

# 众包协作翻译平台

“众包”（Crowdsourcing）是指一种分布式的问题解决和生产模式。美国计算机杂志《连线》记者杰夫·豪对“众包”的定义是：“一个公司或机构把过去由员工执行的工作任务，以自由自愿的形式外包给非特定的（而且通常是大型的）大众网络的做法。”<sup>[1]</sup> “‘众包’这一概念实际上是源于对企业创新模式的反思。传统的产品创新方法是，首先由生产商对市场进行调查，然后根据调查结果找出消费品的需求，最后再根据需求设计出新产品，但这种创新的投资回报率通常很低，甚至血本无归。而如今，随着互联网的普及，消费者的创新热情和创新能力愈发彰显出其巨大的能力和商业价值，以‘用户创造内容’（user-generated content）为代表的创新民主化正在成为一种趋势。”<sup>[1]</sup> 大家所熟知的人人猎头、易到用车等均是国内较为成功的众包平台实例，国外著名的众包平台包括亚马逊的 Mechanical Turk、MIT Climate CoLab、Zooniverse 等。

“众包”模式一经兴起，在语言服务、电商乃至物流行业的发展态势如火如荼。把过去交由专职译员执行的任务，外包给网络上的志愿译员来完成，就形成了所谓的“众包翻译”。本次 C++ 大作业针对众包协作翻译平台展开。

参考国内现有的“众包翻译”平台“译言古登堡计划”（<http://g.yeeyan.org/>），针对一个翻译任务，本次大作业将翻译流程做以简化，抽象成下面的流程图。

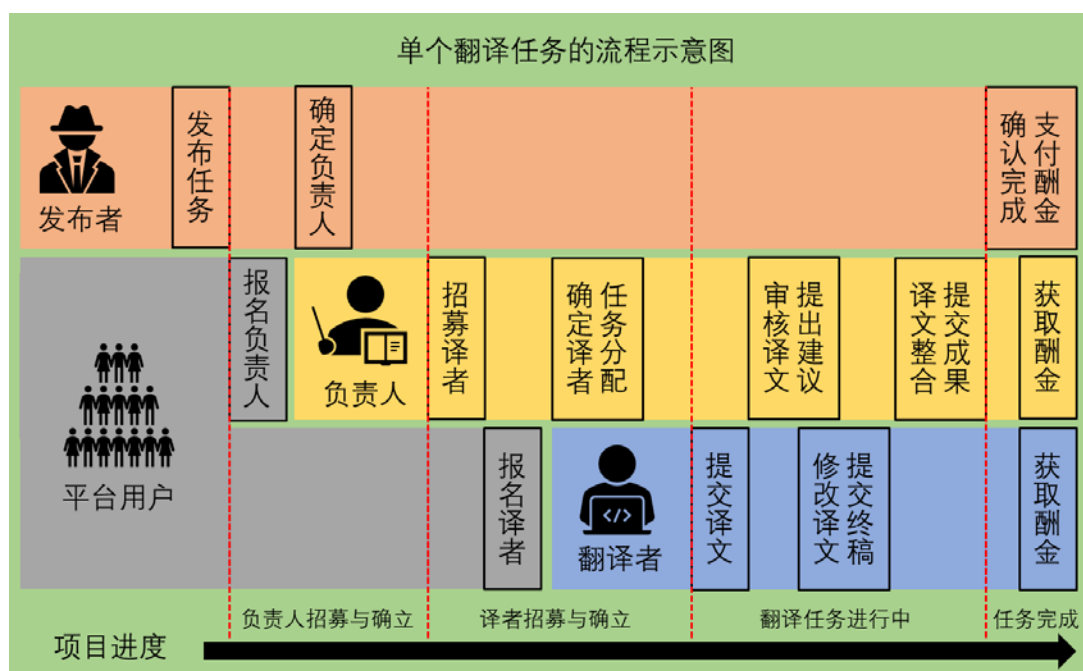


图 1. 众包协作翻译流程图

- 首先，任务发布者在平台上发布任务，例如，英译中翻译 Kevin Kelly 著作《The Inevitable》共 12 章书籍的翻译任务，并在平台上征集负责人。
- 其次，根据报名情况，任务发布者确定该书的负责人，其职能是统筹协调负责整本书的翻译。

<sup>[1]</sup> <https://wiki.mbalib.com/wiki/众包>

- 负责人会为该任务招募并确定合格的译者。此后, 将整个翻译任务拆分成若干个子任务, 例如, 将《The Inevitable》的每一章翻译作为一个子任务, 分配给各个译者。
- 译者翻译所领取的子任务, 负责人对其所提交的译文进行审核并提出修改意见。译者需要根据负责人的修改意见对译稿做相应的改动。
- 所有章节的译文审核通过后, 负责人将全部译稿提交给任务发布者。同时, 任务发布者支付酬金, 负责人和译者得到相应的酬金, 该任务完成。

根据以上的简化流程, 一个翻译任务中所涉及的用户可能承担三种角色, 它们分别是任务发布者, 负责人以及译者。每一个角色在整个翻译任务中会有各自的分工:

- 对于任务发布者 (每个任务只有一人), 他是该任务的发起者, 需要提交待翻译资料, 确立任务负责人, 收取译稿以及支付酬金等;
- 对于负责人 (每个任务只有一人), 他主要负责选取有资质的译者, 并将翻译资料以子任务的形式发放给各个译者。最终需要对译者提交的译稿进行质检, 并将质检后的全文译稿整合提交给任务发布者;
- 对于译者 (每个子任务只有一人), 其主要职能是翻译负责人分配的子任务, 并对负责人的评审反馈给予相应的修改。

需要注意的一点是, 任何用户在一个翻译任务中, 只能承担任务发布者, 负责人以及译者其中之一角色; 而对于不同的任务, **用户在其中的角色可以不同**。每个用户会根据参与过翻译任务的多少而有相应高低的积分。只有积分达到一定水平的用户才有成为项目负责人的权限。

#### 大作业任务具体说明:

1. 要求实现用户、任务 (发布的任务, 负责的任务以及翻译的任务) 等类。**要求体现继承关系**。如果实现了多态, 会有适当的加分。
2. 平台用户需要注册, 每一个用户应该具有唯一的标志 (手机, 用户名或身份证), 同时还应有相应的个人基本信息, 同时提交语言资质证明 (此处仅用于示意即可, 如文字输入“TOFEL 110 分”、“英语专业 8 级”等), 该语言资质证明会决定用户的初始积分 (如何制定标准此处不做要求, 但需要有体现)。
3. 已注册的用户在登录时需要匹配账号和密码。
4. 用户可以查看个人基本信息、账户余额以及积分, 并且能够维护信息以及充值操作 (这里的充值操作只作示意性要求即可)。
5. 所有用户都能够发布翻译任务, 从而成为任务发布者。发布新任务过程中, 用户需要编写任务简介, 上传翻译内容, 指定任务周期以及负责人报名的截止时间等, 还要说明该任务的金额 (为了简化起见, 在翻译内容部分, 我们以几个短句来代替整个翻译任务, 例如翻译任务就是“Hello everyone. This is C++ class.”)。翻译任务应当至少包括英译中和中译英两类。
6. 用户可以在平台中查询所有其参与的任务。根据任务的状态 (是否完成) 以及用户在任务中承担的角色 (译者, 任务发布者, 负责人), 每个任务应该有不同的标识。并且对于承担着不同角色的任务, 用户进入该任务之后能执行不同的功能。
  - 对于该用户发布的任务: 在负责人报名截止之前能够查看负责人报名状态, 在负责人报名截止后可以指定负责人。对于负责人提交完成译稿的任务, 该用户能够确认任务完成, 并支付酬金 (此处的支付只是起到示意效果即可, 且该操作之后负责人和译者才能够领取相应的酬金);
  - 对于该用户负责的任务: 该用户能够获取任务原文, 查看当前报名的译者, 当译者报名时间截止后, 需要确定译者并分配章节任务 (这里所谓的章节任务就是之前示

意性翻译任务的划分，例如将翻译任务“Hello everyone. This is C++ class.”相应手动划分为两个子任务“Hello everyone.”以及“This is C++ class.”两个字符串。)在翻译任务进行中，可以查看每一个译者相应的译文，可以给出相应的反馈。具有能够将全部译文合并提交给发布者的接口。在任务结束后，能够领取酬金。

- 对于该用户翻译的任务：用户能够查看翻译子任务，可以查看上次暂存（提交）的翻译文字并可以删改，继续翻译并暂存和提交。同时该用户还应该能够看到负责人给出的历史反馈信息（此处的反馈信息也是示意即可）。在任务结束后，能够领取酬金。
7. 每一个用户能够查看除自己所涉及翻译任务以外的其他任务。每一个任务体现出当前所处的不同状态（例如，“正在招募负责人”，“正在招募译者”，“翻译任务进行中”等）、任务的酬金等。
  8. 对每一个正在进行且该用户目前没有参与的任务，用户可以从平台中获取到该任务的描述。根据任务所处的状态不同。用户可以有不同的操作。
    - 对于正在招募负责人的任务：用户可以在该任务中申请成为任务负责人。对于等级不够的用户和已经报名的用户应该有相应的提示。
    - 对于正在招募译者的任务：用户可以在该任务中申请成为译者。对于已经报名的译者应该有相应的提示。
    - 对于翻译进行中和已经完成任务，页面中只显示任务描述，负责人和译者等信息。
  9. “众包翻译”平台应该具有“消息盒子”功能，在用户有相关操作和用户状态以及用户所涉及的任务有状态变化时，消息盒子都应该及时更新起到提醒功能。
  10. 在该众包协作翻译平台中，为了简单起见，我们默认每一个用户在完成翻译任务且译文审稿合格后其相应的积分都会+1。对于成为负责人的资格，则可以相应的简化为具有翻译完成一定数量的任务（以译者身份）的用户。为了方便调试和展示，允许后台直接修改用户属性，如何实现不作具体要求，不属于考核范围。

#### **加分项目（包括但不限于）：**

1. 个性化推荐；
2. 增加第四方“审核人”角色， he可以和负责人共同审验译者的译稿。成为“审核人”的要求应当高于成为“负责人”的要求。
3. 增加平台功能，平台是独立于用户的其他工作人员。他们的主要工作是对用户提交的“资质证明”进行筛查，同时可以后台修改每个用户的积分。
4. 其它功能，如译者互校、多种（三种及以上）语言之间的互译等。

#### **其它说明：**

1. 程序核心功能应该全部采用 C++类和对象来实现，体现 OOP 的抽象、封装、继承、多态等设计思想和关键技术，要求充分运用课程中所学 C++的面向对象的知识，以单纯的 C 语言写出的面向过程的程序是不能得分的。
2. 需要有基于数据库或文件系统的数据存储机制（不同存储机制得分不同），保证计算机断电后数据不会丢失。用户登录后，相关数据需要先读入内存中再进行操作。对各种数据的操作（查询、修改、添加、删除等）应该在内存中进行，并定期保存至硬盘中。严禁在数据库中直接对数据进行操作。用户登出后，更新的数据需要写入数据库或文件系统中。
3. 本作业不要求实现“真正的多用户系统”（即多用户可同时登录的系统），可以用单用户系统模拟多用户系统，即每次只有一个用户登录系统。其它用户需登录时，之前用户需要

退出。

4. 可以有图形界面,也可以没有图形界面(得分不同)。图形界面开发可以采用 MFC 或 QT。用户界面应做到功能完善、用户友好,图形界面应做到界面美观。
5. 程序应具备一定的容错能力,对用户的不规范操作与误操作有一定的容忍度,不会轻易崩溃。
6. 程序功能完整正确,程序结构良好,代码整洁规范。代码中应该有合理的注释。
7. 大作业要求一人一组单独完成,需提交项目报告(具体要求见模板,提交电子版即可)。项目报告的字数没有要求,其内容应该是实质性的设计内容,最好不要太多粘贴界面之类的图片,报告的内容质量直接决定了大作业的得分。
8. 最终需要提交可以运行的源代码(含注释说明)、可执行程序、项目报告。大作业重点考查对课堂知识的掌握情况和运用能力,合理应用的知识点越多,程序功能越强大,得分就越高。
9. 大作业程序及项目报告要求完全独立完成。大作业答辩时会随机抽查 1、2 处代码请开发者解释,以确认大作业完成的独立性及课程知识的掌握程度。