```
"st data": ST data
"conversations": [
  "from": "human",
  "value": "<st_data>\nGiven the meteorological raster data (ERA5) on U
  Date in China, predict the environmental event that will happen. The
  meteorological raster data: temperature, u&v wind, precipitation are
  encoded as R,G,B channels of an image. It is then encoded using a
  visual encoder. The output should be in the format: [event1, event2,
  event3, ...]. Please analyze the meteorological patterns in China and
  predict the environmental events that will happen. The context of the
  meteorological information: max temperature: Max T, min
  temperature: Min T, mean temperature: Mean T, max wind
  speed: Max W, min wind speed: Min W, mean wind speed:
  Mean W, max precipitation: Max P, min precipitation: Min P
  mean precipitation: Mean P."
  "from": "gpt",
"value": "[v<sub>1</sub>, v<sub>2</sub>, ...]"
```