

## “七巧板”医学术语协作项目

### OMAHA 秘书处

#### 项目背景

在医疗行业中，临床术语的表达带有从业者的习惯因素，不同人通常对医学概念的表达不一致。而医学术语是促进健康医疗信息在语义层面互操作的关键内容之一，对术语采用科学的方法进行管理有助于提高对健康医疗数据挖掘、管理和应用的能力。本项目旨在联合行业中顶尖团队和个人成为“七巧板”中的一份子，通过协作众创的方式形成一套体系性的医学术语，来帮助各类健康医疗数据生态中的企业提高效率，推动开放数据过程中的标准形成。

#### 医学术语集研制意义

临床医学术语可服务于健康医疗信息技术企业和健康医疗数据分析企业。大体上，术语集的应用场景可分为如下三种：

- 1) 临床病历：医学术语集作为医疗机构电子病历系统/临床信息系统中的一部分，支持专业医务人员的日常表达形式，实现语义层面的信息检索和再利用，包括不同层面上信息的抽取、分析、集成；
- 2) 知识表达：医学术语集可以结合知识库，包括参考资料、临床指南、临床决策指南、知识图谱等，为健康医疗信息提供基于术语本体库的语义支撑，用于垂直搜索、临床辅助决策、智能医疗等；
- 3) 集成和分析：帮助对健康医疗数据进行后结构化处理，包括数据清洗、标准化和语义识别，临床术语集可以为处理医疗健康数据的企业提供底层的标准化工具，实现结构化解析、命名实体标准提取、文本标记等。

#### 目前成果

截止 2019 年 12 月 31 日，OMAHA 七巧板医学术语集已积累 79 万概念、101 万术语、265 万关系、81 万映射，涉及领域有疾病诊断、症状、药品、检验检查、手术操作、解剖结构、微生物、基因、基因突变等，并在每个季度第一个月的 20 号持续进行发布。

#### 项目组织方

开放医疗与健康联盟（Open Medical and Healthcare Alliance，OMAHA）是由浙江数字医疗卫生技术研究院发起的行业联盟，联盟旨在了解和梳理行业内从业企业的共同需求，将个别企业难以独立完成或难以取得短期经济效益的可共用、可复用的构件、组件、中间件，通过开放开源、众包协作的研发方式供联盟企业共同使用，从而降低单个企业的研发成本，提高行业产品的标准化规范化。

OMAHA 联盟在成立之初就切入医学术语标准研究的工作，并在 2017 年 6 月正式发起“七巧板”医学术语协作项目，通过联合行业内顶尖企业、学术机构及医务工作者等参与方，共同构建符合中文临床环境的医学本体术语集，并建立起持续的维护更新机制，为整个医疗行业提供稳定可靠的医学术语服务。目前为止，“七巧板”医学术语项目已成功开展十期，吸引 7 家优秀机构及 100 多名医学专家参与医学术语集构建工作。

## 协作要求

“七巧板”医学术语项目已建立了一支拥有 700 多人的志愿者团队，团队成员均为擅长某一医学领域的医护从业者、医学院校教授和学生、优秀企业、病案编码员等。为保证项目质量，仅允许符合一定要求的团队和个人参与该项目，对项目参与者准入的评估条件包含如下内容。

### 团队资质要求

团队是指优秀企业、医学院校、科研机构等以团队作为主体参与该项目的形式。团队资质要求为：

- 团队对联盟（非营利性组织）向行业提供底层的共性基础服务的理念认同；
- 团队对于用协作的方式来建立医学术语集的理念认同；
- 团队的主要发展方向与数据和技术驱动医疗健康决策相关；
- 具有一定的医学术语管理能力，内部有一定的术语积累，并采用分类体系或语义关联组（包括本体）的方式来管理术语；
- 配备项目管理人员，以及医学术语协作人员，协作人员应当具备临床医学类专业背景，且对 SNOMED CT, UMLS, RxNorm, ICD-10 等医学术语的架构和内容有一定的了解；
- 在项目规划的时间内，每名人员投入足够的工作时间在项目上，确保按时完成任务；
- 如参与方为企业，需成为 OMAHA 高级版服务机构，拥有对所有 OMAHA 已形成的术语资源获取的权限。

### 个人资质要求

个人是指临床医护从业者、医学生、编码员，具有临床背景的企业内工作人员以个人为主体参与该项目的形式。个人资质要求为：

- 对用协作的方式来建立医学术语集的理念认同；
- 具有一定临床医学类专业背景，目前正在从事临床护理工作，或在校医学生，或具有医学相关教育背景的非临床工作人员，或病案室编码员；

- 有充足的时间，按时完成相关工作；
- 具有一定的自主学习能力，愿意学习医学术语集构建、工具使用等知识；
- 责任心强，愿意认真对待每一次工作。

## 协作内容

目前，该项目的主要工作内容有如下三项，项目组织方将根据术语集实际情况和协作方能力，不定期组织开展相应的项目。联盟将根据协作内容，邀请不同的团队或个人参与：

### 贡献术语语料

参与的团队或个人可以提供来自于真实世界的术语语料，如临床病历、科研文献、教科书等。术语语料可以是单个的词，也可以具有同义关系的同义词组。

要求：术语语料需来自于真实世界，不可以是杜撰生成。术语语料质量要有保证，不能出现乱码、错别字等。术语语料尽可能是分领域整理，不同领域术语语料不能混杂在一起。

### 参与术语集构建

术语集构建工作主要包括为概念新增术语、为概念新增关系、审核概念下的术语和关系这三项工作。相应的工作主要在 protégé 本体编辑工具或 OMAHA 术语维护平台上开展。工作开展前会进行相应的培训。

要求：参与者需在满足基础资质的条件下，符合当期七巧板项目的领域要求，如开展呼吸领域的术语集构建项目时，参与者需具备呼吸领域的专业背景。参与审核工作的志愿者专业背景必须是临床住院医师及以上职称。

### 参与映射构建

映射工作是指建立 OMAHA 七巧板医学术语集与行业已有术语集之间的连接，如与医保药品编码、ICD-10、ICD-9-CM-3 等之间的映射。

要求：参与术语集与 ICD-10、ICD-9-CM-3 的映射工作的志愿者，需是病案室编码员。其它领域的映射工作，参与志愿者需具备该领域专业背景。

## 协作回报

为感谢所有参与该项目的志愿者团队及个人，联盟建立了相应的协作回报方案，，希望可以表达我们对所有志愿者辛苦付出的感激之情，同时也鼓励更多志愿者加入到我们的队伍中来。

### 贡献榜表彰

按照不同的协作内容类型，联盟设立相应的参与方贡献值计算方案，并按照贡献值情况设立贡献值榜单，在 OMAHA 联盟官网进行发布，并定期通过 OMAHA 联盟公众号进行宣传。贡献值榜单每季度更新一次。

贡献值榜单地址：<http://term.omaha.org.cn/?/termcommunity/contribution/>。

## 报酬回馈

按照不同的协作内容类型，联盟设立相应的报酬回馈方案，在每期项目结束后，按照志愿者完成的工作量情况计算报酬。报酬计算方案如下，并会根据项目实际开展情况作出一定的调整：

- 提供术语语料（无同义关系）：0.1 元/条术语（词）
- 提供术语语料（有同义关系）：0.2 元/条术语（词）
- 为概念新增术语：0.3 元/条术语
- 为概念新增关系：0.4 元/条关系，关系包括属性关系、层级关系
- 审核概念下的术语或关系：0.4 元/条术语或关系
- 新增映射：0.4 元/条映射

以上所有报酬计算均是以“有效成果”为前提，即必须是有价值的成果，如 OMAHA 术语库已包含的、错误的内容等不在报酬计算范围内。以团队为主体参与协作的报酬给予机构，以个人为主体参与协作的报酬给予本人。

## 活动邀请

所有志愿者都可以免费参加 OMAHA 联盟举办的各项行业活动，如城市沙龙、OMAHA 年会等。联盟将在每次会议举办前以短信、邮件、联盟公众号等方式邀请志愿者参加。

## 成果归属和保护期

“七巧板”医学术语协作项目成果知识产权归组织方所有。参与方可以对协作成果进行改进升级，对协作成果及改进升级后成果具有不可撤销的全球范围内永久免费使用权。参与方有权使用协作成果进行产品开发及改进升级，但不得向其他组织机构或企业直接售卖协作成果。参与方利用协作成果独立进行产品开发的成果归乙方所有。

基于 OMAHA 的长期愿景，最终希望通过协作能够形成不断完善的中文医学术语集，因此，协作成果将以合适的形式提供给行业，从而提升行业整体效率和企业术语管理能力。参与方可向组织方申请对项目协作成果设置保护期，保护期限为半年，保护期内所有成果（包括在此期间持续维护更新而取得的成果）不会对行业内的其他非本项目参与方提供，参与方可在此时间段内以此为基础供企业内部研究和使用的。

## 协作联系

团队协作可联系 OMAHA 联盟数字医学知识中心负责人徐美兰女士，邮箱 [meilan.xu@omaha.org.cn](mailto:meilan.xu@omaha.org.cn)。

个人协作可通过扫描下方二维码进行报名，收到报名信息后，OMAHA 联盟同事将与您联系。

