arXiv:2301.04493v1 [cs.DB] 11 Jan 2023

SeaLiT 本体——对 CIDOC-CRM 的扩展，用于海洋史信息的建模与集成

PAVLOS FAFALIOS、ATHINA KRITSOTAKI 和 MARTIN DOERR，希腊 FORTH-ICS 文化信息学中心与信息系统实验室

本文介绍 SeaLiT 本体的构建与应用。这一本体扩展了国际标准化组织标准 CIDOC-CRM，旨在对海事历史信息进行建模和集成。该本体逐步开发而成，遵循自底向上方法，需要分析大量真实的一手数据（档案材料），并借助领域专家（海事史学家）的知识进行验证。我们介绍该本体的规范、RDFS 和 OWL 实现，以及利用这一数据模型集成信息的知识图谱。相关信息来源于大量且种类繁多的档案文献，例如船员名册、水手登记簿、海军舰船登记册和工资名册。我们还介绍了一款基于这些知识图谱运行的应用，它通过用户友好界面支持历史学家探索并定量分析集成数据。最后，我们讨论与该本体的使用、演进和可持续性相关的问题。

CCS 概念：• **计算方法** → **本体工程**；• **信息系统** → **本体**；**信息集成**；**数字图书馆与档案**；**语义网描述语言**。

其他关键词与短语：本体、海事史、CIDOC-CRM、数据集成、语义互操作性 ty

1 引言

海事史是对海上人类活动的研究。它是历史学中的一个宽泛主题，着重探讨人类与全球海洋、海域及主要水道之间的各种关系 [7]。该领域的一大研究方向需要收集并整合来自多种不同历史来源的数据，以便对实证事实进行定性和定量分析，并得出有关潜在影响因素的结论 [5, 16]。

例如，SeaLiT 项目（海事史领域的欧洲研究委员会启动基金项目）[1], 研究了19世纪50年代至20世纪20年代从帆船航行向蒸汽船航行的转变，以及这一转变对地中海和黑海海上航行人群的影响 [2]。该项目的历史学家收集了大量不同类型、不同语言的档案文献，包括船员名册、工资名册、水手登记册、海军船舶登记名册和雇佣记录，这些文献来自不同国家的多个机构（关于该项目的更多信息见第2.1节）。关于同一关注实体（如船舶、港口或船长）的补充信息，可能存在于不同的档案文献中。例如，对于同一艘船，一个来源可能提供其所有者的信息，另一个来源可能提供船舶的建造详情和特征（长度、宽度、吨位、马力等），而其他来源则可能提供有关该船航次和船员的信息。

在此背景下，信息集成对于开展有效的数据分析并得出可靠结论至关重要，例如回答需要合并和汇总信息的问题，如“统计在特定港口抵达、居住地各不相同，且曾在特定类型船舶（如双桅横帆船）上担任船员的水手人数”。此外，基于通用数据模型进行信息集成，可以生成价值高且长期有效的数据，使其能够在特定研究活动或项目之外重复使用，也能与更广泛的（历史科学）社群的其他数据集集成。

[1] https://sealitproject.eu/

作者地址：Pa vlos Fafalios，fafalios@ics.forth.gr；Athina Kritsotaki，athinak@ics.forth.gr；Martin Doerr，martin@ics.forth.gr，Cent re 文化信息技术与信息系统实验室，FORTH-ICS，N. Plastira 100，希腊伊拉克利翁，GR-70013。

为此，本文介绍*SeaLiT* 本体的构建与应用。该本体旨在通过提供一个通用且可扩展的语义框架，促进对海事史信息的共同理解，并支持信息建模与集成。它采用并扩展了 CIDOC 概念参考模型（CRM）（ISO 21127:2014）[2]，将其作为一种形式本体，用于描述发生在空间和时间中的人类活动、事物与事件 [3]。

该本体在设计时考虑了领域专家（一大群海事史学家）的需求和知识，这些需求和知识体现在他们的研究需求、所遵循的推理过程以及所作的例外处理之中。本体采用自底向上的方式开发，通过分析大量异质的一手数据，尤其是从多个国家的主管机构收集的不同类型、不同语言的档案文献，包括船员名册、工资名册、民事登记册、水手登记册、海军舰船登记册、雇佣关系记录、人口普查资料等。所有建模决策均由领域专家验证；在实践中，我们还将他们的数据（转录文本）转换为基于 SeaLiT 本体的丰富语义网络。借助用户友好界面，该语义网络使他们能够找到需要结合不同来源信息才能回答的问题。

我们介绍了设计本体所采用的方法论和步骤，并提供其规范说明、RDFS 和 OWL 实现，以及使用该本体集成从大量不同档案文献转录的数据所构建的知识图谱。我们还介绍了一款数据探索应用程序，该程序基于这些知识图谱运行，目前可帮助海事史学家探索和分析集成数据。

表 1 列出了 SeaLiT 本体及相关资源和信息的主要访问链接。

表 1. SeaLiT 本体的主要访问链接和信息。

SeaLiT 本体规范	https://zenodo.org/record/6797750
SeaLiT 本体的 DOI	10.5281/zenodo.6797750
SeaLiT 本体的命名空间	http://www.sealitproject.eu/ontology/
SeaLiT 本体 RDFS（Turtle）	https://sealitproject.eu/ontology/SeaLiT_Ontology_v1.1_RDFS.ttl
SeaLiT 本体 RDFS（RDF/XML）	https://sealitproject.eu/ontology/SeaLiT_Ontology_v1.1_RDFS.rdf
SeaLiT 本体 OWL（RDF/XML）	https://sealitproject.eu/ontology/SeaLiT_Ontology_v1.1.owl
SeaLiT 知识图谱（KG）	https://zenodo.org/record/6460841
SeaLiT 知识图谱的 DOI	10.5281/zenodo.6460841
基于知识图谱的 ResearchSpace 应用程序	http://rs.sealitproject.eu/
SeaLiT 本体和知识图谱的许可	知识共享署名 4.0

本文其余部分安排如下：第2节介绍本研究的背景，提供必要的背景知识，并讨论相关工作。第3节详细说明我们构建本体所遵循的方法论和原则。第4节介绍该本体，举例说明模型的某一部分如何经过多次修订，以纳入新的历史知识，并提供其规范说明以及 RDFS 和 OWL 实现。第5节介绍该本体在真实环境中的应用。第6节讨论其用途和可持续性。最后，第7节总结本文并概述未来工作。

[2]https://cidoc-crm.org/

2 背景、背景信息与相关工作

2.1 SeaLiT 项目

该本体是在 SeaLiT 项目[3],的背景下开发的。SeaLiT 是一个海事史领域的欧洲项目（欧洲研究委员会启动基金，编号 714437）。该项目研究了19世纪50年代至20世纪20年代间，从帆船航行向蒸汽船航行转变的过程，以及这一转变对地中海和黑海海上航行人群的影响。SeaLiT 的历史学家研究海事劳动力市场、船东、船长、船员与地方社会之间不断变化的关系，以及从帆船航行向蒸汽船航行的转型时期内新商业策略、贸易路线和航行模式的发展。科学研究聚焦的主要概念包括船舶（涵盖类型、用途、尺寸、技术等各类信息）、与船舶相关的人群（水手、船舶所有者、学生、亲属），以及与之相关的历史事件和活动（如航次、招募、付款）。

SeaLiT 项目所考察和研究的档案来源包括手写的船舶航海日志、船员名册、工资名册和雇佣记录，以及不同类型的登记册，例如民事登记册、水手登记册、学生登记册和海军舰船登记册。这些档案来源来自地中海和黑海沿岸国家的不同机构，使用多种语言书写，包括西班牙语、意大利语、法语、俄语和希腊语。SeaLiT 项目研究的完整档案语料库详见项目网站。[4]

2.2 CIDOC-CRM 国际标准化组织标准

SeaLiT本体使用并扩展了 CIDOC-CRM（概念参考模型）[5]，尤其采用其稳定的 7.1.1版，这意味着 SeaLiT 本体的每个类都是 CIDOC-CRM 类的直接子类或其后代。

CIDOC-CRM 是一个高层次的事件中心本体，涵盖人类活动、事物以及发生在时空中的事件，为描述文化遗产文献中使用的隐含和显式概念及其关系提供定义和形式化结构 [3]。它是受控交换文化遗产信息的国际标准（ISO 21127:2014）[6]，旨在作为领域专家和实施人员共同使用的语言，帮助他们为信息系统制定需求，并提供整合来自不同来源的文化遗产信息的方法。

CIDOC-CRM 当前考虑的稳定版本（7.1.1版）包含 81 个类和 160 个独特属性。CIDOC-CRM 中最高层次的区分由顶层概念表示：E77 持久项（相当于哲学概念中的持存体）、E2 时间实体（相当于哲学概念中的延存体），以及进一步的 E92 时空体积概念。E92 时空体积描述的是这样一些实体：其物质在物质世界中具有或构成可识别且边界明确的几何范围，并且该范围可能随时间变化。图 1 展示了 CIDOC-CRM 的高层类之间的关联。

2.3 相关工作

近年来，语义网的方法与技术开始在历史研究中发挥日益重要的作用。 [13]中的综述回顾了语义技术应用于历史研究的最新进展，尤其涉及以下方面：i）知识建模（本体、数据链接）；ii）文本处理与挖掘；iii）搜索与检索；以及 iv）语义互操作性（数据集成、分类系统）。

[3] https://sealitproject.eu/ [4] https://sealitproject.eu/archival-corpus [5] http://www.cidoc-crm.org/ [6] https://www.iso.org/standard/57832.html

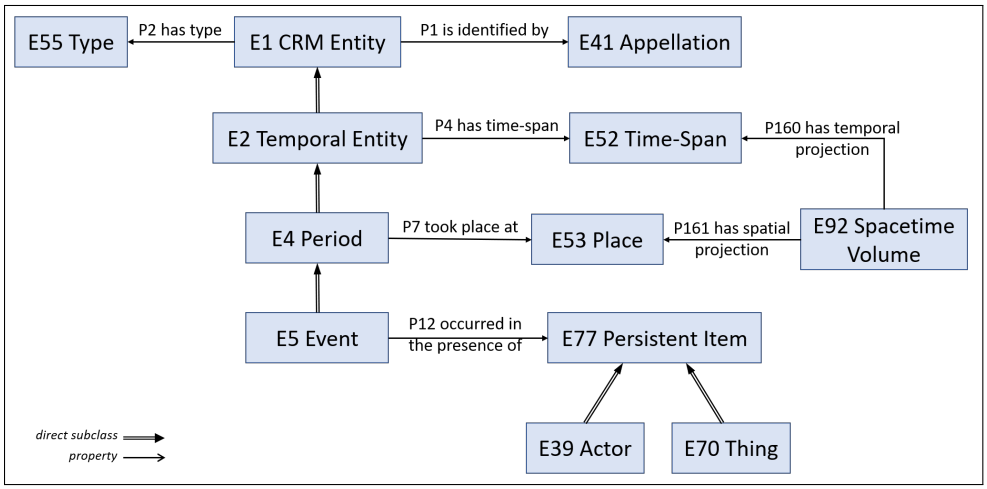

图 1. CIDOC- CRM.

关于用于建模海事史信息的本体，最相关的工作是正在进行的本体管理环境 OntoME 项目 [1]，其旨在为海事史/航海史领域提供数据模型。[7] 该项目由荷兰历史惠更斯研究所、LARHRA 和历史数据联盟合作开展。目前的（草案）模型包含 13 个类和 12 个属性，同时使用 CIDOC-CRM 及其扩展。该本体尚未完成，截至 2022年12月15日仍不可使用。

Conflict[8] 是一个在 SAILS 项目（2010—2013 年）[9] 背景下开发的本体，用于建模描述第一次世界大战的相关概念。所提供的本体版本（0.1）实际上是一个分类法，由 175 个类组成，其中一些类可用于建模海事史相关信息，例如 Ship、Ship_journey、Ship_type 和所有权。同样，也有一些本体可用于建模该模型的其他部分，例如 *GoodRelations* [8],，这是一个用于交换电子商务信息的轻量级本体，可用于涉及产品付款的部分。

我们选择使用 CIDOC-CRM，因为它是文化遗产文献领域的标准本体，在文化遗产、历史和考古学领域得到广泛应用。它与历史这一学科的论述定义域直接相关；历史研究人类及社会在过去的生活。从海事历史研究的角度考察这一范围，可以通过 CIDOC-CRM 提供的现实抽象来表示。例如，我们可以直接利用 CIDOC-CRM 类 E7 活动的（直接或继承）属性，如*'P14* 由.....执行'、*'P4* 具有时间跨度'、*'P7* 发生于'等，并将它们用于描述 SeaLiT 本体中作为 E7 活动子类的类的实例（如航行、抵达、招聘等）。因此，使用 CIDOC-CRM 有助于与相关（现有或未来）且同样采用 CIDOC-CRM 的数据集进行数据集成，同时也能确保数据可持续性，因为 CIDOC-CRM 是一项动态标准，拥有一个非常活跃的社群，持续对其进行维护和改进。

最后，已有大量本体作为 CIDOC-CRM 的扩展而开发，例如用于记录考古科学的 CRMas [14]、用于地理空间信息的 CRMgeo [9]、用于数字对象来源的 CRMdig [18]、用于事实论证的 IAM [4]，以及其他本体。

66

[7]https://ontome.net/namespace/[8]
http://ontologies.michelepasin.org/docs/conflict/index.html [9]
http://sailsproject.cerch.kcl.ac.uk/

3 设计方法与原则

3.1 总体方法

该本体遵循自底向上策略 [6], 逐步构建，以真实的实证数据和信息需求为基础，尤其是数字化的历史记录（转录文本）及其各种形式的相应数据结构，以及由庞大的历史学家群体提出的研究问题。档案材料与研究问题共同定义了建模需求。

我们策略的主要特点概述如下：

- 研究并分析与海事史相关的大量、多样的档案来源。这些资料提供了有关船舶、人员（如水手、船长、船舶所有者、学生）以及相关活动和事件（如航次、招募、付款、教学活动）的历史信息。
- 汇集研究问题及相应的信息需求（能力问题），这些问题和需求可由所考察的档案来源提供答案或重要相关信息。
- 我们与来自不同机构和国家（西班牙、意大利、法国、克罗地亚、希腊）的一大群海事史学家进行了长时间讨论，以征求意见，并了解他们采用的推理过程和例外处理方式。

具体而言，我们的研究方法侧重于从历史学家的视角研究和分析历史来源，并遵循他们各自的研究问题和文献记录实践。为此，我们必须长期咨询所有数据提供者（他们来自不同的研究团队和国家），并深入研究其研究实践和历史数据，以开发和验证模型。因此，该模型根据实际数据值、现有且实际使用的结构化信息来源（如电子表格）以及包含原始信息的历史记录（转录文本）设计而成。在项目期间，为纳入来自新类型来源的新信息，我们多次完善模型中的概念。表2列出了所考察的档案来源，并概述了记录的信息，同时为每种来源提供了一条记录（转录文本）示例。[10]

至于历史学家提出的研究问题和信息需求，大多数涉及汇总信息，例如 特定港口到达的水手按出发地划分的数量、船舶的平均吨位、各国的工资水平、按水手职业划分的移民比例等。其他信息需求涉及检索特定的实体列表（例如 特定时间段内的船舶建造地点）、比较信息（例如 水手在船上服务的时间与在陆地上的时间、船舶上的女性／男性人数等），或检索特定数值（例如 公司在特定年份或若干年间雇用的船员总数）。[11]

在创建本体时，我们采用了定制工程方法 [11]，不过该方法保留了现有方法的大部分特性，例如 HCOME [10] 和 DILIGENT [17]。具体而言：

- 数据驱动／自底向上处理（我们的本体开发策略）
- 领域专家的参与（在我们的案例中为海事史学家）
- 迭代处理（循序渐进、高度迭代的本体开发）
- 协作工程处理（由一小组概念建模专家共同完成）
- 验证与应用（由领域专家进行验证，并在真实环境中应用）
- 详细的版本管理（包含多个中间版本，目前为稳定的 1.1 版）

[10]一个 Web 应用程序 t 用于探索这些档案来源转录文本中数据的工具可在以下网址获取：https://catalogues.sealitproject.e u/

[11]完整的信息 需求列表可在以下网址获取：https://users.ics.forth.gr/~fafalios/SeaLiT_Competency_Questions_InfoNeeds.pdf

表2. 纳入考量的档案来源及记录信息的类型。

档案来源	记录信息概览及转录示例
船员与载重清单（Roll）	船舶（姓名、类型、建造地点、建造年份、登记地点、所有者）、港口的来源地、到达港、目的港、船员（姓名、父亲姓名、出生地、居住地、职业、年龄）、登船港、下船港。 **[示例文本：https://tinyurl.com/4ukzezfe]**
船员名册（Ruoli di Equipaggio）	船舶（姓名、类型、建造地点、建造年份、登记编号、登记港、所有者）、航次（起止日期、持续时间、船员总数）、目的地、出发港、到达港、船员船员（姓名、居住地、出生年份、序列号、职业）、登船港、卸装货港。 **[示例抄本：https://tinyurl.com/2u35frya]**
西班牙船员总名册	船舶（名称、类型、吨位、登记港）、船舶所有者、船员（姓名、年龄、居住地）、航次（起止日期、船员总数）、登船港、目的地。 **[示例抄本：https://tinyurl.com/3axs6ret]**
水手登记簿（Libro de registro de marineros）	海员（姓名、父亲姓名、母亲姓名、出生日期、出生地、职业、兵役组织 地点） **[示例抄本：https://tinyurl.com/2p8kzm6n]**
海事人员登记簿	人员（姓名、父亲姓名、母亲姓名、出生地、出生日期、居住地、婚姻婚姻状况、此前职业、服役机构所在地）。 **[示例转录：https://tinyurl.com/4v6hnwjj]**
海上船员	人员（姓名、父亲姓名、婚姻状况、出生日期、职业、退役原因、工作状态类型）、船舶（姓名）、目的地。 **[示例转录：https://tinyurl.com/2x5cu37n]**
海军船舶登记名录	船舶（姓名、类型、吨位、长度、建造地点、登记地点、所有者）。 **[示例文字稿：https://tinyurl.com/bdhx87tr]**
船舶名录	船舶（姓名、曾用名、类型、登记港、登记年份、建造地点、建造年份、吨位、发动机制造地点、发动机制造商、标称功率、指示功率、所有者）。 **[示例文字稿：https://tinyurl.com/2cphfpef]**
民事登记簿	人员（姓名、职业、出发地、年龄、性别、婚姻状况、死亡地点、死因、相关人员）。关人员）。 **[示例转录：https://tinyurl.com/bdzeja8n]**
海事登记册，拉西奥塔	人员（姓名、出生日期、出生地、居住地、职业、服务行业）、登船地点、下船地点、船舶（名称、类型、导航类型）、船长、资助人。 **[示例文字记录：https://tinyurl.com/fkhyyp4a]**
学生登记册	学生（出发地、职业、雇主公司、宗教、相关人员）、课程（标题、主题、起止日期、学期、学生总人数）。 **[示例文字记录：https://tinyurl.com/mryp6cbb]**
拉西约塔人口普查	居住者（姓名、年龄、出生年份、出生地、国籍、婚姻状况、宗教、职业、工作组织、家庭角色、地址）。 **[示例抄录：https://tinyurl.com/4dzfcbtt]**
俄罗斯帝国人口普查	居住者（姓名、父称、性别、年龄、婚姻状况、等级、宗教、母语、家庭角色、职业、地址）。 **[示例抄录：https://tinyurl.com/43xczvux]**
工资名册（希腊船舶）	船舶（姓名、类型、所有者）、船长、航次（起止日期、总天数、出海天数、在港天数，工资总额、养老金基金总额、净工资总额）、人员（姓名、成人/儿童、识字能力、出生地、职业/职级）、雇佣关系（招聘日期、解雇日期、招聘地点，月工资、工资总额、养老金基金、净工资）。 **[示例记录：https://tinyurl.com/ztjk4jw7]**
俄国蒸汽航运工资名册和贸易公司）	船舶（姓名、所有者）、人员（姓名、父称、成人/儿童、性别、出生日期、等级、登记地点）、招募（港口、文件类型、职级/专业、月薪）。 **[示例 transcription脚本：https://tinyurl.com/y5urjhc9]**
雇佣记录（造船厂：Messageries Maritimes、拉西奥塔）	工人（姓名、性别、出生年份、出生地、居住地、婚姻状况、职业、公司任职情况、车间经理）。 **[示例文字记录：https://tinyurl.com/yc3havkc]**
航海日志	船舶（姓名、类型、电报码、吨位、登记港、所有者）、船长、出发港、目的-港口、航线行程、日历事件类型。 **[示例文字记录：https://tinyurl.com/mrx2re9k]**
账簿	船舶（姓名、类型、所有者）、航次、船长、出发港、目的港、停靠港、交易——（类型、记录地点、供应商、中介、接收方）。 **[示例文字稿：https://tinyurl.com/4uf3bye8]**

3.2 设计步骤与原则

该模型以 CIDOC-CRM 为基础，因为它是一种适用于记录历史信息的标准，能够说明相关的人物、时间、地点和事物。从本体角度出发，我们遵循了以下步骤：

(1) 我们通过创建 CIDOC-CRM 类的子类，并相应地定义属性（其中一些属性是 CIDOC-CRM 属性的子属性），对 CIDOC-CRM 进行了扩展。针对特定类型的档案来源及其对应的信息需求扩展或修订模型后，我们创建了映射，将数据从源模式转换为基于所设计（目标）模型的语义网络（资源框架描述三元组）。这种概念对齐是本体开发过程中的重要步骤，有助于重新设计概念并最终确定模型。

(2) 我们将现有模式中包含的实体分为两类：直接或间接表示事件 的实体，以及表示物体（可移动或不可移动）的实体；再根据它们表示个体还是个体集合，将其归入不同的抽象层级。我们发现，大多数二元关系都可体现为时间实体（例如曾相遇、曾创建等）。这一原则帮助我们从数据结构中识别出隐藏的事件。

(3) 我们根据实体的定义域和值域实体所属的抽象层级，对实体之间现有的关系进行了分类，并据此建立了类层级和属性层级。我们没有针对不同类重复定义同一属性，而是找出该属性适用的最一般（超）类。发现不同类存在重复属性，表明这些属性依赖于某个共同且更一般的概念，而这一概念使这种关系得以成立。找到能够描述这一共同概括的唯一最一般概念后，我们便可以创建一个一般类，将这些属性应用于该类，并通过将原先建模的类指定为新建概括类的子类，使这些关系得以继承（例如 服务报酬 和法律对象关系这两个类）。

(4) 我们针对相关属性找到了相应的类，而非针对相关类寻找属性（例如，属性“航次”对应航行 ，“建造”对应船舶建造 等）。我们确定了每项属性所特有的一般类。换言之，我们找到了一个最具体的类，它概括了所有将该属性作为定义域或值域的类。

(5) 我们通过确定这些概念的身份判据来定义概念，即区分什么属于这些概念的实例，什么不属于。我们识别出独立于属性而存在的类，而不是“任何具有该属性的事物”（例如 服务 概念的情况）。

(6) 所开发的类和关系数量足以回答全局性质的查询。所谓全局查询，是指用户同时向多个数据库（来源）提出的查询，以获得全面的答案，尤其是包含跨数据库连接的查询。还应强调，目标并非对“一切”进行建模，而是为这一特定定义域中的必要且已充分理解的概念建立模型。

本体 的构建遵循了这些原则。其设计与开发是一个迭代过程，历经 th
多次重复 上述步骤。

4 SEALITONTOLOGY

我们首先概述 本体（第 4.1 节），然后介绍一个本体演化示例（第 4.
2 节），最后我们).
介绍本体的规范，以及 RDFS 和 OWL 实现（第 4.3

4.1 本体概述

该本体目前 （1.1 版）包含 46 个类、79 个属性和 4 个属性的属性，可用于描述
ng the descriptions of 关于船舶的 船舶航行、航海人员、雇佣关系和付款、教学 *ng*

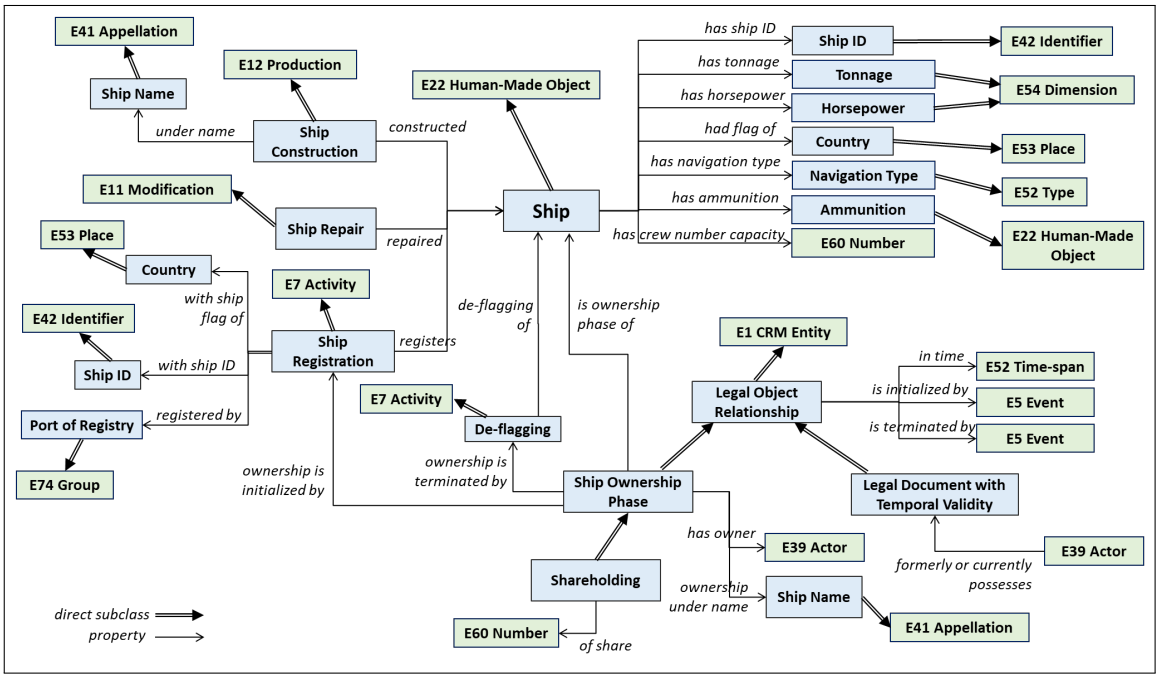

图 2. 船舶信息的建模。

活动，以及大量其他相关活动和特征。附录 A 和 B 分别提供了完整的类层级和属性层级。

图 2 展示了如何对船舶信息进行建模。[12] 船舶（E22 人造物品的子类）是船舶建造活动（E12 生产的子类）的产物；该活动为船舶赋予了船名（E41 称谓的子类）。船舶还具有一些特征，例如马力和吨位（E54 尺寸的子类；除了数值之外，还可以提供相应的计量单位、注释等），并通过船籍港（E74 群体的子类）进行船舶登记（E7 活动的子类），登记时还会注明特定国家（E53 地点的子类）的船旗以及特定的船舶标识（E42 标识符的子类）。将船舶标识建模为一个类，便于纳入有关该标识符的其他信息，例如由哪个机构、何时提供该标识符等（通过将其与 CIDOC-CRM 类 E15 标识符指派关联）。最后，一艘船舶具有一个或多个船舶所有权阶段（法律对象关系的子类），每个阶段均由船舶登记启动，并由撤销船旗活动终止。请注意，所有与活动相关的类（如船舶建造、船舶维修、撤销船旗等）均可使用 CIDOC-CRM 属性*'P4* 具有时间跨度'来提供时间信息。

图 3 展示了本体如何对船舶航行 相关信息进行建模。首先，航行（E7 活动的子类）涉及一艘特定船舶，由一名或多名船长（E39 行动者）驾驶，并具有出发地点 、目的地 以及最终抵达地点 （E53 地点）。然后，船舶航行期间的主要活动包括装载货物、离开地点、经过或穿越地点、抵达地点，以及

[12]名称以字母 'E' 后跟数字开头的类是 CIDOC-CRM 类（图中以绿色框表示）。其他所有类均属于 SeaLiT 本体（以蓝色框表示）。相应地，名称以字母 'P' 后跟数字开头的所有属性均为 CIDOC-CRM 的属性，而其他所有属性均属于 SeaLiT 本体。

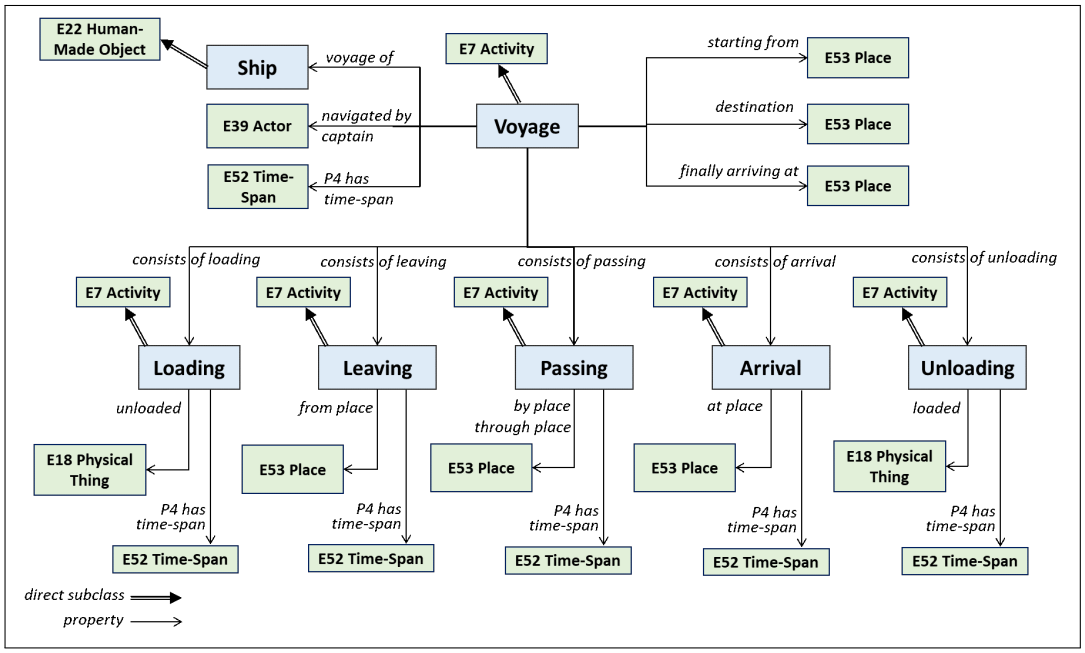

图 3. 船舶航行信息建模。

卸货。这些活动均通过 CIDOC-CRM 属性‘*P4* 时间跨度’关联至 E52 时间跨度。

图 4 展示了该本体如何描述雇佣关系和付款信息。服务报酬（E7 活动的子类）支付给 E39 行动者，以换取特定服务（E7 活动的子类）。[13]服务报酬这一类有两个专业化子类：物品报酬和劳动报酬，而就业这一类是服务这一类的专业化。船员薪酬涉及特定航行，是劳动报酬的专业化。此处，劳动合同（E29 设计或程序的子类）规定了劳动报酬的条件。雇佣关系始于招聘（E7 活动的子类），终于卸载（E7 活动的子类）。

图 5 展示了本体中如何建模有关人员 （航海人员，如船长、船员、学生等）的信息。人员（E21 人）通过民事登记活动进行登记，并获得一个标识符（E42 标识符）。人员拥有名字和姓氏（E62 字符串），工作于某个组织或公司（E74 群体），在特定时间（记录信息时）具有一定年龄（E60 数值），并拥有一组其他属性，尤其是宗教状况、识字状况、性别身份、语言能力、社会地位和职业（均为 E55 类型的子类）。当来源（如民事登记簿或人口普查文件）未提供足够的时间信息，无法推断或观察相应事件时，将 E55 类型用作这些属性／特性的超类（而不是将它们建模为时间实体）是很好的解决方案（SeaLiT 项目的档案来源正是如此）。此外，人员还可以受到惩罚（E7 活动的子类）或获得晋升（E13 属性指派的子类）

[13]我们使用‘ 货币’而非‘付款’一词，因为我们想表明发生了货币交易，例如使用里拉、法郎、 tc.（在古代，支付 可以不使用货币进行，例如使用物品）。

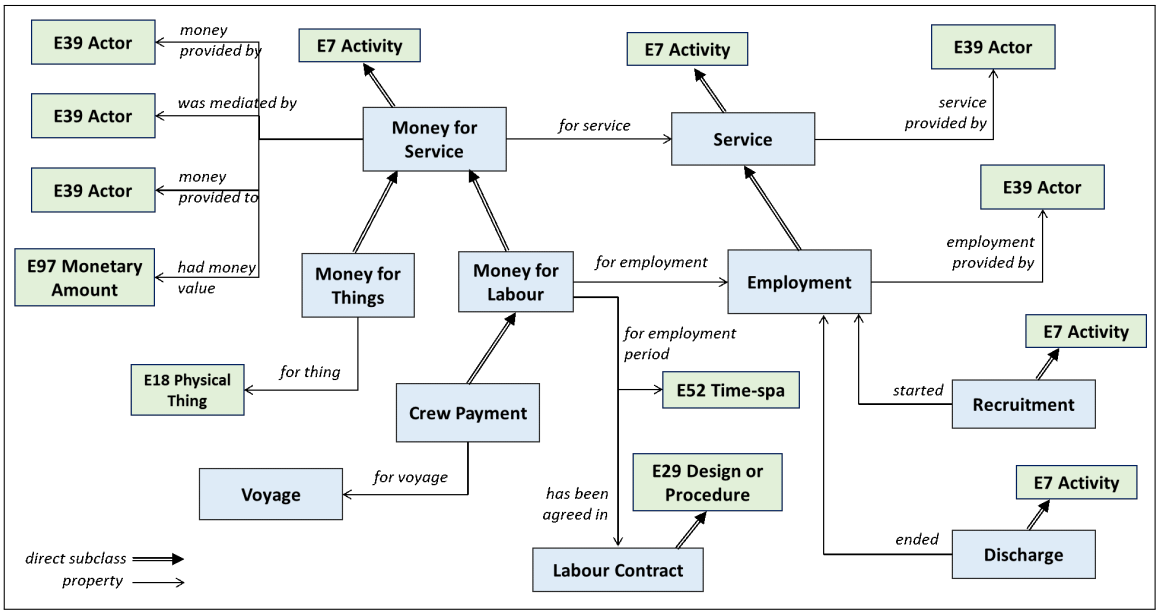

图 4. 雇佣关系和付款信息的建模。

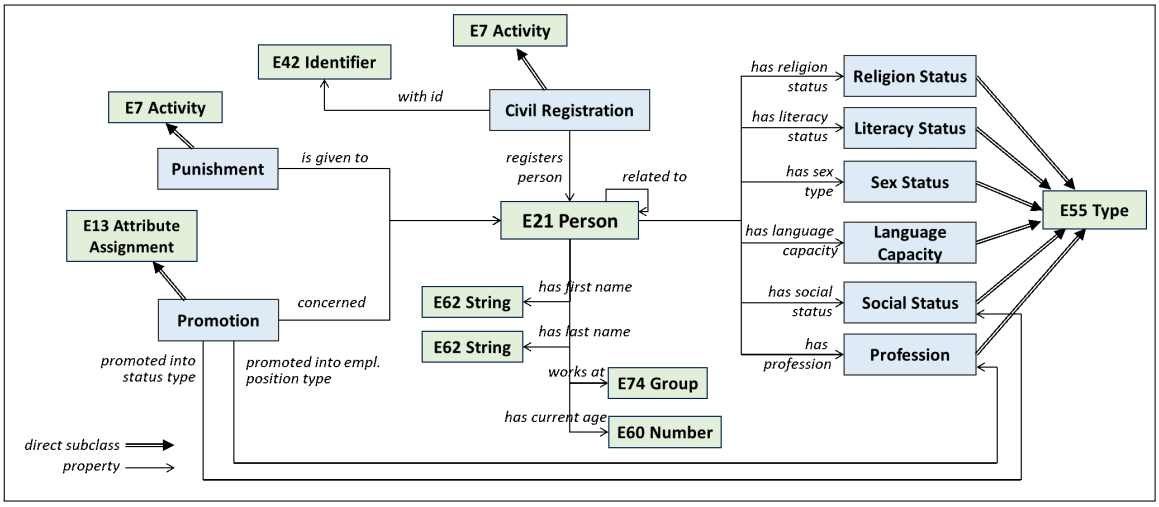

图 5. 人员信息的建模。

。晋升可以与社会地位晋升或工作／职业（职业）晋升相关。

最后，图 6 展示了该本体如何描述与航海相关的教学活动信息。教学单位是一种活动，可细分为课程或章节。它与一个主题（E55 类型的子类）、参与该教学单位的学生（E39 行动者）、参与学生人数（E60 数值）以及一个或多个其他教学单位相连，这通过 CIDOC-CRM

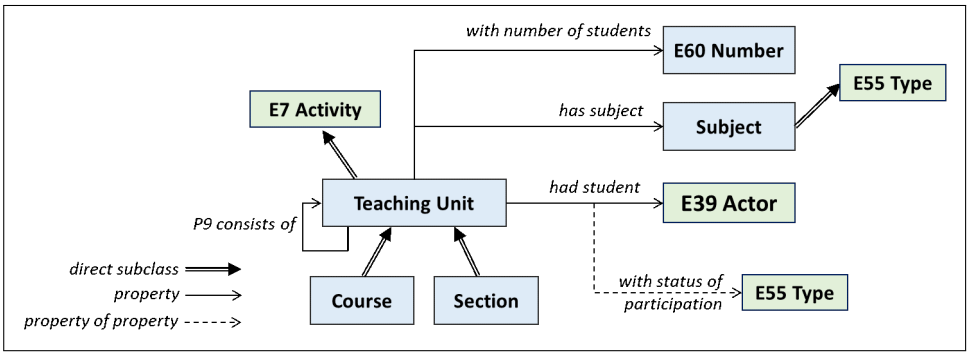

图 6. 教学活动信息建模。

属性 *'P9 consists of'* 实现。特别是，后者可以描述课程由若干章节组成这一信息。

4.2 本体演化示例

本体开发过程历时两年多，其间经历了大量中间版本，最终发布了第一个稳定版本（1.0）。具体而言，本体元素（类和属性）根据以下因素进行了数次修订：(a) 新纳入考量的档案来源带来的新证据；(b) 领域专家（海事史学家）提出的新要求（信息需求）。这些新证据和要求需要定义新元素（例如创建新类或属性），或修订涉及模型某一部分的一组现有元素。

图 7 展示了涉及船舶所有权的本体部分在本体开发过程中如何经过数次修订。

历史学家提出的第一个需求是能够查找每位所有者名下的所有船舶。经过分析的档案材料（船员名册）仅提供了所有者的姓名，其值可能是人员姓名，也可能是公司名称。根据这一证据，创建了属性“有所有者”，用于连接船舶的一个实例与 CIDOC-CRM 类 E39 行动者的一个实例（图 7 中的 v1）。

另一来源（海军船舶登记名册）提供了船舶历任所有者的信息，而新增的需求是能够查找某一时期内每艘船的首任所有者人数。基于这一点，以及二元关系 *has owner* 蕴含／隐藏了一个时间实体这一事实，我们定义了 Ship Ownership Phase 类、用于将船舶与船舶所有权阶段相连的属性 *'has phase'*、用于将所有权阶段与 E52 时间跨度相连的属性 *'in time'*，并修改了属性 *'has owner'*，使其用于将船舶所有权阶段与 E39 行动者相连（图 7 中的 v2）。

一艘船在其生命周期内可能有多个船名，而一位所有者可能拥有多艘同名船（如航海日志和船员与排水量名册所示）。据历史学家称，所有权通常会为船舶指定一个名称；船舶则会在特定时间因所有权状态变更而更名。基于这一历史知识，我们创建了属性 *'ownership under name'*，以便将船舶所有权阶段关联到一个船名（图 7 中的 v3）。

证据表明，船舶所有权属于一种可以推断而无法直接观察的信息。所有权阶段可以通过启动该阶段的船舶登记 活动以及终止该阶段的撤旗活动来追溯。对奥地利劳埃德船队名册中船舶登记的记录，尤其包含船舶的建造地点和日期信息；这些信息与船舶建造后获得的船名共同构成识别船舶的可靠标准。基于此，定义了船舶登记类

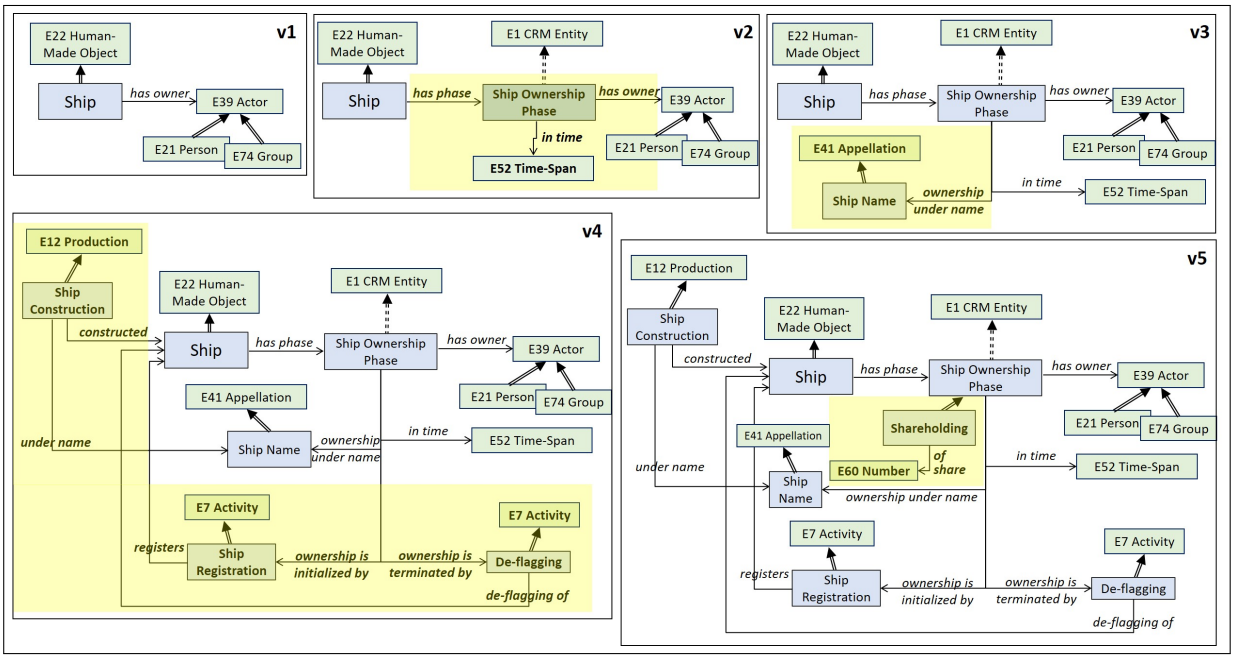

图 7. 用于建模船舶所有权信息的本体演化示例。

（E72 活动的子类）、撤旗类（E72 活动的子类）和船舶建造类（E12 生产的子类），并定义了以下属性：‘登记’（用于将登记活动关联到船舶）、‘所有权由……开始’（用于将所有权阶段关联到登记活动）、‘撤旗对象’（用于将撤旗活动关联到船舶）、‘所有权由……终止’（用于将所有权阶段关联到撤旗活动）、‘建造’（用于将建造活动关联到船舶），以及‘船名为’（用于将建造活动关联到船名（图 7 中的 v4）。

船舶的所有权实际上是一种法律协议，其中所有者持有股份。例如，根据意大利资料来源（海事登记册），船舶所有权被划分为 24 份（“卡拉蒂”）。有时，一位船舶所有者拥有全部 24 份。然而，更常见的情况是，这 24 份由多位船东共同持有。基于这一证据，新建了一个类“持股”，作为“船舶所有权阶段”的专业（子类），并添加属性 ‘股份数’，用于为股份持有阶段指定股份数量（图 7 中的 v5）。

在上一个本体版本中（见图 2），船舶所有权阶段被定义为法律对象关系类的专业化（子类）；同一版本还定义了具有时间效力的法律文件类，其中包括在特定时间跨度内有效的官方文件或法律协议。更一般的法律对象关系类表示其状态和时间跨度没有记录、因而无法直接观察的关系类型。我们只能通过初始化或终止该状态的事件（开始事件和终止事件）来观察这种关系。

4.3 规范说明、RDFS 与 OWL 实现

本体规范及其 RDFS 实现可通过 Zenodo 存储库获取（DOI：10.5281/zenodo.6797750）[14], , 并采用知识共享署名 4.0 许可。本体的命名空间（可解析），指向其 RDFS 实现，网址为：http://www.sealitproject.eu/ontology/。

规范说明文档定义了本体类和属性。对于每个类，文档会提供：i）其超类；ii）其子类（如有）；iii）范围注释（对该类内涵的文字描述）；iv）该类实例的一个或多个示例；v）其属性（如有），并以属性名称及其关联的值域类表示。对于每个属性，规范说明会提供：i）其定义域；ii）其值域；iii）其超属性（如有）；iv）其子属性（如有）；v）范围注释；vi）该属性实例的一个或多个示例；vii）其属性（如有）。如果某个属性有逆属性，则会将逆属性置于括号内，紧跟在属性名称之后。范围注释并非正式的建模构造，而是用于帮助解释某个类或属性预期的含义和应用。它们指涉领域专家（海事史学家）共同认可的概念化，并对不同的可能解释加以区分。

RDFS 实现通过 *'rdfs:comment'* 提供每个类或属性的范围注释。为生成类和属性的 URI，类或属性名称中的空格会替换为下划线字符。逆属性通过 *'owl:inverseOf'* 提供。本体的版本通过属性 *'owl:versionInfo'* 提供，许可证则通过都柏林核心术语*'dc:license'*提供。对于指向在 RDF 中表示为字面量的类的属性（共七个，指向 CIDOC-CRM 类 E60 数值或 E62 字符串），我们将其值域定义为 rdfs:Literal。

我们还提供了该本体的 OWL 实现，其中包含 71 个对象属性、7 个数据类型属性和 1 个对称属性（即属性 *'related to'*）。[15]

由于 RDF 未提供直接表达属性之属性的方法，我们采用属性类 （CIDOC-CRM 提出并实现的方案）作为再化方法，用于编码 SeaLiT 本体中定义的四个属性之属性。使用此方法，每个具有属性的属性都会对应创建一个类。随后，可将该属性类实例化，并与 CIDOC-CRM 的 RDFS 实现所提供的属性 *'P01 具有定义域'* 和 *'P02 具有值域'*配合使用。[16] 例如，图 8 展示了如何利用属性类的概念实现属性 '工作于' 的属性'角色'。首先，提供属性类 PC works at，用于表示属性 '工作于'。在数据生成/实例化期间，会创建该属性类的一个实例，并使用属性 *'P01 具有定义域'* 和 *'P01 具有值域'*分别指向原始属性 '工作于' 的定义域（E21 人的一个实例）和值域（E74 群体的一个实例）。然后，我们可以通过直接将属性之属性'角色' 链接到该属性类实例来表示它。

5 应用

5.1 SeaLiT 知识图谱

SeaLiT 本体已应用于 SeaLiT 项目（参见第 2.1 节），用于将从多个分散的本地海事史信息来源中转录的数据，转换为丰富且连贯的集成数据语义网络（即知识图谱）。此项转换旨在使人们能够针对集成数据提出复杂问题，例如历史学家提出的、需要整合多个来源信息的问题。

具体而言，历史学家使用 FASTCAT 系统 [5]，以表格形式（类似于电子表格）协作转录和记录原始档案文献。在 FAST CAT 中，来自不同来源的数据

[14]https://zenodo.org/record/6797750

[15]https://sealitproject.eu/ontology/SeaLiT_本体_v1.1.owl

[16]https://cidoc-crm.org/rdfs/7.1.1/CIDOC_CRM_v7.1.1_PC.rdf

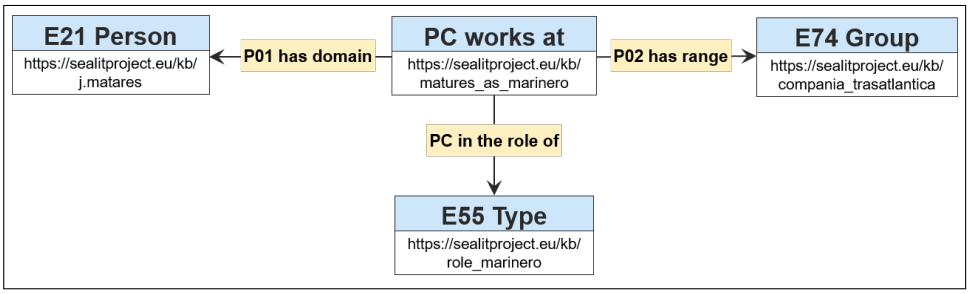

图 8。 Repr使用属性类表示 RDF 中的属性的示例　　属性类。

会被转录为属于特定模板的记录。一条记录通过一组表格整理档案文献的数据和元数据，而模板则表示单个数据来源的结构，即定义数据录入表格。目前，SeaLiT 的历史学家已在 FAST CAT 中创建并填写了 600 多条记录。表2列出了每种不同来源类型（模板）的一条记录示例。

为了基于 SeaLiT 本体将转录数据转换为 RDF，我们为每个不同的 FAST CAT 模板创建模式映射。这些映射定义了如何将 FAST CAT 记录中的数据元素（例如表格中的列）映射到本体类和属性。为创建模式映射并运行转换，我们采用 X3ML 映射定义语言和框架 [12]。转换后的数据（RDF三元组）随后被提取到语义存储库（RDF三元组存储库）中，外部应用程序和服务可以使用 SPARQL 语言和协议访问该存储库。ResearchSpace应用程序（下文将介绍）运行于此类存储库之上，用于支持历史学家对集成数据进行搜索和定量分析。有关 FAST CAT 系统以及数据转录、策展和转换过程的更多信息，读者可参阅 [5] 。

生成的知识图谱可通过 Zenodo 存储库（DOI：10.5281/zenodo.6460841）[17],获取，并采用知识共享署名 4.0许可。该数据集目前包含超过 1,850 万个三元组，提供了约 3,170 艘船舶、92,240 名人 员、935 个法人组织和 5,530 个地点的集成信息。由于包括实例匹配在内的数据整理工作仍在进行中，且新的档案文献正在 FAST CAT 中转录，这些数字可能会在未来版本中发生变化。

5.2 ResearchSpace应用程序

为了帮助历史学家探索 SeaLiT 知识图谱（进而探索集成数据），我们使用了 ResearchSpace [15], , 这是一个开源平台，提供多种功能，包括一个查询构建界面，支持用户通过直观（用户友好）的界面逐步构建复杂查询。随后，用户可以通过不同的可视化方式（例如条形图）浏览、筛选或定量分析结果。该应用程序可通过以下地址访问：http://rs.sealitproject.eu/。

ResearchSpace 的查询构建界面已针对 SeaLiT 知识图谱进行了配置。具体而言，系统定义了以下搜索类别：船舶、人员、法人组织、船员薪酬、地点、航行、课程、记录、来源。用户选择某个类别（例如 船舶）后，系统会显示与之关联的类别列表。选择一个关联类别（例如 地点）后，用户便可选择连接这两个类别的属性（例如 抵达地点）以及一个实例/值（例如 马赛；也就是说，用户正在查询抵达马赛的船舶）。此类属性实际上对应知识图谱中的一条路径，用于连接所选类别的实例。

[17]https://zenodo.org/record/6460841

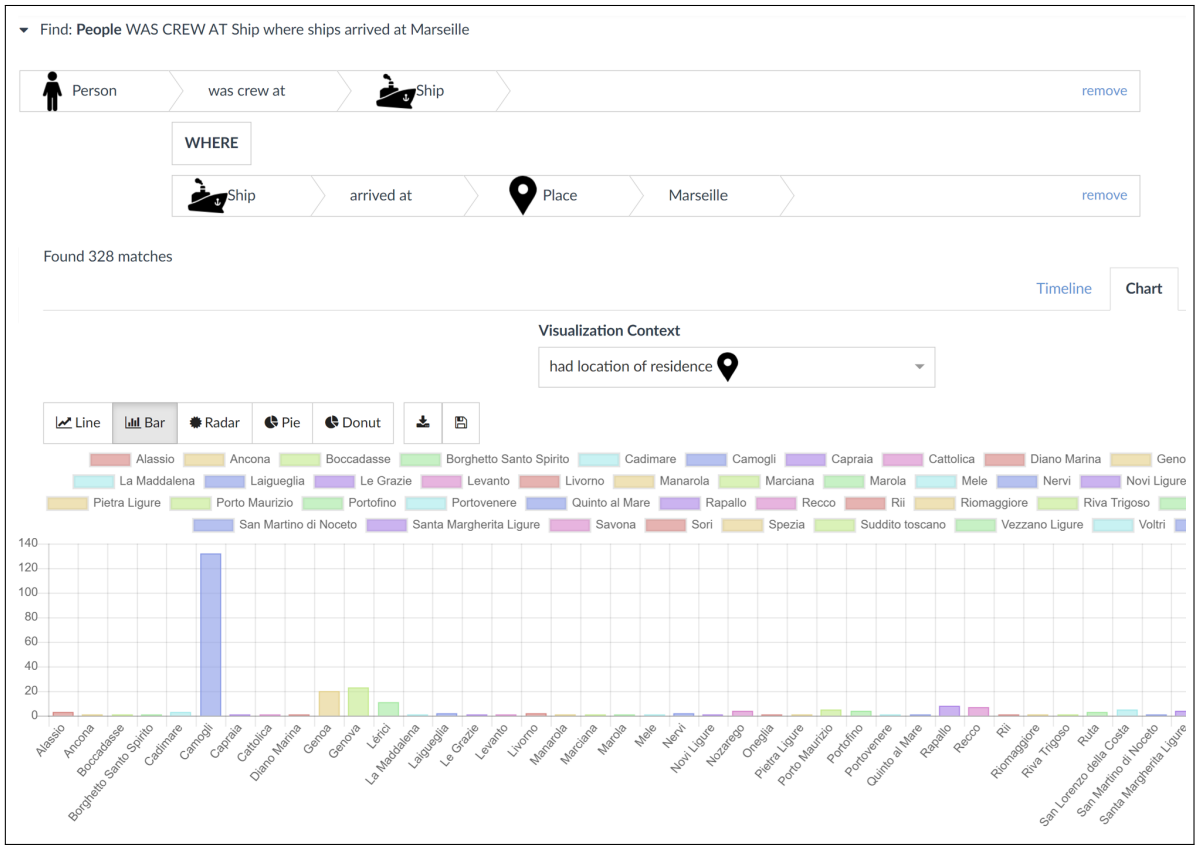

图 9. 在 ResearchSpace 应用程序中构建查询并可视化结果。

图 9 展示了系统的屏幕截图。在此示例中，用户查询了曾在抵达马赛的船舶上担任船员的人员，[18] 并选择按这些人员的居住地 分组，并以条形图呈现结果。从条形图中可以看出，大多数人员的居住地是卡莫利 。这一查询对应于 SeaLiT 历史学家提出的一项实际信息需求。

为检索结果并创建图表，ResearchSpace 会在内部将用户交互转换为 SPARQL 查询，并在 SeaLiT 知识图谱上执行。例如，下面的 SPARQL 查询会检索曾在最终目的地为 马赛 的船舶上担任船员的人员：

```
1 PREFIX crm: <http://www.cidoc-crm.org/cidoc-crm/>2 PREFIX sealit: <
http://www.sealitproject.eu/ontology/>34 SELECT DISTINCT ?person WHERE {5 ?ship
sealit:voyages?voyage .6 ?voyage sealit:finally_arriving_at <
https://rs.sealitproject.eu/kb/location/Marseille> ;7 crm:P14_carried_out_by?person }
```

[18]ResearchSpace 查询链接：https://tinyurl.com/2p8ky96e

为按居住地对人员进行分组并显示图表，执行以下 SPARQL 查询以检索相关数据：

```
1 前缀 crm: <http://www.cidoc-crm.org/cidoc-crm/>2 前缀 sealit: <
http://www.sealitproject.eu/ontology/>3 前缀 rdfs: <http://www.w3.org/2000/01/rdf-schema#>45
选择 不同的 ?location?locationName（计数（?person）作为?numOfPersons）其中 {6 ?ship sealit:
航次?voyage .7 ?voyage sealit:最终_抵达_于 <https://rs.sealitproject.eu/kb/location/Marseille> ;8
crm:P14_由_执行_者?person .9 ?person crm:P74_拥有_当前_或_曾经_居住地 ?location .10 ?location
rdfs:标签 ?locationName .11 } 按 ?location ?locationName 分组 按 ?locationName 排序
```

此类查询也可以利用 RDFS 推理规则，例如基于 子类关系 和子属性关系的规则。一个例子是使用 CIDOC CRM 属性 *'P9 consistsof'* 来获取特定船舶的所有航行相关活动（从某地点离开、抵达某地点、经过或穿越某地点、装载货物、卸货），如下方的 SPARQL 查询所示：

```
1 PREFIX crm: <http://www.cidoc-crm.org/cidoc-crm/>2
PREFIX sealit: <http://www.sealitproject.eu/ontology/>3
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>45
SELECT DISTINCT ?activity?activityName WHERE {6 <
SHIP-URI> sealit:voyages?voyage .7 ?voyage crm:P9_
consists_of ?activity .8 ?activity rdfs:label ?activityName }
```

在这种情况下，我们利用属性*'P9 consists of'* 是以下属性的超属性这一事实：*'consists of leaving'*、*'consists of arrival'*、*'consists of passing'*、*'consists of loading'*和*'consists of unloading'*。

利用 ResearchSpace 平台对集成数据进行查询，能够（直接或间接）回答的历史学家研究问题／信息需求的类型，主要取决于实际档案材料的内容——这些材料经过转录，并依据 SeaLiT 本体转换为资源描述框架（RDF）——而较少取决于本体本身。具体而言，本体的设计考虑了社区需求和实物证据。因此，如果转录文本（以及由此转换得到的数据）中存在回答某项信息需求（或查找与该需求相关的重要信息）所需的数据，那么便可以通过检索重要的相关信息，完整或部分地回答该问题。例如，在 SeaLiT 的案例中，有些转录文本（FAST CAT 记录）包含未填写完整的表格，这是因为某些档案文献没有提供相应信息，或者历史学家在数据转录时没有填写相关栏目（计划稍后补填）。在这种情况下，缺失信息是回答相关信息需求所必需的，因此这些需求无法得到满足。未来，如果出现本体无法建模的新类型信息（及相应的信息需求），就会对本体进行扩展或修订，并发布新版本。

关于不完整信息，实体属性缺失（例如，某艘船的建造地点未知）在历史档案研究中十分常见，但对历史学家而言，这些信息也很重要，因为它们可能影响定量分析结果的解读。我们的 ResearchSpace 配置将缺失信息表示为“未知”值，例如在条形图中显示“未知”列。

6 使用与可持续性

如前所述，这一本体是在 SeaLiT 项目背景下创建并使用的，用于将从海事史档案文献中转录的数据转化为丰富的语义网络。语义网络中的集成数据使一大群海事史学家能够（通过 ResearchSpace 平台提供的用户友好界面）对转录材料进行定量和定性分析，并发现与其研究需求相关的重要信息。

在 SeaLiT 项目结束后，预计相关活动仍将继续开展，这将依托希腊研究与技术基金会（FORTH）旗下两个参与机构的紧密合作：地中海研究所（SeaLiT 项目的协调方）和计算机科学研究所（Sea LiT 项目的数据工程合作方）。尤其是，一旦需要转录新的档案材料类型并将其整合到 SeaLiT 知识图谱中，本体就会扩展。

通过参与相关社群，我们确保了本体的长期可持续性，尤其是 CIDOC-CRM 特别兴趣小组[19] 和历史数据联盟[20], 。后者是一个国际联盟，旨在为历史研究中的数据建模、整理和管理建立一套通用方法。目前，其他（正在进行的）历史／档案研究领域项目已对使用（并可能扩展）该本体表现出兴趣。此外，模型中关于雇佣关系和付款的部分正被考虑用于创建一个新的 CIDOC-CRM 家族模型，以描述社会交易和纽带（CIDOC-CRM 特别兴趣小组对此已有相关讨论；参见议题 420 和 557[21]）。

7 结论

我们介绍了 SeaLiT 本体的构建与应用。SeaLiT 本体是 CIDOC-CRM 的扩展，旨在对海事史领域的数据进行建模与集成。该本体通过提供一个共同且可扩展的语义框架（通用语言），促进人们对海洋史信息形成共同理解，从而实现基于证据的信息集成。我们提供了该本体的规范、RDFS 和 OWL 实现，以及利用该本体将大量、多样的档案文献整合为丰富语义网络的知识图谱。我们还介绍了一个实际运行的应用（ResearchSpace 部署），它基于这些知识图谱运行，支持海事史学家通过用户友好界面探索和分析集成数据。

近期，我们计划：a）根据新的数据建模需求，研究本体可能的扩展；b）改进规范文档中类和属性的范围注释，并补充更多示例（随后发布新版本体）；c）创建并公开该本体的 JSON-LD 上下文，以便在基于 Web 的编程环境中使用。

致谢

本研究获得欧盟地平线 2020研究与创新计划的资助，资助来源包括：i）玛丽·斯克沃多夫斯卡-居里资助协议编号 890861（ReKnow 项目）；以及 ii）欧洲研究委员会（ERC）资助协议编号 714437（SeaLiT 项目）。

参考文献

[1] Francesco Beretta。2021年。历史研究面临的挑战：利用协作本体管理环境（OntoME）使数据符合 FAIR 原则。语义网 12, 2 (2021)，279–294。[2] Apostolos Delis。2020年 32, 2 2020 464 478。地中海海事史十字路口上的航海生活。《国际海事史杂志》( )，–。

[19]https://www.cidoc-crm.org/sig-members [20]
http://dataforhistory.org/members [21]
https://cidoc-crm.org/issue_摘要

[3] Martin Doerr。2003年。CIDOC 概念参考模块：一种实现元数据语义互操作性的本体方法。人工智能杂志 24, 3 (2003), 75–75。[4] Martin Doerr、Athina Kritsotaki 和 Katerina Boutsika。2011年。事实论证——一种用于构建断言的核心模型。《计算与文化遗产期刊》（*JOCCH*）3, 3 (2011)，1–34。[5] Pavlos Fafalios、Kostas Petrakis、Georgios Samaritakis、Korina Doerr、Athina Kritsotaki、Yannis Tzitzikas 和 Martin Doerr。2021年。FAST CAT：面向数字人文语义互操作性的协作式数据录入与整理。*ACM*《计算与文化遗产》期刊(2021年)。[6] Fabien Gandon。2002年。分布式人工智能与知识管理：企业语义网中的本体与多智能体系统。博士论文。尼斯·索菲亚·安蒂波利斯大学。[7] John B Hattendorf。2012年。当今海事史。历史视角 50, 2 (2012)，34–36。[8] Martin Hepp。2008年。Goodrelations：一种用于描述网络上产品、服务及其供应信息的本体。载于国际知识工程与知识管理会议。Springer，329–346。[9] Gerald Hiebel、Martin Doerr 和 Øyvind Eide。2017年。CRMgeo：CIDOC-CRM 的时空扩展。《国际数字图书馆杂志》18, 4 (2017)，271–279。[10] Konstantinos Kotis 和 George A Vouros。2006年。以人为中心的本体工程：HCOME 方法论。《知识与信息系统》10, 1 (2006)，109–131。[11] Konstantinos I Kotis、George A Vouros 和 Dimitris Spiliotopoulos。2020年。用于演进现存及复用本体的本体工程方法论：现状、趋势、发现与建议。《知识工程评论》35 (2020)。[12] Yannis Marketakis、Nikos Minadakis、Haridimos Kondylakis、Konstantina Konsolaki、Georgios Samaritakis、Maria Theodoridou、Giorgos Flouris 和 Martin Doerr。2017年。X3ML 映射框架：文化遗产及其他领域的信息集成。《国际数字图书馆杂志》18, 4 (2017)，301–319。[13] Albert Meroño-Peñuela、Ashkan Ashkpour、Marieke Van Erp、Kees Mandemakers、Leen Breure、Andrea Scharnhorst、Stefan Schlobach 和 Frank Van Harmelen。2015年。用于历史研究的语义技术：综述。语义网 6, 6 (2015)，539–564。[14] Franco Niccolucci。2017年。使用 CIDOC CRM 记录考古科学。《国际数字图书馆杂志》18, 3 (2017)，223–231。[15] Dominic Oldman 和 Diana Tanase。2018年。通过在 Re-searchSpace 中连接研究人员、数据与实践，重塑知识图谱。载于国际语义网会议。Springer，325–340。[16] Kostas Petrakis、Georgios Samaritakis、Thomas Kalesios、Enric García Domingo、Apostolos Delis、Yannis Tzitzikas、Martin Doerr 和 Pavlos Fafalios。2021年。海事史档案来源的数字化、整理与可视化：以十九世纪和二十世纪的船舶航海日志为例。*Drassana* 28 (2021)，60–87。[17] H Sofia Pinto、Christoph Tempich 和 Steffen Staab。2009年。在分布式世界中使用 DILIGENT 开展本体工程与演进。载于本体手册。Springer，153–176。[18] Maria Theodoridou、Yannis Tzitzikas、Martin Doerr、Yannis Marketakis 和 Valantis Melessanakis。2010年。通过扩展 CIDOC CRM 对来源信息进行建模和查询。《分布式与并行数据库》27, 2 (2010)，169–210。

SEALITONTOLOGY 的类层次结构

E1 CRM 实体 - E2 时间实体 - -
E4时期 - - - E5 事件 - - - - E7 活
动 - - - - - 航行 - - - - - 抵达 - - - - -
离开 - - - - - 通过 - - - - - 装载 - - -
- - 卸货 - - - - - 撤销船旗 - - - - -
卸载 - - - - - 民事登记 - - - - - 船
舶登记

- - - - - E11 修改 - - - - - - 船舶维修 - - - - - - E12 生产 - - - - - - - 船舶建造 - - - - - 服务报酬 - - - - - - 物品报酬 - - - - - - 劳动报酬 - - - - - - - 船员薪酬 - - - - - 教学单位 - - - - - - 课程 - - - - - - 章节 - - - - - 服务 - - - - - - 就业 - - - - - E13 属性指派 - - - - - - 晋升 - - - - - 惩罚 - - - - - 招聘 - E53 地点 - - 国家 - E54 尺寸 - - 马力 - - 持续时间 - - 吨位 - E77 持久项 - - E39 行动者 - - - E74 群体 - - - - 船籍港 - - E70 事物 - - - E71 人造物 - - - - E24 人造实体 - - - - - E22 人造物品 - - - - - - 船舶 - - - - - - 弹药 - - - - E28 概念对象 - - - - E55 类型 - - - - - - 语言能力 - - - - - - 识字状况 - - - - - - 航海类型 - - - - - - 职业 - - - - - - 宗教状况 - - - - - - 性别类型 - - - - - - 社会地位 - - - - - - 主题 - - - E72 法律对象 - - - - E90 符号对象 - - - - - E41 称谓 - - - - - - 船名 - - - - - - E42 标识符 - - - - - - - 船舶标识 - - - - - E73 信息对象

- - - - - - E29 设计者 程序 - - - - - - - 劳动合同 -
法律对象关系 - - 船舶所有权阶段 - - - 持股 - -
具有时间效力的法律文件

B SeaLiT 本体的属性层级

属性	定义域	值域
P1 由……标识（标识……）	E1 CRM 实体	E41 称谓
- 具有船舶标识（船舶标识用于标识）	Ship	船舶标识
P2 具有类型（是……的类型）	E1 CRM 实体	E55 类型
- 具有导航类型（是导航类型的）	Ship	导航类型
- 具有语言能力（是语言能力的）	E21 人	语言能力
- 具有识字状况（是……的识字状况）	E21 人	识字状况
- 具有社会地位（是……的社会地位）	E21 人	社会地位
- 具有性别类型（性别类型所属）	E21 人	性别类型
- 具有职业（职业所属）	E21 人	职业
- 具有宗教状况（是宗教状况的）	E21 人	宗教状况
- 具有主题（是主题的）	教学单位	主题
P9 由……组成（是……的一部分）	E4时期	E4时期
- 由离开组成（离开是……的一部分）	航行	离开
- 由抵达组成（抵达是其一部分）	航行	抵达
- 由通过组成（通过是其一部分）	航行	通过
- 由装载组成（装载是其一部分）	航行	装载
- 由卸货组成（卸货是其一部分）	航行	卸货
P12 发生于……在场时（在场于）	E63 存在的开始	E77 持久项
- P92 使之产生（被带入……）	E63 存在的开始	E77 持久项
- - P108 已产生（由……产生）	E12 生产	E24 物理人造物
- - - 建造（由……建造）	船舶建造	Ship
- P11 有参与者（参与于）	E5 事件	E39 行动者
- - P14 由……执行（执行）	E7 活动	E39 行动者
- - - 由船长驾驶（驾驶）	航行	E39 行动者
- - - 登记者（负责……）	船舶登记	船籍港
- - - 资金提供者（提供资金）	服务报酬	E39 行动者
- - - 由……中介（担任中介）	服务报酬	E39 行动者
- - - 向其提供的款项（收到款项）	服务报酬	E39 行动者
- - - 提供服务者（所提供的服务）	服务	E39 行动者
- - - - 提供就业者（所提供的……）	就业	E39 行动者
- - 有学生（学生属于）	教学单位	E39 行动者
- P31 修改了（被修改）	E11 修改	E18 物理事物
- - 已维修（由……维修）	船舶维修	Ship
- 航行于（航次）	航行	Ship
P15 受……影响（影响）	E7 活动	E1 CRM 实体
- P17 由……促成（促成）	E7 活动	E1 CRM 实体
- - 用于航行（促成的付款）	船员薪酬	航行
P43 具有尺寸（是……的尺寸）	E70 事物	E54 尺寸
- 具有吨位（是……的吨位）	Ship	吨位

- 具有马力（是.....的马力）
P46由.....组成（构成.....的一部分）
- 具有弹药（属于弹药）
P90具有数值
- 具有持续时间数值
P107i 是现任或前任成员（拥有.....）
- 工作于（是其工作地点）
P140 将属性赋予（曾由其赋予属性）
- 涉及（曾由其晋升）
P141 指派（由.....指派）
- 晋升为身份类型（身份类型为.....）
- 晋升至就业职位类型（.....）
P173 开始时间早于或等于.....的结束时间
- P174 开始时间早于.....的结束时间（结束时间晚于.....）
- - P175 开始时间早于或同时于.....的开始时间（.....）
- - - 开始（由.....开始）
- - P184 结束时间早于或同时于.....的结束时间（结束.....）
- - - 结束（由.....结束）
P191 持续时间（是.....的持续时间）
- 具有持续时间（持续时间为）
具有旗帜（曾是旗帜）
具有船员人数容量
以姓名命名（名称为）
具有船旗（为船旗）
具有船舶标识（船舶标识属于）
登记（由.....登记）
拥有所有者（是该阶段的所有者）
- 有股东（按份额参与）
是所有权阶段（具有所有权阶段）
- 是股份持有阶段（具有持股）
以名称为依据的所有权（具有所有权的名称）
由.....初始化（初始化.....）
- 所有权由.....初始化（初始化.....）
由.....终止（终止.....）
- 所有权由.....终止（终止.....）
拥有所有者（是该阶段的所有者）
- 有股东（按份额参与）
所持股份
发生于（时间）
曾经或目前拥有（曾经是.....）
撤销船旗（在.....时撤销船旗）
最终抵达（是抵达地点）
始于（是起始地点）
目的地（是其目的地）
装载（由其装载）
卸载（由其卸载）
在某地（是抵达地点）
来自某地（是离开地点）
途经地点（是通过地点）
经过地点（即经过发生的地点）

Ship
E18 物理事物
Ship
E54 尺寸
持续时间
E39 行动者
E21 人
E13 属性指派
晋升
E13 属性指派
晋升
晋升
E2 时间实体
E2 时间实体
E2 时间实体
招聘
E2 时间实体
卸载
E52 时间跨度
E52 时间跨度
Ship
Ship
船舶建造
船舶登记
船舶登记
船舶登记
船舶所有权阶段
持股
船舶所有权阶段
持股
船舶所有权阶段
法律对象关系
船舶所有权阶段
法律对象关系
船舶所有权阶段
船舶所有权阶段
持股
持股
法律对象关系
E39 行动者
撤销船旗
航行
航行
航行
装载
卸货
抵达
离开
通过
通过

马力
E18 物理事物
弹药
E60 数值
E60 数值
E74 群体
E74 群体
E1 CRM 实体
E21 人
E1 CRM 实体
社会地位
职业
E2 时间实体
E2 时间实体
E2 时间实体
就业
E2 时间实体
就业
E54 尺寸
持续时间
国家
E60 数值
船名
国家
船舶标识
Ship
E39 行动者
E39 行动者
Ship
Ship
船名
E5 事件
船舶登记
E5 事件
撤销船旗
E39 行动者
E39 行动者
E60 数值
E52 时间跨度
具有临时效力的法律文件
Ship
E53 地点
E53 地点
E53 地点
E18 物理事物
E18 物理事物
E53 地点
E53 地点
E53 地点
E53 地点

用于服务（服务提供方）	服务报酬	服务
- 用于就业（雇佣方）	劳动报酬	就业
具有货币价值（是.....的价格）	服务报酬	E97 货币金额
适用于就业时期（是.....的就业时期）	劳动报酬	E52 时间跨度
已在.....中达成约定（是.....的约定）	劳动报酬	劳动合同
针对事物（事物的）	物品报酬	E18 物理事物
有名字	E21 人	E62 字符串
有姓氏	E21 人	E62 字符串
有当前年龄	E21 人	E60 数值
具有 ID（ID 的）	民事登记	E42 标识符
登记人员（人员被登记）	民事登记	E21 人
给予（受到惩罚）	惩罚	E39 行动者
与.....相关	E21 人	E21 人
学生人数为	教学单位	E60 数值