



2024

高能物理AI平台

张易于, 杜然
高能物理研究所
计算中心



■ 人工智能三要素

■ 高能物理AI平台

- 模型服务、数据、培训

■ 高能物理AI算力平台

- 平台架构
- 平台建设状态与使用方法
 - 算力层、软件生态层、调度管理层、用户服务层
- 平台管理规则、算力分配与计费方案
- 平台上线时间计划

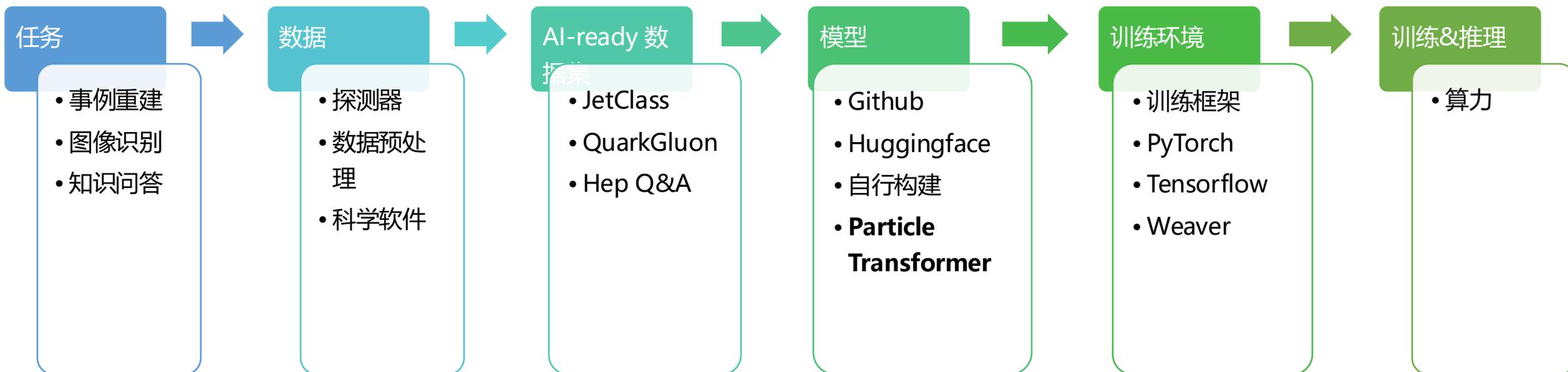
人工智能三要素



- 人工智能三要素：数据、算法、算力



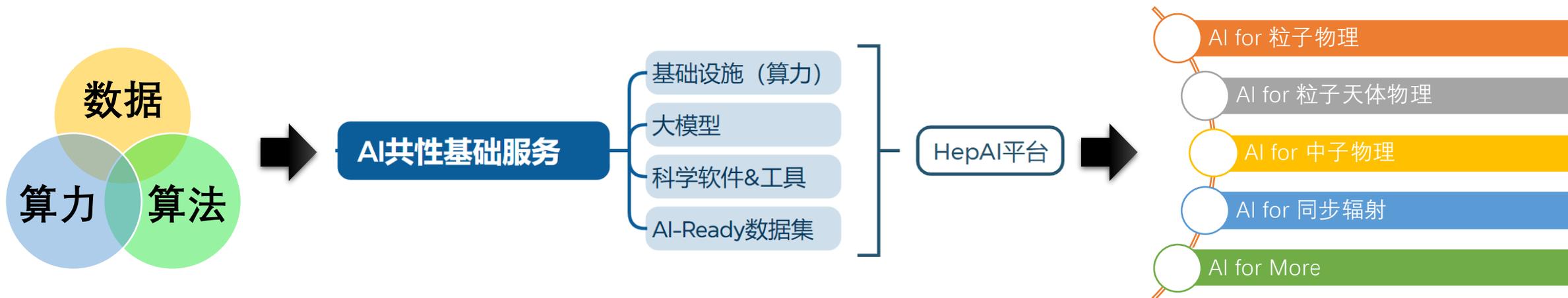
- AI研究流程：复杂、零碎，需要共性基础服务简化流程



面向科学发现的人工智能平台



- 高能物理AI平台：围绕 **数据-算法-算力**
- AI共性基础服务（面向领域科研活动）：算力基础设施、大模型、工具、数据集



高能物理人工智能平台HepAI

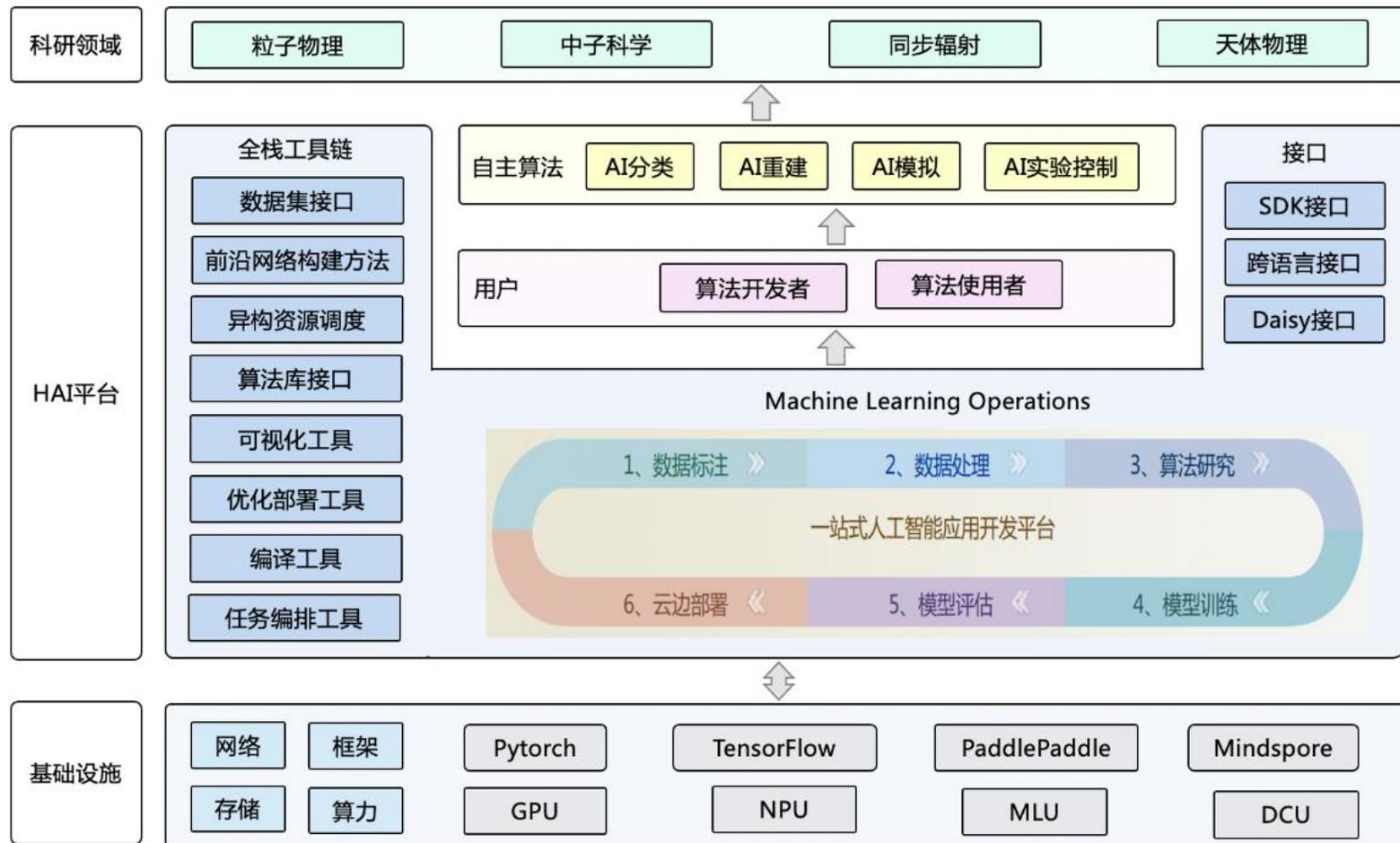


- 人工智能平台能加速多学科场景下的科学研究，简化模型迭代和流动，是发展AI算法和应用的**共性基础设施**。
<https://ai.ihep.ac.cn>

- 人工智能平台本身是**软件系统**，**承载AI算法模型**，**接入AI算力**，**打通数据通道**。

- 人工智能平台不仅仅是一个单纯的算力平台

- 算法：集成支持、自动化、优化、更新、维护
- 数据：管理、处理、安全
- 算力：资源优化
- 培训、共享与合作



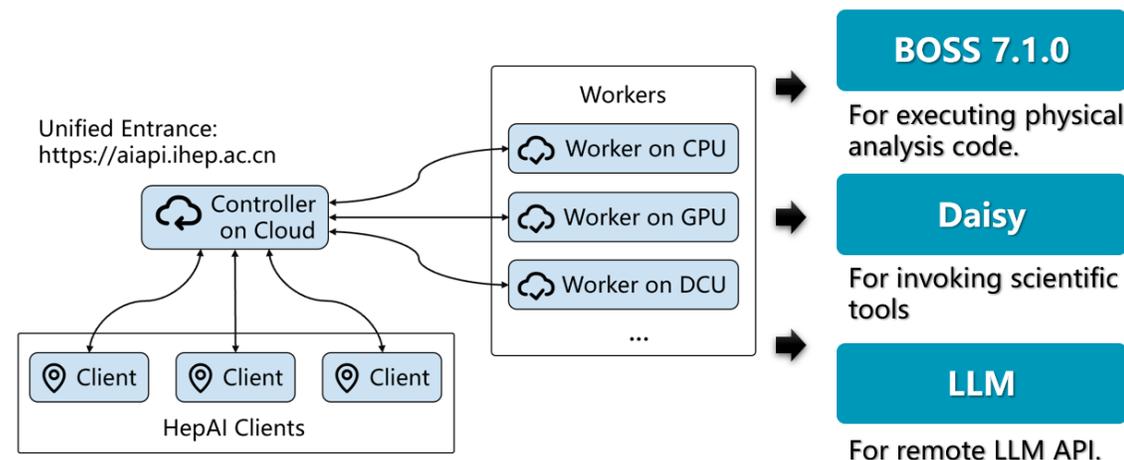
HepAI平台架构

AI算法模型、科学软件&工具



- 通过自研分布式部署框架部署先进AI算法模型和科学软件工具，实现直接调用。

- 大语言模型：ChatGPT, GPT-4, Xiwu, etc
- 科学软件：BOSS, Daisy
- AI算法模型：SAM(分割一切), ParticleNet&ParT(粒子鉴别), PtychoNN(叠层成像重建)
- 科研工具：PDF解析, PDF翻译, 论文润色, 科研绘图, etc



具体应用

- 天文警报信息汇集平台
- 增材制造缺陷分析标注工具
- GPT学术优化app
- HelpDesk运维助手
- "赛博士"科研智能体

- 通过HepAI-MaaS(Model as a Service)服务加速AI模型的应用创新

- 旧模式**：基于AI模型进行应用开发时，需要下载代码、模型、考虑算力、解决环境依赖等问题。
- 新模式**：基于HepAI-MaaS服务，使用API-KEY直接调用AI模型，而无需考虑算法部署、算力资源等问题。

大模型服务 辅助科研活动



- 基于大语言模型的生成式服务 HaiChat → Dr. Sai
 - 2023年8月正式上线，高能所注册用户2000+，30天日均活跃用户突破900人。
 - 科研模式变革中：工作遇到问题查资料 → 先问大模型
- 代码生成
- HaiAcademic

```
https://chat.ihep.ac.cn/v3

# headers={"Authorization": f"Bearer (hepa_api_key)",
# json=data,
# stream=True,
# )可以访问, 但aiohttp.ClientSession r = await session.request 失败, 为什么

OI openal/gpt-4o-mini
如果您能够使用 "requests.post" 成功访问 API, 但在使用 "aiohttp.ClientSession" 时遇到问题, 可能是因为以下一些原因。我们可以逐一检查这些可能性, 以解决您面临的问题。

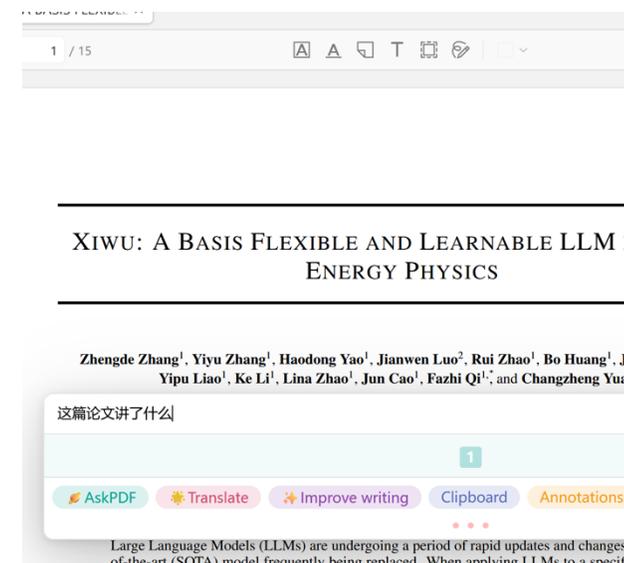
1. 请求头问题
确保使用 "aiohttp" 时您正确设置了所有请求头, 包括 "Authorization"。例如:

python
import aiohttp

OI openal/gpt-4o
当您使用 "requests.post" 可以成功访问, 而使用 "aiohttp.ClientSession" 的 "await session.request" 失败时, 可能的原因有以下几种:

1. Session 配置:
- 确保 "aiohttp.ClientSession" 的配置和 "requests.post" 相符, 特别注意请求头、认证信息和其他参数。
2. Event Loop:
- 确保您的代码在适当的事件循环中运行。"aiohttp" 基于异步 I/O, 确保
```

```
css # index.css > root
1 root {
2   --nav-background-color: #111;
3   --background-color: #ed5565;
4   --bars-color: #ffcc99;
5   --bars-color-secondary: #034f84;
6   --nav-font-color: #fff;
7   --button-color: #fff;
8   --button-color-secondary: #000;
9   --input-range-color: #ed5565;
10 }
11
12 body {
13   margin: 0;
14   padding: 0;
15   background: var(--background-color);
16 }
17
18 nav {
19   width: 100%;
20   background: var(--nav-background-color);
21   display: flex;
22   flex-wrap: wrap;
23   justify-content: space-evenly;
24 }
```



- 数据收集, 改进模型, 建立生态, 保持优势

辅助文献阅读



- AI-Ready的数据集

- 直接用于ML/DL的训练和测试，可**加快科学研究**进程；确保数据被广泛使用，**避免数据孤岛**现象；统一数据格式，有助于**促进跨学科**协作。

- 现状

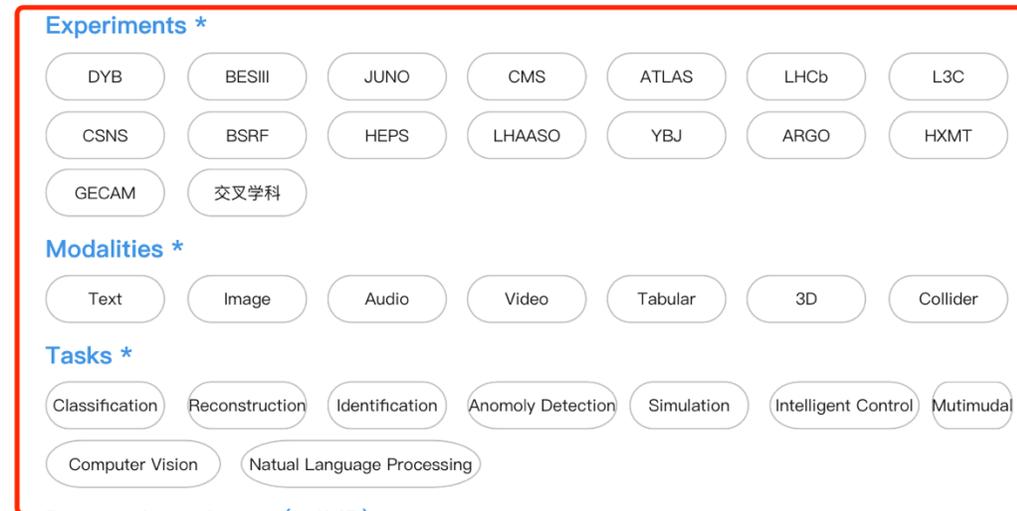
- 大部分：统一的格式，统一的存储，但不是AI-Ready的
- 少量：增材制造缺陷数据集，BESIII Tracking数据集，etc

- 数据来源

- 公开数据集
- 领域共享数据集

- 共享服务

- 依托AI平台提供数据集共享、发现服务
- 标记数据集知识产权与成果影响力



培训、共享与合作



- 代码逐行解读

- 代码快速复现

- 代码托管

- Gitlab

- 讲座、报告

- ML研讨会

- 公开课程

```
# Multi-Head Attention Module
This computes scaled multi-headed attention for given query, key and value vectors.
Attention(Q, K, V) = softmax_{seq} (frac{QK^T}{sqrt{d_k}}) V
In simple terms, it finds keys that matches the query, and gets the values of those keys.
It uses dot-product of query and key as the indicator of how matching they are. Before taking the softmax the dot-products are scaled by 1/sqrt{d_k}. This is done to avoid large dot-product values causing softmax to give very small gradients when d_k is large.
Softmax is calculated along the axis of the sequence (or time).
• heads is the number of heads.
• d_model is the number of features in the query, key and value vectors.
90 def __init__(self, heads: int, d_model: int, dropout_prob: float = 0.1, bias: bool = True):
96 super().__init__()
99 self.d_k = d_model // heads
101 self.heads = heads
104 self.query = PrepareForMultiHeadAttention(d_model, heads, self.d_k, bias=bias)
105 self.key = PrepareForMultiHeadAttention(d_model, heads, self.d_k, bias=bias)
106 self.value = PrepareForMultiHeadAttention(d_model, heads, self.d_k, bias=True)
```

Run the training pipeline

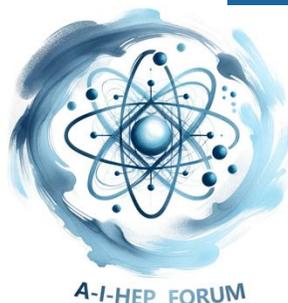
Next, you run the DAG to start the training job by invoking the metl parameters:

- dataset: The dataset resource to train the model.
- model_display_name: The human readable name for the trained model.
- training_fraction_split: The percentage of the dataset to use for training.
- test_fraction_split: The percentage of the dataset to use for testing.
- validation_fraction_split: The percentage of the dataset to use for validation.
- budget_milli_node_hours: (optional) Maximum training time in millihours (1000 = hour).
- disable_early_stopping: If True, the entire budget is used. Else, the training job is stopped if the model does not improve on the model objective measurements.

The run method when completed returns the model resource.

The execution of the training pipeline will take upto 60 minutes.

```
[ ] model = dag.run(
    dataset=dataset,
    model_display_name="salads",
    training_fraction_split=0.8,
    validation_fraction_split=0.1,
    test_fraction_split=0.1,
    budget_milli_node_hours=20000,
    disable_early_stopping=False,
)
```



IHEP ML workshop and ML group kick-off

基于AIGC的核脉冲检测与生成技术研究

C305, Main building

纳米相干多模态成像方法的AI研究

C305, Main building

机器学习在中子散射数据分析中的应用

C305, Main building

中子散射实验的蒙特卡洛模拟和实验参数的贝叶斯优化



2024年完整的从零开始人工智能



2024年人工智能入门天花板教程

异构算力基础设施



- 目前高能所算力硬件共250+英伟达GPU，以**V100**为主，少量**A100**（非AI研究独占）。

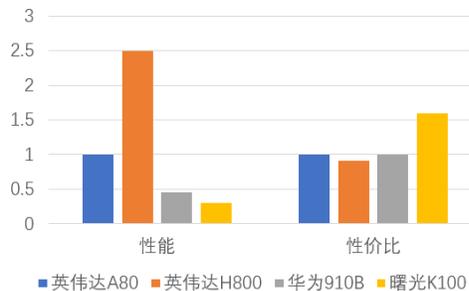
AI研究趋势

- 研究项目增加、模型大小变大
- 算力需求、时长增加
- 算力价格昂贵
- 国产化替代必要性增强

算力类型



A800, H800, NPU和DCU对比



- 模型、数据、培训需要算力的支撑

任务	算法	参数量	训练时间	算力	数据	科学软件		
硬件设计与优化	DNN	100	sec-min	CPU、单卡		Gen4 Root Python C++ Matlab BOSS ...		
	DNN, CNN GNN, VAE ParticleNet	10k	Min-hour	中端单卡	100-10k			
在线处理	事例触发 在线重建 数据压缩							
	事例重建, 鉴别, 预测	ResNet Transformer Diffusion	10M	Hour-day	高端GPU 多GPU		10k-10M	
数据分析	理论、方程求解	Transformer SAM Clip-parT	1B	Day-week	多GPU集群		10M-10B	
	科学大模型 中子实验数据重整 同步辐射结构化大模型 射线成像图像通用大模型		13B	3800 tokens/s	NV A800 1*8卡		1B tokens	
			Llama2 Owen Baichuan	13B	1400 tokens/s		DCU K100 1*8卡	1B tokens
			70B	2400 tokens/s, 14 days	DCU K100 8*8卡		1B tokens	
		70B	7700 tokens/s, 4.5 days	DCU E100 8*8卡	1B tokens			
运维与服务	卫星星座智能化控制 智能调光、方案实时优化 物理实验智能系统运维	DNN, CNN DQN	10k-1B		单卡 多GPU		5000-1M	实验控制 接入
	智能值班员 科研智能体 科研助手 智能实验	agent	-	-	推理算力	-		
系统集成与应用								



■ 人工智能三要素

■ 高能物理AI平台

- 模型服务、数据、培训

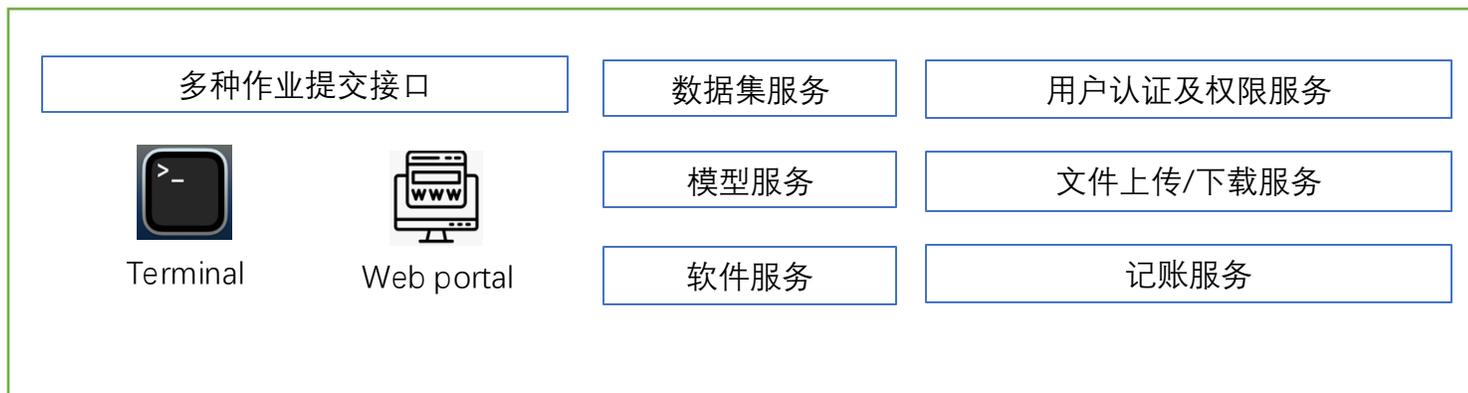
■ 高能物理AI算力平台

- 平台架构
- 平台建设状态与使用方法
 - 算力层、软件生态层、调度管理层、用户服务层
- 平台管理规则、算力分配与计费方案
- 平台上线时间计划

平台架构 – 院网络中心合作，分四层建设



用户服务层



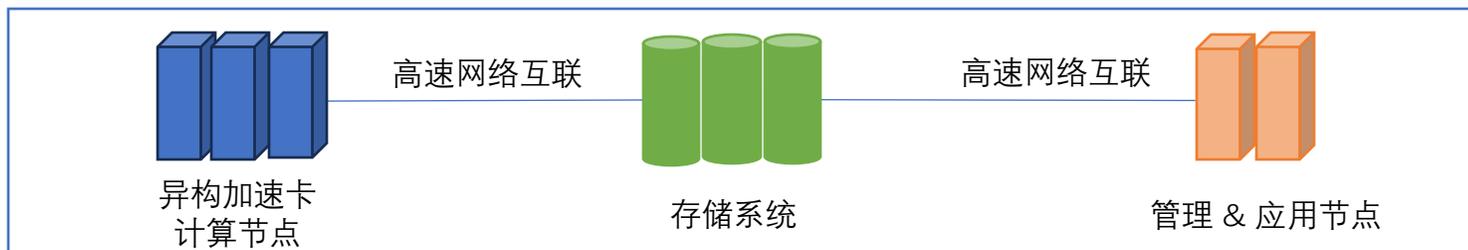
调度管理层



软件生态层



算力层



监控与报警服务

算力层 - 设备



用途	设备名称	数量 (台/套)	配置
计算	A800 GPU服务器	1台	<ul style="list-style-type: none"> 8 * A800 80GB PCI-e NVIDIA GPU卡 2 * Intel(R) Xeon(R) Gold 6430(32 core) 1TB 内存, 7.68 TB NVME 本地硬盘 2 * 200Gbps IB网卡, 1 * 25Gbps 以太网卡
	L40 GPU服务器	1台	<ul style="list-style-type: none"> 8 * L40 48GB PCI-e NVIDIA GPU卡 2 * Intel(R) Xeon(R) Gold 6430(32 core) 1TB 内存, 7.68 TB NVME 本地硬盘 2 * 200Gbps IB网卡, 1 * 25Gbps 以太网卡
	DCU服务器	4台, 待增加2台	<ul style="list-style-type: none"> 8 * K100AI 64GB PCI-e 国产海光DCU卡 2 * Intel(R) Xeon(R) Gold 6430(32 core) 1TB 内存, 7.68 TB NVME 本地硬盘 2 * 200Gbps IB网卡, 1 * 25Gbps 以太网卡
存储	存储服务器及阵列	1套	<ul style="list-style-type: none"> 全闪, 可用容量 ~ 200TB, 挂载点/aifs
网络	IB交换机	1台	<ul style="list-style-type: none"> 200Gbps 带宽, 连接计算节点与存储系统
	以太网交换机	1台	<ul style="list-style-type: none"> 25Gbps 带宽, 连接计算节点与存储系统
管理和应用	CPU服务器	4台	<ul style="list-style-type: none"> 2 * Intel(R) Xeon(R) Gold 6430(32 core), 256GB 内存 1 * 25Gbps以太网卡, 480GB SATA SSD + 1TB NVME 本地硬盘

算力层 – 存储系统，可用容量200TB，全闪



• 个人目录

事项	说明
路径	/aifs/user/home/<username>
用途	存放个人datasets/models等文件
权限	用户个人可读/写
限额(quota)	默认值500GB、30万个文件，可按需修改

• 公共目录

事项	说明
路径	/aifs/public/data
用途	存放公开datasets/models等文件
权限	普通用户可读、不可写，公共数据管理员可读/写
限额(quota)	按需定制

软件生态层



- AI所需基础库文件、包管理软件、框架软件，基于AlmaLinux 9.4
- 可使用module工具加载软件环境
- 后续可按需增加

```
----- /cvmfs/slurm.ihep.ac.cn/alma9/modulefiles -----
anaconda/24.3.0      epics/7.0.7          intel_oneapi/compiler-rt32/2024.1.0  intel_oneapi/ifort32/2024.1.0      intel_oneapi/tbb32/2021.12
cmake/3.18.4         fftw/3.3.10-gcc11    intel_oneapi/compiler-rt32/latest    intel_oneapi/ifort32/latest        intel_oneapi/tbb32/latest
cmake/3.26.4         gcc/7.5.0            intel_oneapi/compiler/2024.1.0       intel_oneapi/intel_ipp_ia32/2021.11 intel_oneapi/vtune/2024.1
cmake/3.29.1         gcc/9.5.0            intel_oneapi/compiler/latest         intel_oneapi/intel_ipp_ia32/latest  intel_oneapi/vtune/latest
cp2k/2023.1-gcc11    gcc/10.4.0           intel_oneapi/compiler32/2024.1.0     intel_oneapi/intel_ipp_intel64/2021.11 lammps/2024.02.07
cp2k/2024.1-gcc11   gcc/11.4.0           intel_oneapi/compiler32/latest       intel_oneapi/intel_ipp_intel64/latest lapack/3.11.0-gcc11
cuda/11.0-cvmfs      gcc/12.3.0           intel_oneapi/dal/2024.0.0            intel_oneapi/intel_ippcp_ia32/2021.11 lume-astro/0.6.1
cuda/11.1-cvmfs      gimic/2.0            intel_oneapi/dal/latest              intel_oneapi/intel_ippcp_ia32/latest molpro/2015-gcc11
cuda/11.2-cvmfs      gromacs/2023.4-gcc11 intel_oneapi/debugger/2024.1.0       intel_oneapi/intel_ippcp_intel64/2021.11 mpi/mpich/4.1.3
cuda/11.7-cvmfs      hdf5/1.10.11        intel_oneapi/debugger/latest         intel_oneapi/intel_ippcp_intel64/latest mpi/mpich/4.2.1
cuda/12.2-cvmfs     intel_oneapi/advisor/2024.1 intel_oneapi/dev-utilities/2024.0.0  intel_oneapi/mkl/2024.1            mpi/mvapich/2.3.7
cuda/12.4           intel_oneapi/advisor/latest  intel_oneapi/dev-utilities/latest    intel_oneapi/mkl/latest           mpi/mvapich/3.4.3
cuda/12.4-cvmfs     intel_oneapi/ccl/2021.12.0  intel_oneapi/dnnl/3.4.0              intel_oneapi/mkl32/2024.1         openmpi/4.1.4-gcc11
cuDNN/8.0.5-cuda11.0 intel_oneapi/ccl/latest    intel_oneapi/dnnl/latest             intel_oneapi/mkl32/latest         orca/5.0.4-gcc11
cuDNN/8.1.1-cuda11 intel_oneapi/compiler-intel-llvm/2024.1.0 intel_oneapi/dpct/2024.1.0          intel_oneapi/mpl/2021.12         python/2.7.18
cuDNN/8.9.7-cuda11 intel_oneapi/compiler-intel-llvm/latest  intel_oneapi/dpct/latest            intel_oneapi/mpl/latest         python/3.7.16
cuDNN/8.9.7-cuda12 intel_oneapi/compiler-intel-llvm32/2024.1.0 intel_oneapi/dpl/2022.5             intel_oneapi/mpl/latest         python/3.8.19
cuDNN/9.0.0-cuda11 intel_oneapi/compiler-intel-llvm32/latest intel_oneapi/dpl/latest             intel_oneapi/oclfgpa/2024.1.0   python/3.9.18
cuDNN/9.0.0-cuda12 intel_oneapi/compiler-rt/2024.1.0       intel_oneapi/dpl/latest             intel_oneapi/oclfgpa/latest     python/3.10.14
elegant/1.0         intel_oneapi/compiler-rt/latest  intel_oneapi/ifort/2024.1.0        intel_oneapi/tbb/2021.12        python/3.11.8
                    intel_oneapi/compiler-rt/latest  intel_oneapi/ifort/latest          intel_oneapi/tbb/latest
```

调度管理层



- 基于Slurm 23.11.06, AlmaLinux 9.4
- 资源管理：按功能分区

PARTITION	AVAIL	TIMELIMIT	NODES	STATE	NODELIST
control	up	infinite	1	drain	aislurm01
login	up	infinite	1	drain	ailogin001
gpu	up	infinite	2	idle	aigpu[001-002]
dcu	up	infinite	4	idle	aidcu[001-004]

- 算力分区
- 根据加速卡类型划分
- dcu算力分区和gpu算力分区

- 作业调度：基于QOS的优先级调度，作业中指定QOS即可获得对应的调度服务

QOS名称

作业优先级

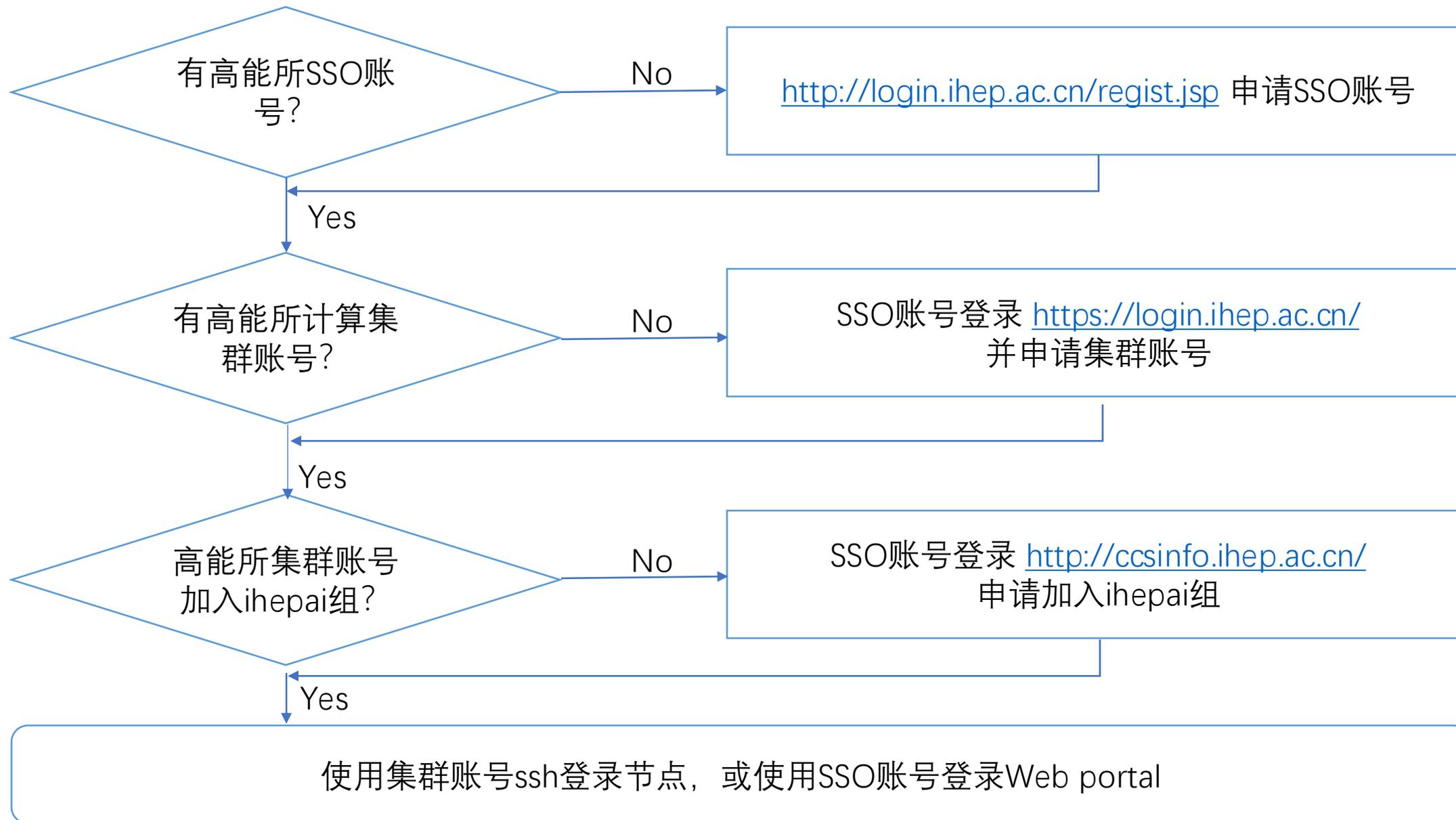
作业可使用的最大资源限制

作业最大运行时间

可提交的最大作业数量

Name	Priority	MaxTRESPU	MaxWall	MaxSubmitPU
gpunormal	0	cpu=32,gres/gpu:a800=8,gres/gpu:l40=8,gres/gpu=8,mem=640G	2-00:00:00	8
gpudebug	20	cpu=32,gres/gpu:a800=8,gres/gpu:l40=8,gres/gpu=8,mem=640G	00:15:00	16
dcunormal	0	cpu=48,gres/dcu:k100ai=12,mem=960G	2-00:00:00	16
dcudebug	20	cpu=64,gres/dcu:k100ai=16,mem=1280G	00:15:00	24

用户服务层 – 平台申请与用户认证



用户服务层 – 两种使用方式



- 终端登录：使用集群账号

```
└─$ ssh duran@ailogin001.ihep.ac.cn
duran@ailogin001.ihep.ac.cn's password:
#####
#                               Welcome to IHEP AI Platform                               #
# Any Question, Please contact http://helpdesk.ihep.ac.cn/ #
#####
Last login: Thu Oct 17 22:15:55 2024 from 10.100.0.157
[duran@ailogin001 ~]$
```

- Web portal登录：使用SSO账号

Welcome to the IHEP AI Platform

AI training and inference are supported with DCU and GPU cards

32	16
TOTAL DCU Cards	TOTAL GPU Cards
384	6TB
TOTAL CPU Cores	TOTAL Memory

Login

Add your credentials

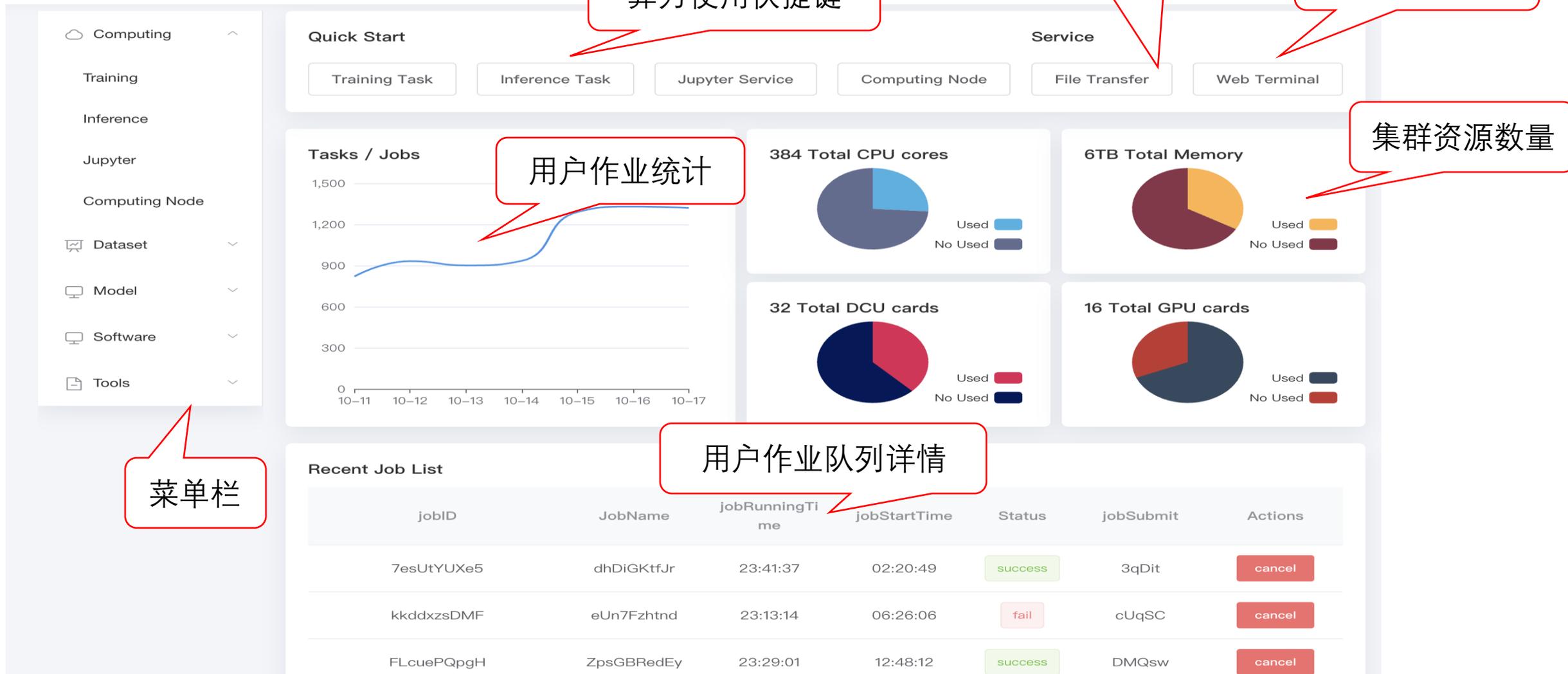
账号

密码

[登录](#) [忘记密码?](#)

没有账号, 点击[注册](#)

用户服务层 - Dashboard



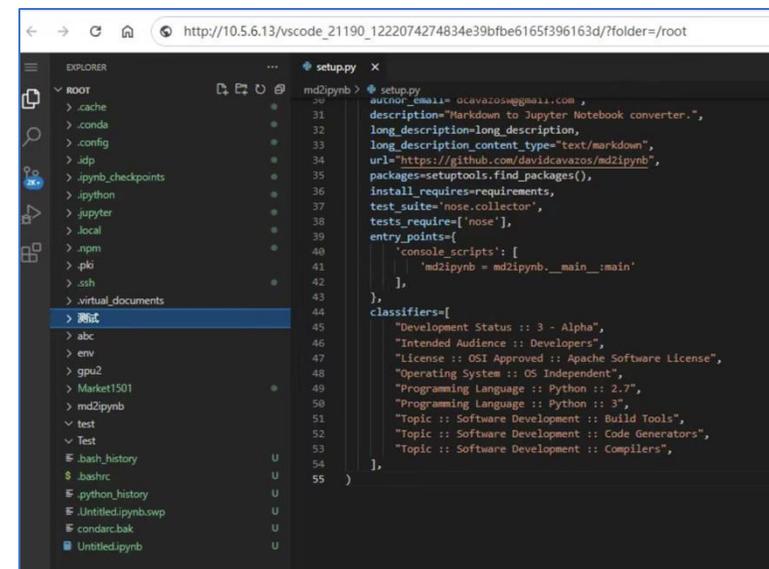
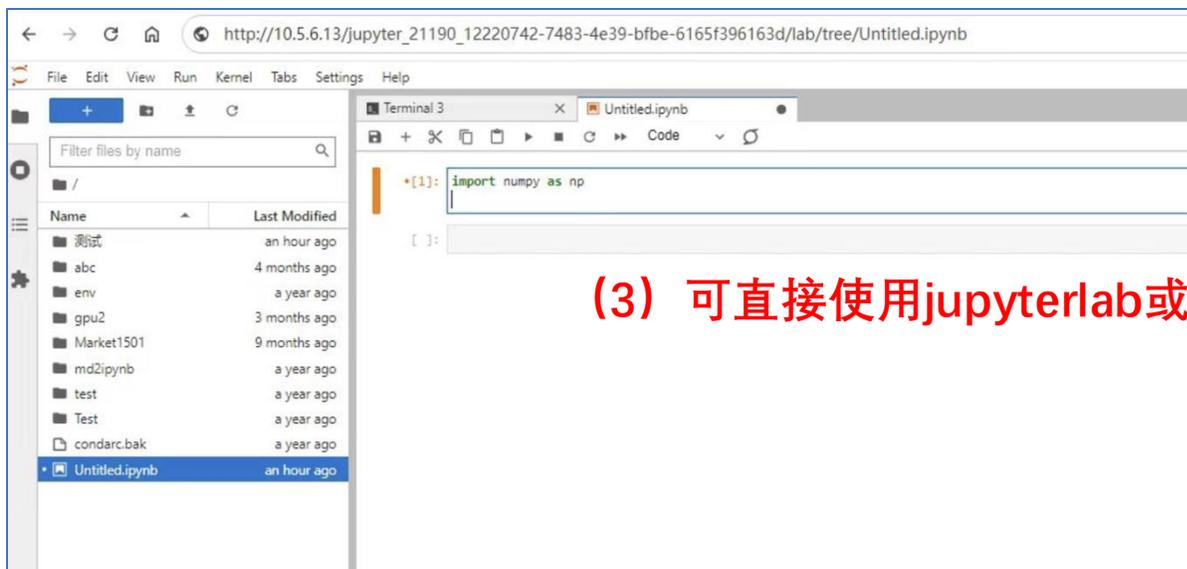
用户服务层 - 算力使用



(1) 填写算力使用相关信息

基于网页形式的、结合了编写说明文档、数学公式、交互计算和其他富媒体形式的工具。简言之，网页应用是可以实现各种功能的工具

(2) 等待资源分配和服务启动





- 查看公开 & 私有的数据集/模型/软件
 - 公开数据集: /aifs/public/data/datasets

📁 hep_text_v1.0	Oct 11, 2024, 8:16:01 PM	drwxr-xr-x
📁 JetClass	Oct 11, 2024, 8:14:26 PM	drwxr-xr-x
📁 OmniFold	Oct 11, 2024, 2:36:20 PM	drwxr-xr-x
📁 QuarkGloun	Oct 11, 2024, 7:26:12 PM	drwxr-xr-x
📁 TopLandscape	Oct 11, 2024, 7:26:45 PM	drwxr-xr-x

- 公开模型: /aifs/public/data/models

📁 Xiwu	Oct 11, 2024, 7:51:00 PM	drwxr-xr-x
--------	--------------------------	------------

用户服务层 - 文件上传/下载



The screenshot shows the 'Jobs Admin' interface. On the left is a navigation menu with items: 文件管理, 文件浏览, 文件上传, 资源监控, 日志管理, 机时管理, and 用户中心. The main content area has a header 'Drop files here to upload or' and two buttons: 'select files' and 'select folder'. Below the buttons, a file named 'public_cnrc.tar' is shown with a size of 2.9 GB and a progress indicator of 3% 7.2 MB / s 6 minutes. A red callout box points to the 'select files' and 'select folder' buttons with the text '点击上传文件/文件夹'.

用户服务层 – web terminal



- web terminal可直接ssh到登录节点，提交作业

```
Connecting...

[duran@ailogin001 ~]$
[duran@ailogin001 ~]$ pwd
/aifs/user/home/duran
[duran@ailogin001 ~]$ squeue
      JOBID PARTITION     NAME     USER  ST       TIME  NODES NODELIST(REASON)
       383      dcu   dcutest   duran   R        0:09      1 aidcu002
       384      dcu   dcutest   duran   R        0:09      1 aidcu002
       381      dcu   dcutest   duran   R        0:12      1 aidcu002
       382      dcu   dcutest   duran   R        0:12      1 aidcu002
       380      gpu   gputest   duran   R        0:32      1 aigpu001
[duran@ailogin001 ~]$
```



- **面向全所**所有AI计算需求的用户、课题组提供计算服务
- **有偿运行**
 - 偿还所里投资
 - 支付水电费
 - 保证平台可持续发展
- 所里资源可以满足需求的前提下，**不再支持**购买**外部算力**使用
- **鼓励**课题组/应用使用平台算力，支持购买硬件资源、贡献到平台中**共享使用**
 - 购买计算资源的**技术参数可咨询**计算中心
 - 可根据贡献度大小提高资源使用上限 & 作业优先级大小

算力分配方案 & 收费方案



• 算力分配方案

- 普通用户：提供初始作业优先级和资源使用限制，后续**可按需调整**
- 有硬件资源贡献的用户：
 - 节点数量**满足一定规模**的用户/组可以**按需保留**部分节点专用，其他节点需**共享使用**
 - 专用节点可**按需配置**作业优先级及资源使用限制
 - 有节点贡献到共享资源分区的用户/组，可根据贡献度大小**提高**作业优先级和资源使用限制

• 收费方案

- 平台**统一记账**，费用收取与科研计划处配合进行
- 计算资源（cpu/gpu/dcu）单价**不高于**商算平台
- 有共享硬件资源贡献的用户，根据贡献机时数量和实际消耗机时数量**核减**
- 水电费理论上由收取的计算费**覆盖**，保证平台的**可持续运行**

平台上线时间计划



时间节点	开放使用目标群体
当前	人工智能工作组测试调优
2024.10月底	面向全所使用
2024.11月底	面向高能物理领域使用



Thanks

Backup



- 人工智能平台本身是**软件系统**，承载**AI算法模型**，接入**AI算力**，打通**数据通道**。
- 承载AI算法模型
 - **生成式AI系统HaiChat**，2023.08已上线，高能所用户1500+，日活700，提升日常工作效率。
 - **赛博士科研智能体**，提升物理分析效率，计划24年上线。
 - 下一步：支持开源的Particle Transformer、Segment Anythong Model等SOTA模型，直接体验前沿技术；**部署所内自主开发的AI模型**，可直接调用展示成果。
- 提供AI算力
 - 下一步：提供部署AI模型的算力，用于开发和模型共享。
- 打通数据通道
 - 当前：AI开放数据平台（4个AI-Ready数据集）
 - 下一步：接入更多AI-Ready科学数据集共享、发现门户；与DOMAS等对接，接入光源/高能物理数据；