

房地产咨询公司数据标准探索

一 引言

企业信息化建设是提高企业核心竞争力的重要途径。随着云计算、大数据、移动互联网等技术的蓬勃发展，企业信息化进入转型升级的阶段，逐渐由流程驱动转向数字驱动。2020 年初的新冠肺炎疫情对企业的复工造成不小的影响，远程协同办公的需求在客观上倒逼企业加速完善信息化系统。

对房地产咨询公司而言，开展房地产估价、房地产经纪、信息咨询等业务有着广泛的数据需求；同时，各个业务部门在长期发展中积累了大量的数据。如何有效整合和利用公司内外部业务相关数据，挖掘出更多的潜在价值是房地产咨询公司推进信息化的重点。要想实现数据价值与核心业务的深度融合，首要任务是建立一套完善的数据标准体系。

目前，在房地产及相关领域，政府有关部门建立了各种数据标准，如《房地产市场信息系统技术规范》、《房地产市场基础信息数据标准》、《不动产登记数据库标准（试行）》、《土地利用规划数据库标准》

等；各个行业组织也在积极推进行业数据标准建设，如房地产经纪行业自律发展组织“中房经联”发布了《房地产房源数据库标准行业建议版 1.0 版》；部分从业人员就房地产行业整体或细分业务领域的数据标准建设进行了有益的探索。

然而，企业数据标准建设的出发点与政府、行业组织不同，其数据标准的范围应该主要集中在业务的核心数据部分，而不是将房地产领域所有数据悉数纳入考虑范围。本文以房地产咨询公司为例，对构建数据标准时使用的思路和逻辑进行论述，以展示站在企业角度进行数据标准建设的过程。

二、房地产咨询公司数据特征

“标准”的意义是为了在一定范围内获得最佳秩序，想要在企业范围内构建数据秩序，首先需要研究其产生和使用数据的特征。

（一）数据在细分领域上有所侧重

不同于政府和行业组织，公司的数据围绕业务产生。由于房地产咨询公司市场定位各异，如一些偏重商办物业、一些偏重居住物业，因而产生的数据也不尽相同，会集中在个别细分领域。所以，在数据标准构建初期，企业需要聚焦当前或近期开展的业务领域，没有必要将房地产行业内大量的数据对象都纳入研究范围。

（二）数据之间内涵存在差异且多变

公司开展的业务种类繁多，数据需求也不尽相同，这往往导致数据之间的内涵会有较大差异。例如，同样是搜集某写字楼的数据，抵押评估业务更关注产权单元信息，而市场定位咨询业务更关注现状信息。此外，随着政策制度调整、市场环境变化以及公司的不断发展，业务类型也在更新迭代，这使得同步产生的数据的内涵不断变化。因此，企业在构建数据标准时，应考虑标准对多类型业务的适应性，避免因为业务类型调整而导致数据与标准脱节。

（三）数据在执业空间范围内零散分布

单个业务项目的对象多数为单一对象，如评估业务多数为某套房屋，规划业务多数为某个村镇。公司

在承接业务时空间上也存在随机性，这使得业务产生的数据在执业空间范围内零散分布。而公司的执业空间范围通常较大，例如我司深耕上海，但仅上海就有 214 个街道镇、14000 多个小区、6000 余条道路、6430 平方公里。广大的空间范围，零散分布的数据，如何让这些数据之间产生有序、易识别的联系也应是构建数据标准的重点关注内容。

- 第一范式：数据表中的每一列(属性)，必须是不可拆分的最小单元。
- 第二范式：满足第一范式的基础上，要求表中的所有列，都必需依赖于主键。
- 第三范式：满足第一、第二范式的基础上，任何非主属性不依赖于其它非主属性。

（二）数据标准构建实践

接下来以房地产咨询公司实践为例，对该思路进行详细阐述。

1.设定通用的数据基础对象

房地产咨询公司开展的业务中，房地产是核心，它既有具象的实体，也有抽象的产权。考虑到企业虽然无法全面、持续地了解房地产权属和实体的对应关系，而实体对象是长期、客观存在的，企业可以直接进行探知和了解，我们选择从实体开始设定。首先设定两个基础对象，即土地和自然栋。土地指地球表面的特定区域，自然栋指土地上独立的、有固定界限的建筑物和构筑物。

对于建筑物，业务往往关注它的实物特征，而房屋的建筑实物特征在自然栋内并不均匀，这种现象在办公楼、商场等物业类型中尤其普遍。为了更好地描述实物特征差异，我们将自然栋进行逻辑拆分，在只考虑实物特征差异的前提下，将自然栋划分为连续、内部匀质的自然子栋。

随着项目的推进，通常会获得一些房地产的产权数据，对于这些可以明确实体和产权对应关系的自然子栋便将继续进行逻辑拆分。根据最小产权划分在自然子栋内拆分，同时保证实体的连续性，由此形成产权子栋。同样，对于可以明确产权的土地，便可以形成产权宗地。

以上就是我们针对房地产设定的五个基础对象（见图 1）：土地、产权宗地、自然栋、自然子栋、产权子栋。这些对象依次递进，对房地产从实体到产权进行了归纳。



图 1

2.将业务数据与基础对象匹配

完成基础对象设定后，将业务中产生的各项数据与基础对象进行匹配。对于不能直接匹配上的，可以通过基础对象组合成次级对象，再进行匹配。这样保证了所有数据都在一个逻辑体系下运作，为各个数据项形成了有序的逻辑联系。例如通过自然栋可以组合成住宅小区、通过产权子栋可以组合成逻辑栋等（见图 2）。

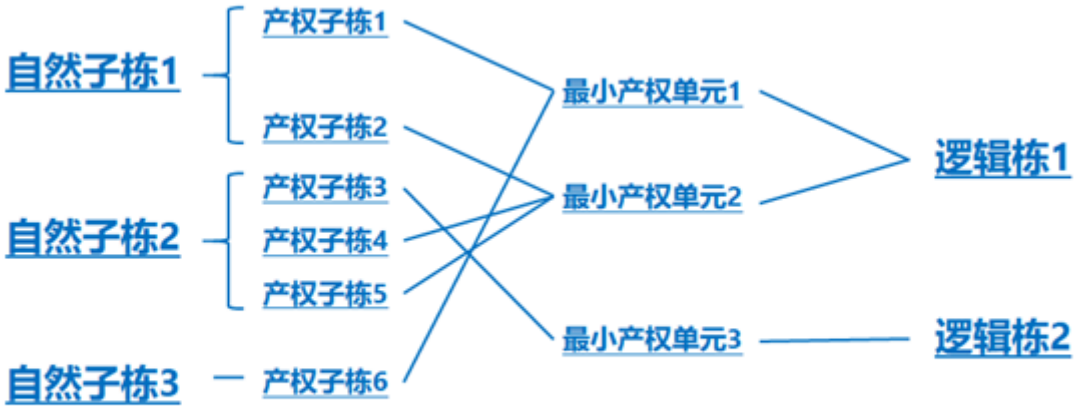

图 2

3.为各对象分级构建属性

接着为各个数据对象分级构建属性，将属性分为三类：界定属性、通用描述属性、业务主题属性。首先构建界定属性，界定属性描述的是数据对象最基础的特征，通过这些属性的组合来区分对象内的每一条记录；其次是通用描述属性，此类属性是对各项业务的数据需求进行归纳，用于描述比较常用的对象特征；最后是业务主题属性，这类属性由各个项目团队根据业务需求补充，自行构建。

以自然栋的属性表为例（见表 1），首先通过底面投影、实际总楼层、建筑物状态来区分出唯一的自然栋，随后通过归纳各类涉及自然栋的数据需求，列有涵盖名称、涵盖地址、主体结构、建筑总面积等通用描述属性，而业务主题属性则在实际业务中自行补充。

表 1

自然栋数据标准			
属性名称	**定义**	**属性类型**	**分类**
ID	自然栋唯一标示。	字符串	界定
底面投影	自然栋 1 层在地表平面的投影面。	GIS 面	界定
地上实际总楼层		整形数值	界定
地下实际总楼层		整形数值	界定
建筑物状态	分为历史、期房、现房。	单选	界定
主体结构	钢结构、钢和钢筋混泥土结构、钢筋混泥土结构、混合结构、砖木结构、其他结构	单选	通用描述
建筑总面积	自然栋的建筑总面积，单位为平方米。	数值	通用描述
竣工时间	自然栋的建筑竣工验收通过时间。	日期（年月日）	通用描述
涵盖地址	自然栋涉及的地址。	字符串	通用描述
涵盖名称	自然栋涉及的名称。	字符串	通用描述
……			通用描述
物业管理费用	物业业管理公司向物业管理使用人或所有人提供物业管理服务收取的费用，单位为元/平方米*月。	整形数值	写字楼估价业务
客梯数量	客用电梯数量。	整形数值	写字楼估价业务
货梯数量	货用电梯数量。	整形数值	写字楼估价业务
……			写字楼估价业务
……			其他业务

4.强调建设地理空间属性

房地产有两个显著的特征，第一它在位置上是固定的，不可移动；第二它是客观排他的，在同样的空间、时间上不可能出现其他的房地产。因此，我们在构建各对象的属性时，特别强调地理空间属性的建设，使其作为各对象的核心差异特征，例如自然栋关注建筑的底面投影、商办综合体关注管理范围边界等。这样借助一些地图引擎可对数据对象中每条记录的空间位置进行直观地显示（见图 3）。



图 3

四、结语

数据标准成功与否取决于能否落地实施，本文提供了一种以业务价值为导向的数据标准构建思路。在该思路中，首要任务是设定基础对象，其他所有对象都在其上构成。在完成数据与对象匹配后，公司内所有数据就形成了逻辑联系（见图 4）。在此基础上，根据业务实际需求可以对次级对象、业务主题属性等进行动态调整和扩展，确保数据标准跟上当前业务推进的步伐，甚至满足未来业务发展的需要。

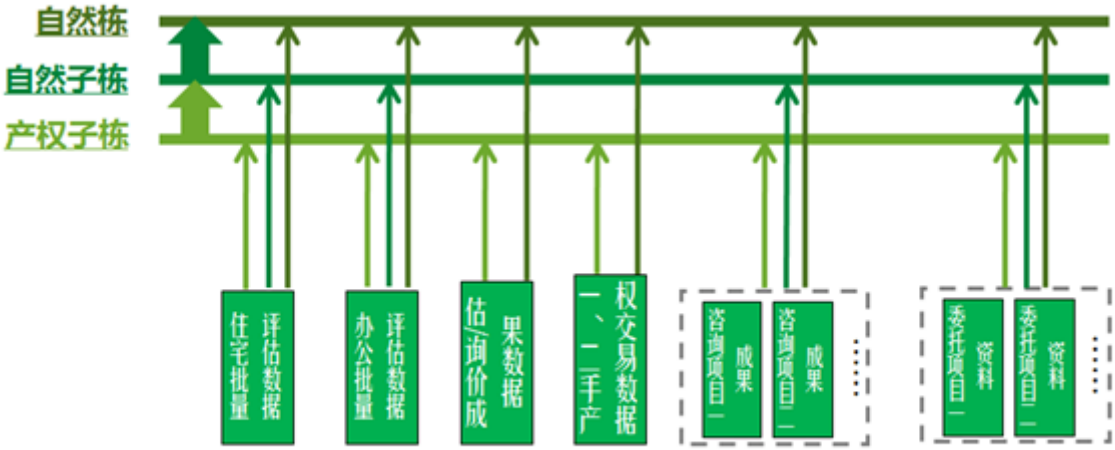

图 4

此外，我们在数据标准中强调使用地理空间属性。地理空间属性可以让数据在地图上具现，通过地图实现直观地查询数据，了解数据和数据之间的空间关系（见图 5）。并且量化的地理空间属性也可以直接用来进行区位条件的运算。



图 5

围绕业务需求进行数据标准构建，让公司有了数据整合、数据共享的基础，进而可以有效减少数据重复建设的工作、提高工作效率，也能为公司业务创新提供有力支撑。

当然，本文只是在数据标准构建思路上进行了探讨，数据标准的具体内容要根据公司的实际业务情况进行调整，并且数据标准也仅仅是信息化的第一步，后续还需要配套的程序系统辅助，才能实现高效的落地。同时，常态化的维护也需要企业建立配套的管理机制。这些相关问题都需要在实践中研究解决。

免责声明
本报告刊载的一切资料及数据，虽力求精确，但仅作参考之用。本刊分析内容并不反映个别情况，对于特定项目读者应自行寻求专业意见。城市测量师行保留一切版权，未经许可，不得转载。

©城市测量师行 2020 年

主办单位：城市测量师行——上海城市房地产估价有限公司
编辑部地址：上海市北京西路 1 号新金桥广场 11、15-18F
邮编：200003
网址：www.surea.com
电子邮箱：surea@surea.com
电话：86 21 63589988

更多楼市资讯、研究报告可以扫描下方二维码阅读



更多价格、更多功能可扫描下方二维码查询、使用

