

# TẠO SINH HÌNH ẢNH TỪ VĂN BẢN VỚI QUY TRÌNH: GHI LẠI, LẬP KẾ HOẠCH VÀ TẠO SINH VỚI MLLMS

Lê Thị Như Ý<sup>1,2</sup>

<sup>1</sup> Trường Đại học Công nghệ thông tin,  
ĐHQG HCM

<sup>2</sup> 21522818@gm.uit.edu.vn

## What ?

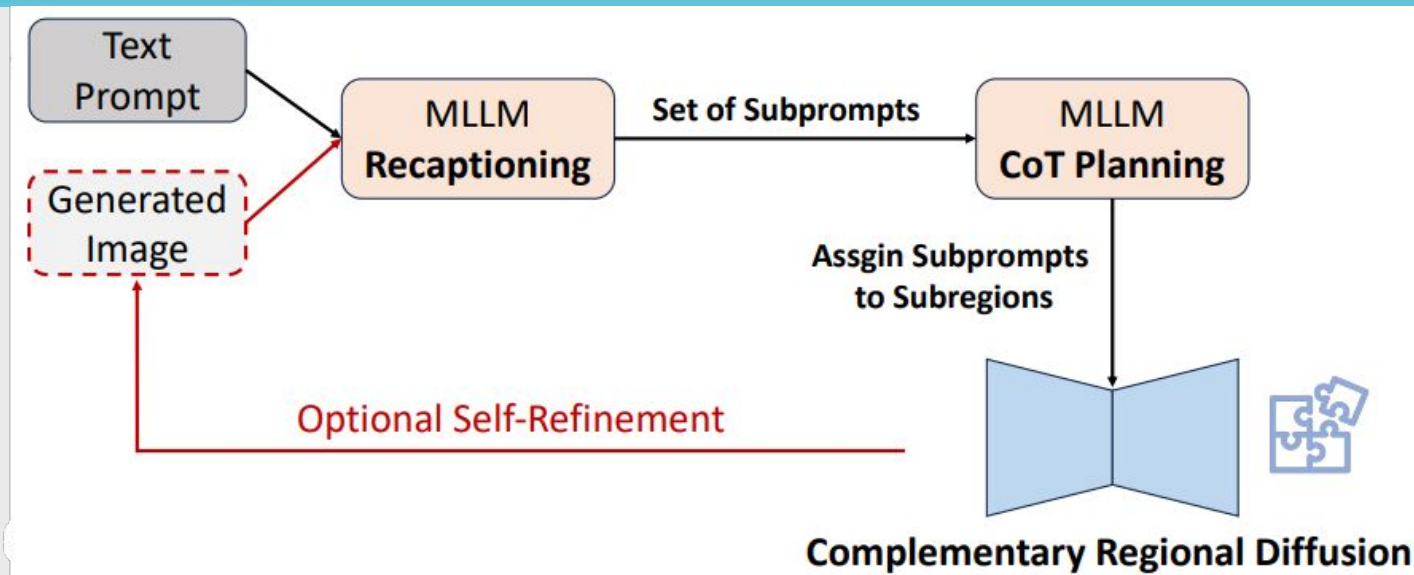
Chúng đề xuất một framework mới có vai trò tạo sinh hình ảnh từ văn bản:

- Sử dụng MLLMs phân rã bài toán thành các yêu cầu nhỏ. Hợp thành hình ảnh kết quả từ các khu vực được tạo sinh nhỏ hơn.
- Xây dựng giải pháp tạo sinh hình ảnh hiệu quả từ những yêu cầu văn bản phức tạp.

## Why ?

- Nhu cầu chuyển đổi thông tin văn bản thành hình ảnh luôn cần thiết trong xã hội phát triển ngày nay. Việc này giúp nhiều lĩnh vực thuận tiện và hoạt động hiệu quả hơn về mặt minh họa, sáng tạo.
- Hầu hết các mô hình hiện nay gặp khó khăn trong việc xử lý các yêu cầu văn bản phức tạp: **không hiểu hết ý nghĩa** hay **không có sự thống nhất tự nhiên**.

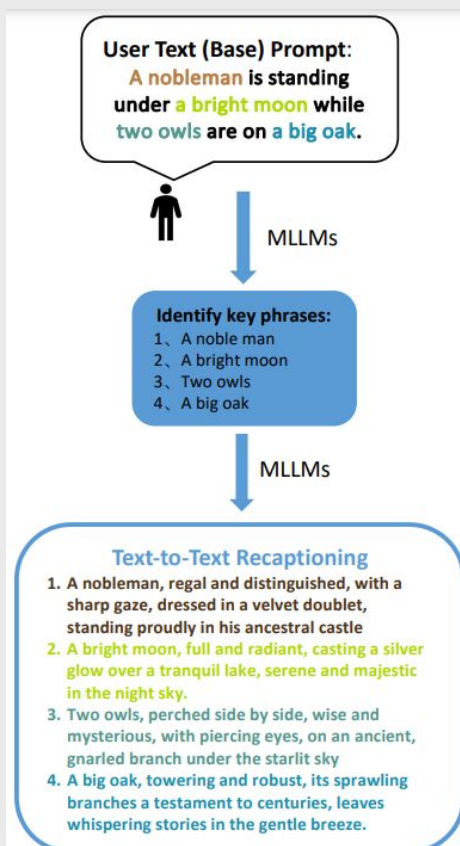
## Overview



## Description

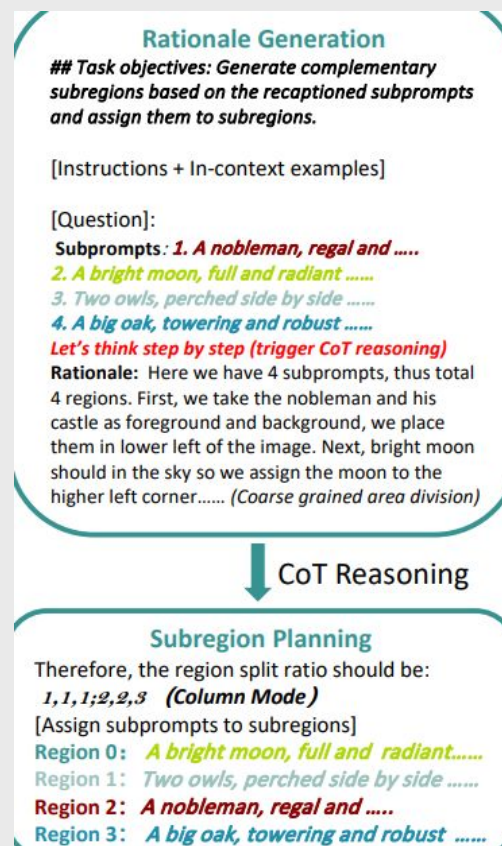
### 1. Ghi lại

Sử dụng Mô hình Ngôn ngữ Lớn đa dạng để phân rã câu văn thành các câu văn phụ khác nhau.



### 2. Lập kế hoạch

Phân chia không gian hình ảnh thành các khu vực con bổ sung và gán yêu cầu phụ khác nhau cho mỗi khu vực con.



### 3. Tạo sinh

Đề xuất lan truyền bổ sung khu vực để tăng cường tính linh hoạt và chính xác của quá trình tạo sinh văn bản thành hình ảnh đồng nhất.

