

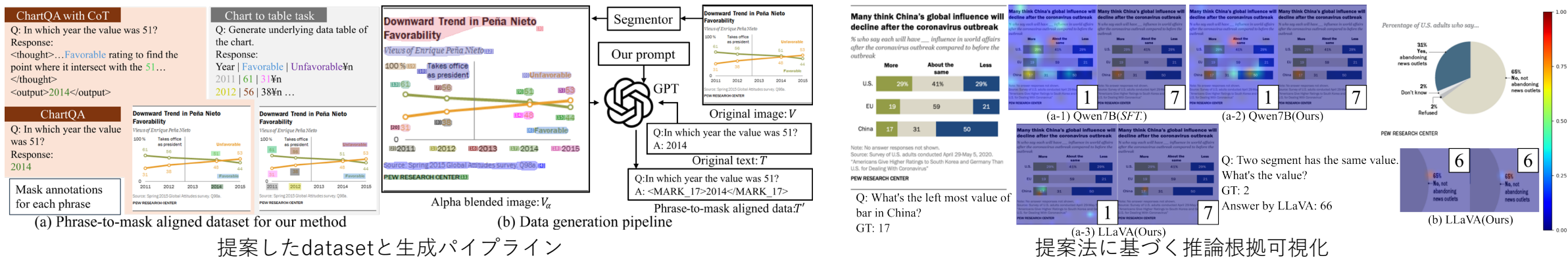
ABCI 3.0開発加速利用 (2025年度) 成果概要 (公開用)

課題名：
宇宙機システム開発におけるLLM技術適用のFeasibility Study

実施時期：2025年4月-2026年3月
所属機関名：宇宙航空研究開発機構
代表者氏名：伊藤光一郎

成果概要：
図認識のVQA (visual question answering) にむけたVLLMの再学習において、物体検出の要素を取り込む検証を行った。具体的には、図画像領域とテキスト内部のフレーズとの対応付けを課す学習を通常の指示学習に付随させることで、VLLMがトークン生成に際して参照すべき画像領域に制約を課した。実験では複数の公開VLLMに対してchartQAデータセットでの検証を行い性能向上を示したほか、推論根拠の可視化に寄与する点を示した。

- 成果のポイント：
- ・VQAにおけるトークン生成時、VLLMが画像内の特定の関連領域を参照していることを明らかにした。
 - ・学習において上記事実の制約を課す仕組みを導入するため、VLLM指示学習にphrase grounding (物体検出) のコンセプトを導入、学習に付随させた。
 - ・上記学習を実現させるため、画像領域とVQAテキストフレーズを対応付けたデータセットの作成方法を提案した。
 - ・提案手法を複数の公開VLLMに適用し、ChartQAベンチマークを用いた検証によってその有効性を実証した。
 - ・また、提案の根拠としたVLLMの特定領域参照についても評価し、その性能向上を示した。



成果についてより詳細な情報を提供しているWebページ、発表論文などの情報：
 「VLM出力に対する画像トークン寄与度の図認識学習利用への検討」, 伊藤光一郎, 金子泰之, 橋本真太郎, 石濱直樹, 人工知能学会2025
 「図認識に向けたVLM注意機構の接地利用の検討」, 伊藤光一郎, 金子泰之, 橋本真太郎, 石濱直樹, 画像の認識・理解シンポジウム2025
 "Phrase-Grounding-Aware Supervised Fine-Tuning for Chart Recognition via Side-Masked Attention", Koichiro Ito, To appear in CVPR2026