

2023 年山东省“技能兴鲁”职业技能大赛  
—第三届山东省人工智能融合创新职业技能竞赛

人工智能训练师

赛  
项  
规  
程

山东省人工智能融合创新职业技能竞赛组委会

2023 年 8 月

# 目 录

一、赛项名称 .....	4
二、项目简介 .....	4
(一) 项目概要 .....	4
(二) 竞赛目的 .....	4
(三) 相关文件 .....	5
(四) 参赛对象 .....	5
(五) 竞赛方式 .....	5
三、选手应具备的能力 .....	6
四、竞赛模块 .....	8
(一) 理论知识竞赛 .....	8
(二) 实操技能竞赛 .....	9
(三) 命题方式 .....	10
五、评分规则 .....	10
(一) 理论知识竞赛成绩评定 .....	10
(二) 实操技能竞赛成绩评定 .....	11
1、评价分（主观） .....	11
2、测量分（客观） .....	11
(三) 评分流程说明 .....	12
六、项目特别规定 .....	13
(一) 违规行为 .....	13
(二) 赛场纪律 .....	14

七、竞赛相关设施设备 .....	15
(一) 实操技能竞赛设备 .....	15
(二) 竞赛场地平面图 .....	16
(三) 场地设备工具 .....	16
八、健康和安​​全 .....	17
九、绿色环保 .....	18
十、样题 .....	18

第三届山东省人工智能融合创新职业技能大赛

注：本项目技术描述是对本竞赛项目内容的框架性描述，正式比赛内容及要求以竞赛当日公布的赛题为准。

## 一、赛项名称

2023年山东省“技能兴鲁”职业技能大赛—第三届山东省人工智能融合创新职业技能竞赛—人工智能训练师。

## 二、项目简介

### （一）项目概要

人工智能训练师项目，主要考察人工智能训练师新职业从业人员的职业能力，包括人工智能相关理论知识以及数据采集和处理、数据标注、智能系统运维、业务分析、智能训练、智能系统设计等技能操作知识。本项目要求参赛选手能够根据竞赛题目的要求进行业务需求分析，利用提供的智能计算实训平台，基于相关理论及技能操作知识，分别对文本类数据、图像类数据、语音类数据进行处理与应用，完成典型场景应用系统的相关功能，通过真实的工作任务实施考察选手的综合职业能力。

本项目参考人力资源社会保障部等部委发布的“人工智能训练师”新职业信息的要求，结合当前新职业从事的工作领域和技能要求，参照人力资源和社会保障部等部委发布的《人工智能训练师国家职业技能标准》，同时结合当前人工智能的主流技术及典型行业应用场景设计本次竞赛的工作内容和考评标准。

### （二）竞赛目的

为深入贯彻落实习近平总书记对技能人才工作的重要指示精神，深入实施人才强国战略，推进“十四五”时期新职业领域技术技能人才队伍建设，增强新职业从业人员社会认同度，改善新职业人才供给质量结构，振兴新职业的创新能力和核心竞争力，促进就业和经济发展。充分发挥职业技能竞赛在促进技能人才培养、推动开展新职业技能培

训和弘扬工匠精神方面的重要作用。

根据人力资源和社会保障部公布的人工智能训练师国家职业标准，结合企业职业岗位对人才培养需求，充分对接相关行业新业态发展、人工智能产业链延伸交叉与技术进步需求。

通过组织人工智能训练师职业技能竞赛，进一步加深人工智能相关技术技能发展趋势的了解与认识，并向社会展示人工智能技术、推广人工智能技术与应用，促进人工智能技术竞赛和人工智能技术人才培养工作科学有序地发展。

### （三）相关文件

本项目规程是对本竞赛项目内容的框架性描述。除阅读本文件外，开展本技能项目竞赛还需配合其他相关文件一同使用：《关于组织开展 2023 年山东省“技能兴鲁”职业技能大赛的通知》。

### （四）参赛对象

竞赛设职工组和学生组两个组别，均为双人赛。凡从事相关工作的企（事）业单位职工及相关专业全日制在籍学生，均可报名参加相应赛项和组别的竞赛。各单位每组别限报 2 队，每队最多 2 人参赛，学生组每支参赛队的指导教师限报 2 名。已获得“山东省技术能手”的人员，不得以选手身份参赛。具有全日制学籍的在校学生不得以职工身份参赛。

### （五）竞赛方式

竞赛内容分理论知识竞赛和实操技能竞赛两部分，参赛选手须独立完成理论知识竞赛，双人协作共同完成实操技能竞赛。

理论知识竞赛统一在教室内进行，不分场次，竞赛时间为 60 分钟，满分 100 分，最终的理论成绩为两位选手的理论成绩平均值；

实操技能竞赛根据参赛队伍的抽签顺序在竞赛工位内进行，竞赛

时间为 120 分钟，满分 100 分。

各参赛队的总成绩=理论知识竞赛成绩\*20%+实操技能竞赛成绩\*80%。最终的竞赛名次按总成绩高低排定。

### 三、选手应具备的能力

参考《人工智能训练师国家职业技能标准》文件内容，具体为标准中人工智能应用产品实现这一职业方向的相关内容，参赛选手应当具备以下知识点和技能点。

	能力描述
1	人工智能相关理论知识
理论知识	个人需要了解和掌握以下理论要点： <ul style="list-style-type: none"><li>● 人工智能职业道德、职业守则相关知识。</li><li>● 人工智能产业应用相关知识。</li><li>● 人工智能发展现状及趋势相关知识。</li><li>● 人工智能热点问题和前沿研究相关知识。</li><li>● 人工智能行业相关法律、法规知识。</li><li>● 人工智能安全与隐私保护原则与标准相关知识。</li><li>● 人工伦理道德治理发展趋势、相关原则及标准。</li><li>● 人工智能专业技术知识：数学基础、编程基础、数据处理、软件工程、计算平台、机器学习、深度学习。</li><li>● 计算机操作知识。</li><li>● 常用办公软件使用知识。</li><li>● 人工智能文档规范、代码规范、质量保障规范等相关知识。</li><li>● 数据采集、标注、清洗、质量控制等数据工程相关知识。</li></ul>
2	文本类数据处理及应用
基本知识	<ul style="list-style-type: none"><li>● Python 编程基础：变量与数据类型、控制结构、函数与模块、类与对象。</li><li>● Python 常用科学计算库：NumPy、SciPy、Matplotlib、Pandas、Scikit-learn 等。</li></ul>

	<ul style="list-style-type: none"> <li>● Python 常用的运行环境：PyCharm、Jupyter Notebook、Linux 命令行等。</li> <li>● 经典机器学习的算法及其典应用场景。</li> <li>● 业务数据的类型及采集方式。</li> <li>● 数据清洗的概念、方法。</li> <li>● 常用的文本数据标注工具及方法。</li> <li>● 常用机器学习算法的模型构建、训练、优化、评估等知识。</li> </ul>
工作能力	<ul style="list-style-type: none"> <li>● 掌握 PyCharm、Jupyter Notebook、Linux 命令行等常用开发环境的使用方法。</li> <li>● 掌握 Python 语法，具备一定的编程能力。</li> <li>● 掌握原始业务数据的采集方法：数据库读取、靶站爬取、基于特定的设备或工具采集等。</li> <li>● 掌握数据清洗的方法：处理业务数据中的缺失值、错误值、重复值等，规范及统一格式。</li> <li>● 掌握数据可视化的方法。</li> <li>● 掌握常用数据标注工具的使用及标注方法。</li> <li>● 掌握常用机器学习模型的原理及典型应用场景。</li> <li>● 掌握模型的构建、训练、优化、评估、预测能力。</li> </ul>
3	<b>图像类数据处理及应用</b>
基本知识	<ul style="list-style-type: none"> <li>● Python 编程基础及其常用库。</li> <li>● PyCharm 开发环境的使用方法。</li> <li>● Linux 操作系统及常用命令。</li> <li>● 经典的机器学习算法原理及其典型应用场景。</li> <li>● 深度学习技术及其应用场景。</li> <li>● 基于 YOLOv5 目标检测模型训练自定义数据模型的方法。</li> <li>● 主流的人工智能开发框架/平台：PyTorch、Caffe、TensorFlow、Paddle Paddle。</li> <li>● OpenCV 图像处理相关知识。</li> <li>● 图像数据采集、预处理、数据加载方法。</li> </ul>

	<ul style="list-style-type: none"> <li>● 常用的图像数据标注工具及其标注方法。</li> <li>● 常用深度学习模型的选择、训练、调参、预测、模型评估、模型部署知识。</li> </ul>
工作能力	<ul style="list-style-type: none"> <li>● 掌握常用集成开发环境的使用方法及常用功能。</li> <li>● 掌握图像数据采集及预处理能力。</li> <li>● 掌握图像数据标注工具的使用及标注方法。</li> <li>● 掌握模型选择能力。</li> <li>● 掌握模型训练、参数优化能力。</li> <li>● 掌握模型评估能力。</li> <li>● 掌握模型部署与测试能力。</li> </ul>
4	语音识别技术应用
基本知识	<ul style="list-style-type: none"> <li>● Python 编程基础及其常用库。</li> <li>● 常用 Python 开发环境的使用方法。</li> <li>● 语音识别的概念、分类、方法、性能指标。</li> <li>● 语音识别技术的典型应用场景。</li> <li>● 常用的离线语音识别引擎及其使用方法。</li> <li>● 常用的在线语音识别引擎及其使用方法。</li> </ul>
工作能力	<ul style="list-style-type: none"> <li>● 掌握常用 Python 开发环境的使用方法及常用功能。</li> <li>● 掌握 Pyttax3 语音识别引擎的安装、使用方法。</li> <li>● 掌握 Google Speech API 语音识别引擎的安装、使用方法。</li> <li>● 掌握百度语音识别引擎的使用方法，能够基于百度语音识别技术进行语音识别、语音合成、实时语音转文字、语音播报等，并能够使用语音识别技术灵活的搭建人工智能综合项目。</li> </ul>

## 四、竞赛模块

### （一）理论知识竞赛

竞赛方式：以笔试形式进行，参赛选手须独立完成理论考核，计分采用百分制，竞赛时间为 60 分钟，竞赛成绩占总成绩的 20%，60 分及以上为及格。

竞赛内容：“人工智能训练师”国家职业技术技能标准的基础理论及相关知识等，包括但不限于单选题、多选题、判断题。

## （二）实操技能竞赛

竞赛方式：以技能实操方式进行，双人协作完成，计分采用百分制，竞赛时间为 120 分钟，竞赛成绩占总成绩的 80%，60 分及以上为及格。竞赛时间到后，选手停止操作并有序离场，由裁判评分，并由工作人员进行设备复位。

竞赛内容：实操技能竞赛包含 3 个模块，涉及数据采集和处理、数据标注、智能系统运维、业务分析、智能训练、智能系统设计等技能操作知识。须利用提供的智能计算实训平台，分别对文本类数据、图像类数据、语音类数据进行处理与应用，完成典型场景应用系统的相关功能实现。

模块编号	模块名称	分值
模块 A	文本类数据处理及应用	30 分
模块 B	图像类数据处理及应用	50 分
模块 C	语音识别技术应用	20 分
总计	/	100 分

### ■ 模块 A：文本类数据处理及应用

基于典型的应用场景考核文本类数据的处理流程及应用，完成数据采集、数据清洗、数据可视化、模型构建及训练、模型评估及预测等内容。其中，数据采集包括通过数据库读取、数据靶站爬取等多种方式采集文本数据；数据清洗包括数据完整性检查，并进行缺失值、错误值、重复值的处理以及数据格式的规范统一；数据可视化主要对选取的数据中的特征变量进行聚合处理，并使用图形化方式展示（扇

形图、热度图、柱形图)；模型构建与训练主要包括数据集切割、构建算法模型及训练、模型评估及预测等。

#### ■ 模块 B：图像类数据处理及应用

基于典型的应用场景考核图像类数据的处理流程及应用，通过深度学习技术，进行算法模型的微调训练，完成数据采集与处理模型选择与训练、模型评估与预测等内容，实现对目标物体的识别与检测功能。其中，数据采集与预处理考核内容为数据采集、数据预处理、数据标注及数据集划分；模型选择与训练考核预训练模型的选择与导入、文件及参数配置、模型训练；模型评估与模型预测主要考核能够对模型的性能进行评估与优化，并调用已训练好的模型对目标物体进行实时检测。

#### ■ 模块 C：语音识别技术应用

基于典型的应用场景考核语音识别技术的应用，通过语音识别技术，进行语音识别、语音合成、实时语音转文字、语音播报，完成对已识别物体对象的实时播报等内容。能够基于离线语音识别引擎或在线语音识别引擎灵活的搭建人工智能综合项目。

### （三）命题方式

本赛项的竞赛试题不提前公开，试题由竞赛组委会组织相关专家按照赛项规则统一命制。为使参赛选手能更好的训练及了解本赛项的命题方式和方向，组委会将统一公布理论知识竞赛与实操技能竞赛的样题。

## 五、评分规则

参赛选手的成绩评定由竞赛组委会的裁判组负责。

### （一）理论知识竞赛成绩评定

根据评分标准统一阅卷、评分与计分，成绩汇总后经 3 名裁判共同签名确认后提交裁判长确认。

## （二）实操技能竞赛成绩评定

本项目评分标准为评价分（主观）、测量分（客观）。凡可采用客观数据表述的评判称为测量，凡需要采用主观描述进行的评判称为评价。按各模块评分表分别设置评分小组，由裁判长指定各组裁判人员，分别对各模块进行评分。各评分小组负责所有选手同一指标的现场评分，并签字确认评分结果。

### 1、评价分（主观）

评价分(Judgement) 打分方式：3 名裁判为一组，各自单独评分，根据选手完成比赛的成果，按权重表的标准给出权重分。裁判相互间分差必须小于等于 1 分，否则需要给出确切理由并在小组长或裁判长的监督下进行重新打分。

权重表如下：

权重分值	要求描述
0 分	未达到要求，包括“未做尝试”
0.5 分	基本达到要求
1 分	较好的达到要求

测量分评分准则样例表（模型训练过程中各项指标的评估）：

权重分值	要求描述
0 分	不接受（未进行模型训练，无评估指标输出）
0.5 分	基本达到要求（已进行模型训练，各项指标未完全收敛）
1 分	较好的达到要求（已进行模型训练，各项指标收敛度较好）

### 2、测量分（客观）

测量分（Measurement）打分方式：每组由 3 名裁判构成。每个组所有裁判一起商议，在对该选手在该项中的实际得分达成一致后最终只给出一个分值。测量分以测试功能或查看配置情况来评分，测量分只有两种结果，即得分和不得分。

测量分评分准则样例表：

子项	示例	正确分值	得分值
A1	检查数据集完成性情况（检查各列的缺失值），对缺失值进行补全操作，以保证数据的完整性。	2	2
B2	配置训练图片及测试图片的引用路径、待训练目标的类别数量、类别名称。	2	2
B3	配置模型训练的相关超参数。	2.5	2.5

### （三）评分流程说明

1、裁判员以小组的形式进行评判工作，裁判员小组的分组和分工由裁判长执行。

2、在评判过程中，所有的评判结论必须由评判小组集体决定。

3、评判工作分为客观测量评分和主观评价评分两个部分。测量评分：针对比赛结果如选手的代码文件、数据截图按《评分表》细则进行测量评价。主观评价评分：针对选手比赛作品的主观判断进行评价，同时对一处指标进行 0-3 等级归类评分，分数由裁判根据规定计算得出并记录到选手《评分表》。

4、关于成绩并列，当比赛现场出现选手总成绩并列时，裁判组首先将按照模块评分优先级不同的方式决定选手总成绩排名，评分优先级由大到小排序：模块 A > 模块 B > 模块 C，评分优先级比较仍不

能区分选手总成绩排名时，由评分裁判对该组排名相同选手的比赛模块所有主观评分项（评价）进行综合评价投票，投票领先的选手总成绩排名在前。

5、竞赛结束后，各裁判填写各参赛选手成绩汇总表并签名确认后提交裁判长确认。

## 6、技能操作部分

内容描述：在规定时间内，按任务书要求实现竞赛内容，竞赛结束，停止一切操作。

评判方式：实操技能竞赛过程中由裁判对选手的比赛操作时间进行计时并记录。裁判组在竞赛规定的结束时间后，分组对参赛队伍进行考评，裁判员按照评分表中要求，对提交结果进行打分评判。

## 六、项目特别规定

### （一）违规行为

选手有下列情形将在比赛过程中视作违规行为，竞赛裁判团队将对违规人员予以扣分或取消比赛资格处理：

1、选手在比赛进行期间未经允许使用非赛事配备的存储设备或通讯设备的，选手在比赛进行期间与其他参赛队或本项目赛场外的人员建立通信联系，选手比赛进行期间未经允许与本项目赛场外的人员建立通信联系；

2、选手使用自行携带未经裁判批准的工具或设备的；

3、赛场作弊、代赛等舞弊现象的；

4、选手在比赛过程中未经许可脱离本选手工位；

5、因操作不当导致事故或者损坏竞赛设备。

## （二）赛场纪律

1、选手着装要求大方得体，不穿拖鞋、短裤，所穿服饰不体现参赛选手的身份；

2、在比赛开始的前一天，选手有权熟悉赛场环境和设备，本日和比赛日均禁止带任何工具、设备入场；

3、在比赛开始前选手在工位内对赛事提供的设备、工具、材料的数量进行确认。设备、工具、材料数量确认后，在裁判宣布开始前禁止触碰竞赛设备或开启电源，否则按违规进行扣分处理；

4、比赛期间选手禁止携带存储及通信设备，如带到赛场，需要交给本单位场外人员保管或由赛场工作人员集中保管；

5、选手上交的电子文档由工作人员用赛场指定 U 盘进行拷贝传递或指定网络上传，设计成果如有需要由工作人员打印并由选手确认签字；

6、各参赛单位场外人员在比赛过程中严禁与任何选手交谈或做出任何提示、影响、干扰行为，如被发现将相应扣除当事人所在参赛队的成绩；

7、比赛期间，选手需要通过提示牌与现场裁判进行应答或举手交流，本单位裁判需要回避，由其他单位裁判员前去处理；

8、场内现场裁判执裁过程中，除选手示意并经裁判长同意，禁止主动进入选手工位内，禁止接近本单位选手；

9、选手如怀疑设备问题，且有明确证据确认损坏由非选手因素造

成，可向裁判提交书面说明，经技术人员判断和裁判长裁决认可，可更换设备，并由裁判长裁决是否补时和补时长度，没有明确证据确认损坏由非选手因素造成设备损坏的，不予更换设备和补时；

10、比赛现场发布的试卷禁止带出场外，竞赛结束后由现场裁判统一收回存档；

## 七、竞赛相关设施设备

### （一）实操技能竞赛设备

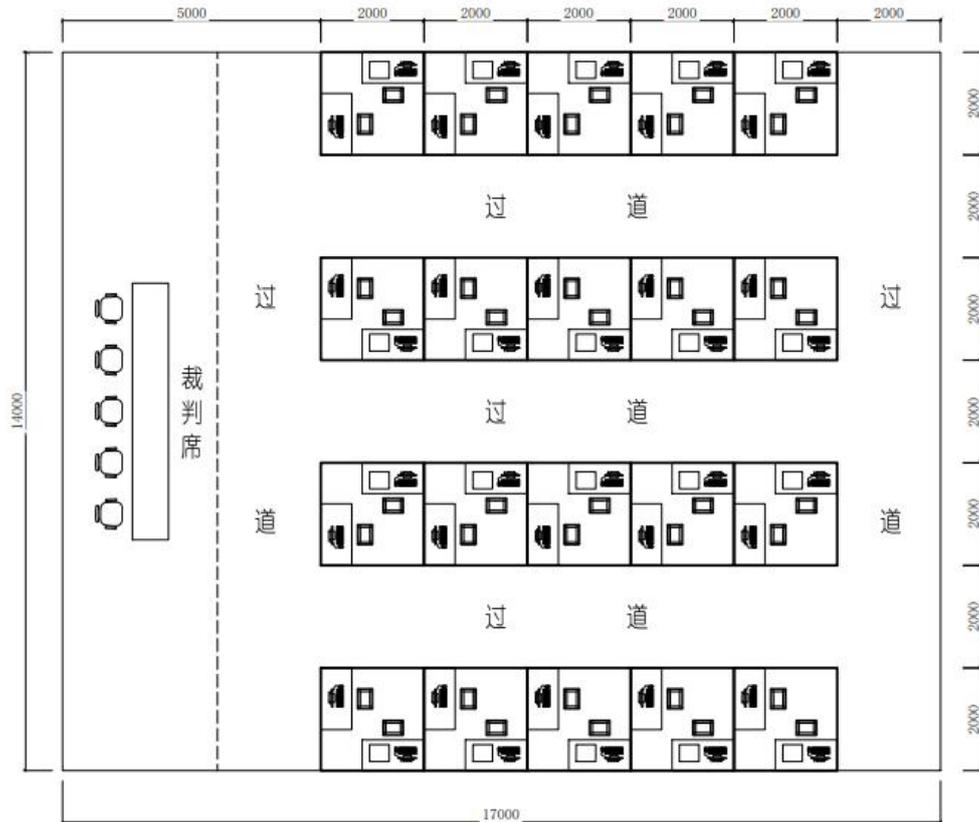
本赛项的竞赛平台为智能计算实训平台，该平台基于软硬一体化设计理念，无需外部算力支撑，可满足多用户同时登录使用。平台基于 KVM 虚拟化技术以及 GPU 虚拟化技术，具备统一管理、AI 开发引导、用户开发数据统计、采集靶标站、Web VNC、Web SSH 等功能。集成涵盖数据采集、数据预处理、数据标注、模型选择、模型训练与调参、模型推理测试、模型部署等全流程的技术开发平台。依托此平台，能够适配“人工智能训练师”国家职业标准，支持基于国家职业标准任务要求的实训体系，支撑“人工智能训练师”省级职业技能竞赛及职业技能等级认定。



智能计算实训平台

## (二) 竞赛场地平面图

实操技能竞赛的赛场设置有 20 个竞赛工位，每个工位互不干扰。同时在外围办公区域设置有检录区、休息区、裁判区和宣告区。



竞赛场地平面图（仅供参考，以实际竞赛场地布局为准）

## (三) 场地设备工具

场地竞赛设备均由组委会提供，单工位竞赛设备清单如下表所示，如有疑问请与技术支持单位联系。

序号	设备名称	规格描述	单位	数量
1	智能计算实训平台	支持文本类、图像类、语音类数据处理及应用实践，涵盖数据采集、数据清洗、数据标注、智能训练、模型推理测试、模型部署等全流程的技术开发平台。	套	1
2	竞赛电脑	I7-12700 代，16G 内存，256G 固态+1T 机械硬盘，23 寸显示器。	台	1

3	工位桌椅	工位：钢木及结构，参考尺寸（长宽高）： 600mm*1200mm*750mm； 座椅：弓形靠背椅 2 把	套	1
4	U 盘	16G，USB3.0 接口	个	1

## 八、健康和安

1、赛场必须留有安全通道，必须配备灭火设备。赛场应具备良好的通风、照明和操作空间的条件。做好竞赛安全、健康和公共卫生及突发事件预防与应急处理等工作。

2、赛项执委会须在赛前对本赛项全体裁判员、工作人员进行安全培训，并在赛前对选手进行培训，避免发生人身伤害事故，建立完善的安全事故防范制度。

3、参赛专家、裁判、工作人员及指导教练、选手入住承办单位统一安排宾馆、注意饮食卫生、乘坐承办单位统一安排的大巴接送赛场及宾馆之间的往返。

4、在赛项承办单位内提供工作人员咨询服务、赛场布局图、消防设施分布情况等，张贴安全提示和赛场标识、路线标识，确定设置安保人员地点和当日现场所需的安保服务人员数量。

5、参赛选手公平竞赛，杜绝舞弊，遵守赛场纪律；遵守设备操作规程，安全、文明参赛；着装规范整洁，爱护设备，保持竞赛环境清洁有序。

6、承办单位配备有医务服务、餐饮等后勤保障服务。

7、所有人员应服从组委会管理及工作人员的指挥、调动，按照比赛秩序表提供的安排准时入场，准时参赛、准时离场。

8、严格控制与参赛无关的易燃易爆以及各类危险品进入比赛场地，不许随便携带书包进入赛场。

9、如遇特殊或紧急情况，按照疏散方向标识，指挥赛场人员安全有序撤离。

## 九、绿色环保

1、竞赛任何工作都不应该破坏赛场周边环境。

2、环境整洁卫生，体现绿色环保。严格遵守竞赛规则，具有安全意识和卫生健康意识，平时养成良好的工作习惯，遵守职业规范，严禁明火，禁止吸烟。

3、所有竞赛相关人员必须注意保持场地整洁。交通路线、走廊、楼梯尤其是紧急疏散通道、灭火器及其他救生设备必须保持周边无障碍，且不得移除。必须立即清理地板上的电缆、杂物、废弃物等可能造成绊倒的类似物体，不再使用的材料，必须马上整理打包。每天比赛结束后，选手要做好个人赛位的卫生，工作人员要保障赛场整体的环境卫生，体现安全、整洁、有序。

4、赛场设置排烟除尘系统，尽可能地减少和控制烟尘。提倡绿色环保的理念，所有可循环利用的材料都应分类处理和收集。

## 十、样题

本赛项的样题包括理论知识竞赛样题及实操技能竞赛样题，竞赛组委会将另行发布。