010] 为了实现以上目的,本发明采用以下技术方案:
- [0011] 一种基于风险图谱的数据治理系统,包括:
- [0012] 构建模块,用于构建基于风险知识、风险信息、风险数据的风险图谱;
- [0013] 治理模块,用于根据构建的风险图谱对待治理数据进行处理。
- [0014] 进一步的,所述构建模块具体包括:
- [0015] 风险数据模块,用于获取风险数据相对应的数据因子,并根据获取的数据因子构建风险数据图谱;
- [0016] 风险信息模块,用于根据获取的数据因子生成风险信息,并根据风险数据图谱以及生成的风险信息构建风险信息图谱;
- [0017] 风险知识模块,用于根据生成的风险信息确定风险知识,并根据风险信息图谱以及确定的风险知识构建风险知识图谱。
- [0018] 进一步的,所述风险知识模块中确定的风险知识包括人的风险知识集、管理的风险知识集、物的风险知识集、环境的风险知识集。
- [0019] 进一步的,所述风险信息模块中还包括对风险信息进行扩展处理。
- [0020] 进一步的,所述治理模块中对待治理数据进行处理具体为通过风险知识对风险信息、风险信息对风险数据的图谱进行约束对待治理数据进行处理。
- [0021] 相应的,还提供一种基于风险图谱的数据治理方法,包括:
- [0022] S1. 构建基于风险知识、风险信息、风险数据的风险图谱;
- [0023] S2. 根据构建的风险图谱对待治理数据进行处理。
- [0024] 进一步的,所述步骤S1具体包括:
- [0025] S11. 获取风险数据相对应的数据因子,并根据获取的数据因子构建风险数据图谱;
- [0026] S12. 根据获取的数据因子生成风险信息,并根据风险数据图谱以及生成的风险信息构建风险信息图谱;
- [0027] S13. 根据生成的风险信息确定风险知识,并根据风险信息图谱以及确定的风险知识构建风险知识图谱。
- [0028] 进一步的,所述步骤S13中确定的风险知识包括人的风险知识集、管理的风险知识集、物的风险知识集、环境的风险知识集。
- [0029] 进一步的,所述步骤S12中还包括对风险信息进行扩展处理。
- [0030] 进一步的,所述步骤S2中对待治理数据进行处理具体为通过风险知识对风险信息、风险信息对风险数据的图谱进行约束对待治理数据进行处理。
- [0031] 与现有技术相比,本发明通过设计数据、信息、知识分层的图谱,通过知识层对信息层、信息层对数据层的图谱约束规则,结合数据(数据类及类规则)间、信息(信息类及类规则)间的图谱组合规则,形成数据治理的管理机制。形成了业务治理和数据治理的高度统一,可以为更好的治理认知提供方向,从而实现数据治理真正目标:业务创新。且通过知识层、信息层到数据层的知识图谱式的双向逐级治理,可以更全面、有效地进行数据治理,确保知识的价值。

附图说明

- [0032] 图1是实施例一提供的一种基于风险图谱的数据治理系统结构图；
[0033] 图2是实施例一提供的风险图谱示意图；
[0034] 图3是实施例四提供的高校治理结构示意图。

具体实施方式

[0035] 以下通过特定的具体实例说明本发明的实施方式,本领域技术人员可由本说明书所揭露的内容轻易地了解本发明的其他优点与功效。本发明还可以通过另外不同的具体实施方式加以实施或应用,本说明书中的各项细节也可以基于不同观点与应用,在没有背离本发明的精神下进行各种修饰或改变。需说明的是,在不冲突的情况下,以下实施例及实施例中的特征可以相互组合。

[0036] 本发明的目的是针对现有技术的缺陷,提供一种基于风险图谱的数据治理系统及方法。

[0037] 实施例一

[0038] 本实施例提供一种基于风险图谱的数据治理系统,如图1所示,包括:

[0039] 构建模块11,用于构建基于风险知识、风险信息、风险数据的风险图谱;

[0040] 治理模块12,用于根据构建的风险图谱对待治理数据进行处理。

[0041] 在构建模块11中,构建基于风险知识、风险信息、风险数据的风险图谱。

[0042] 风险图谱主要为数据治理的管理机制,以实现待治理数据的治理。

[0043] 具体包括:风险知识模块(知识体系到风险知识四个子集,形成风险知识树)、风险信息模块(风险信息细化,风险信息之间形成图谱)、风险数据模块(从信息到数据形成更复杂的图谱),如图2所示。

[0044] 风险数据模块,用于获取风险数据相对应的数据因子,并根据获取的数据因子构建风险数据图谱;

[0045] 获取要形成信息的数据因子,包括数据1、数据2、数据3……数据n。

[0046] 根据获取的数据形成风险数据相关的图谱。

[0047] 在本实施例中,风险数据包括数据类、数据对象等,其中数据类型如:火灾探测器,数据对象如:房间编码为0101002的火灾探测器。

[0048] 风险信息模块,用于根据获取的数据因子生成风险信息,并根据风险数据图谱以及生成的风险信息构建风险信息图谱;

[0049] 风险信息是基于风险数据而生成的,如:风险数据为探测器编码,则风险信息可以为探测器编码对应的信息。

[0050] 在本实施例中,风险数据生成风险信息中包括风险数据本身的信息,还包括风险数据的词经过名词、形容词、谓词等方式组合生成的风险信息。

[0051] 当根据数据生成信息后即产生了相应的职责。

[0052] 需要说明的是,风险信息中包括与风险数据相对应的信息,也包括基于风险数据穷举、扩展后生成的信息。

[0053] 风险知识模块,用于根据生成的风险信息确定风险知识,并根据风险信息图谱以及确定的风险知识构建风险知识图谱。

[0054] 风险知识模块包括多个分项的风险知识,如人的风险知识集、管理的风险知识集、物的风险知识集、环境的风险知识集。

[0055] 人的风险知识集包括人需要处理的与风险信息相关的事件,且通过风险图谱还可以确认相关人员处理事件的结果,将责任到人。

[0056] 物的风险知识集包括产品本身的相关事件。

[0057] 管理的风险知识集包括管理的人员、产品等的相关事件。

[0058] 环境的风险知识集包括天气、周围环境等的相关事件。

[0059] 其中,一个风险信息可以对应一个或多个分项的风险知识。

[0060] 在本实施例中,数据、信息、知识均是可以进行扩展的,如扩展数据就会组成新的信息;扩展信息就会生成新的数据;扩展知识就会生成新的信息。

[0061] 在治理模块12中,根据构建的风险图谱对待治理数据进行处理。

[0062] 通过设计数据、信息、知识分层的图谱,通过知识层对信息层、信息层对数据层的图谱约束规则,结合数据间、信息间的图谱组合规则,形成数据治理的管理机制。

[0063] 在本实施例中,通过知识层、信息层到数据层的知识图谱式的双向逐级治理,可以更全面、有效地进行数据治理,确保知识的价值和信息应用的可扩展性,以确保数据治理和单位治理的高度融合。

[0064] 实施例二

[0065] 本实施例提供一种基于风险图谱的数据治理系统与实施例一的不同之处在于:

[0066] 本实施例以火警报警为例进行说明。

[0067] 在风险数据图谱中,涉及到的风险数据类包括探测点位置、探测器编码、探测器信号、探测装置、探测信号传输线、值班人员、巡逻人员、维保人员、监管人员、管理人员、联动装置等。

[0068] 在风险信息图谱中,经过风险数据类生成的风险信息类包括以下数据的有效组合:探测器编码、信号有触发;信号有触发、传输线;传输线、值班人员;巡逻人员、维保人员;值班人员、联动装置;值班人员、监管人员;监管人员、管理人员等。

[0069] 经过风险数据拓展生成的风险信息包括:重要数据的采集、设置、录入、验证、调整相应的责任人;验证人的验证周期;故障的修复周期。

[0070] 在风险知识图谱中,总的风险知识包括让每一次火警都处理及时、正确的技术管理实现模式。

[0071] 分项的风险知识:每一次火警,人的风险可控;每一次火警,物的风险可接受;每一次火警,环境的风险可控;每一次火警,管理的风险可控。

[0072] 人的风险可控:值班人员及时通知巡查人员现场确认;值班人员及时接收巡查人员确认情况;值班人员正确处理确认情况;等等。

[0073] 物的风险可控:火灾报警主机正确联机;火警及时上报;火警正确定位;火警的误报率可接受(部分也是人的风险);等等。

[0074] 环境的风险可控:大旱、大雪等恶劣气候可以正常;一些特殊环境:海边大雾、中药热灸房等导致误报增加;等等。

[0075] 管理的风险可控:人的培训是否到位;一旦有额外风险时有没有应付机制;(这里管理的风险有不少也是人的风险);同义词、近义词信息管理;等等。

[0076] 在火警报警中,风险知识对风险信息的规则有:每一次火警都要从探测点及时地反映到值班人员处,要求每个探测点位置准确;火警无法及时从探测点传到值班人员处,则需要通知维保人员来检修,要能在规定时间内完好有效;设计人员需要让每一个探测器编码唯一对应到探测点位置上;值班人员要对每一个信息进行处理,要在规定的时间内通知巡逻人员,并监督巡逻人员在规定时间内确认;巡逻人员要在规定的时间内确认等等。

[0077] 风险信息对风险数据的规则可以按上面的规则分解到对风险数据的采集规则:火警探测点状况;探测装置状况;探测信息传输线状况;维保人员状况;维修进度;所有探测器编码位置唯一、准确;值班人员情况;通知时间;确认时间巡逻人员情况;确认时间等。对不同数据如名词、形容词、谓词等的信息组合可以用语义组合的机制。

[0078] 最终形成的设计规则具体为:火警探测点的误报控制;探测装置的报警和传输;信息传输的双向反馈机制;维保人员维修能力、服务能力;维修到位的定时检测和维保人员主动登记;探测器编码唯一或主机号、探测器编码唯一;探测器位置有定位机制或纠错机制(这一点有难度);当前值班人员的姓名、平时值班的绩效考核情况;每一次火警时间、有效通知时间、有效确认时间;巡逻人员的姓名、平时巡逻的绩效考核情况;每一次火警通知时间(要求确认,要求有具体位置)、有效确认时间。

[0079] 对于过程中出现的问题(隐患、故障或事故),可以认为是信息处理不足,需要对验证数据类(信息),同时需要验证相关的数据。比如某一个房间的火警出现,但是值班人员的界面上没显示具体位置,即没通知正确,就验证数据类信息:每一个探测器是否都正确对应对了位置,结果发现该房间编号为01010202的探测器没对应好“楼宇名、楼层和房间位置”,再验证当时输入环节、验证环节的相关人员,并要求对其他负责的数据都进行验证,同时可以通过管理人员、监管人员对主体责任人进行问责或追责。

[0080] 对于以上引入这些规则后,就可以更好地创新了,对信息进行指标化,比如信息传输路线完好,可以安装检测器,在每一个回路末端安装一个试验装置,每一小时都要同步一次数据,如果后端没有及时收到,即认为有故障;比如不同人员采用的名词不一样,即可以按同义词、近义词集来进行工具自动规范所用名词,如火灾探测器、烟感、烟雾感应点都是同义词,温感则是烟感的近义词,但是火灾探测器的同义词;比如值班人员机器化,直接把这些信息有效处理机制的机器化经验证后,即可替代。

[0081] 本实施例从数据到信息产生了责任,数据的验证也是因为信息需要验证,需要产生者负责、需要每一个调整者负责;从而信息也进行了完善;没有责任人的所有数据。在数据治理上也通过风险信息和风险知识实现了业务的高度融合,对业务要求有更好、更精细的技术认知,从而为今后的业务创新提供了全面的提升机制。

[0082] 实施例三

[0083] 本实施例提供一种基于风险图谱的数据治理系统与实施例一的不同之处在于:

[0084] 本实施例以环境治理为例进行说明。

[0085] 在风险数据图谱中,涉及到的风险数据包括本单位的污染源排放导致的可能大的环境事故(不同时间点污染源检测浓度、区域、扩展通道)等。

[0086] 在风险信息图谱中,经过风险数据生成的风险信息包括各环境要素(大气、水、土壤、海洋、核、固体废物与化学品、噪声、自然生态)、污染源、污染治理设施、巡查人员、值班人员、各岗位等。示例:水中重金属超标,值班人员的处理是否到位。

- [0087] 在风险知识图谱中,总的风险知识包括确保环境治理生态化。
- [0088] 分项的治理风险知识:确保人的环境治理风险到位;确保物的环境治理风险到位;确保环境影响的治理风险到位;确保管理的环境治理风险到位等。
- [0089] 确保人的环境治理风险到位:执法人员对单位的监管到位、单位对辖区的巡查到位、单位值班人员对辖区的报警管理到位等。
- [0090] 确保物的环境治理风险到位:污染源的污染物监测到位、大气环境的监测到位、水环境的监测到位、土壤环境的监测到位等。
- [0091] 确保环境影响的治理风险到位:地震、异常天气导致的环境异常等。
- [0092] 确保管理的环境治理风险到位:人的培训是否到位;一旦有额外风险时有没有应付机制;(这里管理的风险有不少也是人的风险)等。
- [0093] 在本实施例中,通过知识层、信息层到数据层的知识图谱式的双向逐级治理,可以更全面、有效地进行数据治理,确保知识的价值。
- [0094] 实施例四
- [0095] 本实施例提供一种基于风险图谱的数据治理系统与实施例一的不同之处在于:
- [0096] 本实施例以单位管理软件数据治理为例进行说明。
- [0097] 在风险数据图谱中,涉及到的风险数据包括绩效、人员管理。
- [0098] 在风险信息图谱中,经过风险数据生成的风险信息包括绩效信息、人员管理信息。
- [0099] 在风险知识图谱中,总的风险知识包括人的绩效管理到位。
- [0100] 分项的知识:人员对绩效管理的风险;物对绩效管理的风险;环境对绩效管理的风险;绩效管理本身的风险
- [0101] 人员对绩效管理的风险:不同人员对绩效目标的关心程度;不同人员对绩效达标的自我管理的能力;不同人员对绩效闭环的相关行为
- [0102] 物对绩效管理的风险:对尽责者需要物质激励、失责者需要物质惩罚;某些工种是否可以被自动化程度的机器替代
- [0103] 环境对绩效管理的风险:大的经济环境对客户的影响;
- [0104] 绩效管理本身的风险:绩效管理是否需要及时失责提醒,需要提醒到谁
- [0105] 在本实施例中,总体为如图3所示的高校治理,将共建共治共享(民主化治理、法治化治理、民生化共享)的治理机制(以责联网方式结合)、各业务部门的业务治理和信息部门的IT治理三大块整合到数据治理中,真正实现知识、信息和数据层的风险全方位融合,通过知识层、信息层到数据层的风险图谱式的双向逐级治理,确保知识形成真正有价值的数据资产之后,还形成认知创新和业务创新。
- [0106] 实施例五
- [0107] 本实施例提供一种基于风险图谱的数据治理方法,包括:
- [0108] S1. 构建基于风险知识、风险信息、风险数据的风险图谱;
- [0109] S2. 根据构建的风险图谱对待治理数据进行处理。
- [0110] 进一步的,所述步骤S1具体包括:
- [0111] S11. 获取风险数据相对应的数据因子,并根据获取的数据因子构建风险数据图谱;
- [0112] S12. 根据获取的数据因子生成风险信息,并根据风险数据图谱以及生成的风险信息

息构建风险信息图谱；

[0113] S13. 根据生成的风险信息确定风险知识,并根据风险信息图谱以及确定的风险知识构建风险知识图谱。

[0114] 进一步的,所述步骤S13中确定的风险知识包括人的风险知识集、管理的风险知识集、物的风险知识集、环境的风险知识集。

[0115] 进一步的,所述步骤S12中还包括对风险信息进行扩展处理。

[0116] 进一步的,所述步骤S2中对待治理数据进行处理具体为通过风险知识对风险信息、风险信息对风险数据的图谱进行约束对待治理数据进行处理。

[0117] 需要说明的是,本实施例提供的一种基于风险图谱的数据治理方法与实施例一类似,在此不多做赘述。

[0118] 与现有技术相比,本实施例通过设计数据、信息、知识分层的图谱,通过知识层对信息层、信息层对数据层的图谱约束规则,结合数据间、信息间的图谱组合规则,形成数据治理的管理机制;且通过知识层、信息层到数据层的知识图谱式的双向逐级治理,可以更全面、有效地进行数据治理,确保知识的价值。

[0119] 本文中所描述的具体实施例仅仅是对本发明精神作举例说明。本发明所属技术领域的技术人员可以对所描述的具体实施例做各种各样的修改或补充或采用类似的方式替代,但并不会偏离本发明的精神或者超越所附权利要求书所定义的范围。

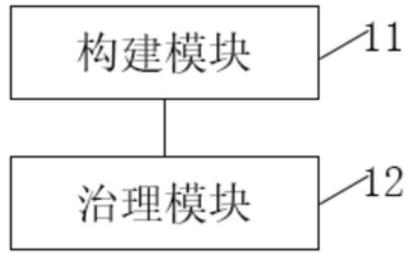


图1

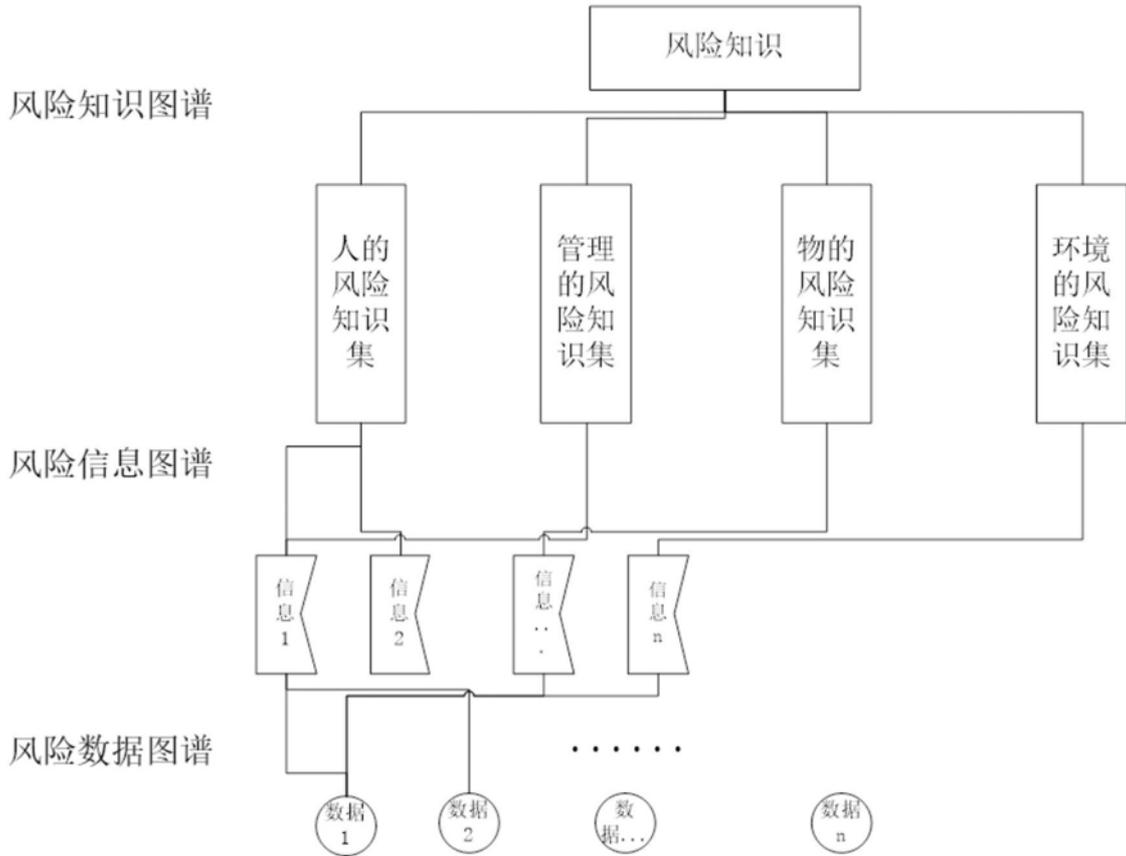


图2

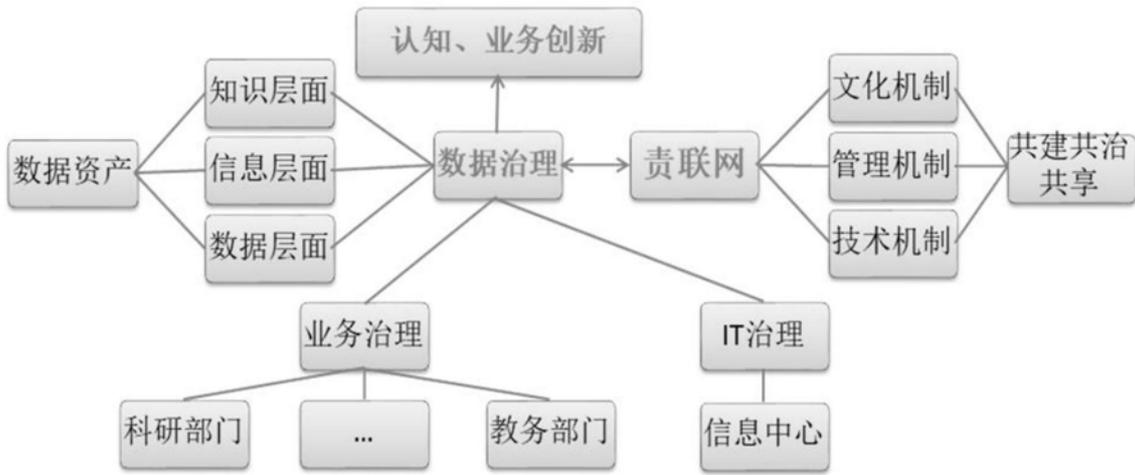


图3