Template 1: Complex Bangla Privacy Rewrite Data Generation

Each example must follow this JSON structure:

{
  "sentence": "A natural, fluent Bangla sentence (can be multi-clause) containing a private or sensitive detail. Use richer sentence structures such as cause-effect, indirect mention, or background context.",
  "personinfo": "A short Bangla sentence stating only the core sensitive/personal detail in direct form.",
  "r_d": "Rewrite by fully removing the personal info, while keeping the sentence fluent and grammatically correct.",
  "r_o": "Rewrite by replacing the sensitive part with a generalized, context-aware Bangla expression (e.g., 'একজন বন্ধু', 'একটি প্রতিষ্ঠান'). Keep the rest of the sentence structure unchanged."
}

Example:

{
  "sentence": "আমি টিচার্স ট্রেনিং প্রোগ্রাম থেকে সরে এসেছিলাম যখন প্রশিক্ষক আমার ওটিজম নির্ণয় ক্লাসে প্রকাশ করেছিলেন।",

  "personinfo": "আমার ওটিজম নির্ণয়।",

  "r_d": "আমি শিক্ষক প্রশিক্ষণ প্রোগ্রাম থেকে সরে এসেছিলাম।",

  "r_o": "আমি শিক্ষক প্রশিক্ষণ প্রোগ্রাম থেকে সরে এসেছি যখন একজন প্রশিক্ষক আমার মানসিক অবস্থা প্রকাশ করেছিলেন।"

}

Explanation:

This template was designed to generate **richer and more varied Bangla sentences** that better reflect real-life discourse. By encouraging the use of multi-clause structures, background context, and indirect mentions, it produces more challenging data for privacy-preserving rewriting tasks. Such complexity helps models learn how to remove or obscure sensitive details without oversimplifying sentence structures, thereby improving generalization to natural text.

Template 2: Two-Sentence Bangla Privacy Rewrite Data Generation

Each example must follow this JSON structure:

{
  "sentence": "Two natural and connected Bangla sentences where one sentence contains a specific private detail (e.g., name, relation, job, health, education, behavior).",
  "personinfo": "A short Bangla sentence that directly reflects the private detail mentioned in the main sentence.",
  "r_d": "Rewrite by removing only the part that matches personinfo. Keep the sentence fluent and natural. Use correct pronouns (তিনি/তাকে/তার, or সে/তার/তাকে) depending on relation and age.",
  "r_o": "Rewrite by replacing only the sensitive part with a generalized, context-aware Bangla phrase (e.g., 'একজন আত্মীয়', 'একজন পরিচিত', 'একজন শিক্ষক'). Keep the rest unchanged or very close to original."
}

Example:

{

  "sentence": "আমার মা এখনো অসুস্থ। আজ তাকে নিয়ে হাসপাতালে গিয়েছিলাম।",

  "personinfo": "মাকে হাসপাতালে নিয়ে গিয়েছিলাম।",

  "r_d": "তিনি এখনো অসুস্থ।",

  "r_o": "আমার পরিবারের একজন সদস্য এখনো অসুস্থ। আজ তাকে চিকিৎসা করাতে নিয়ে গিয়েছিলাম।"

}

Explanation:

This template was designed for **simpler, two-sentence structures** that still capture sensitive details clearly. It ensures data diversity by including short, connected sentences rather than complex multi-clause constructions. This format is especially useful for evaluating how well models handle **direct and explicit personal information**, focusing on precise deletion or contextual obscuring while preserving fluency.