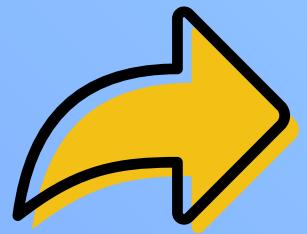
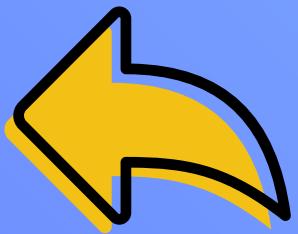


# كيف نُعلم الذكاء الاصطناعي

## فهم اللهجة الشامية الدمشقية؟



01



**يهدف هذا المشروع إلى توثيق**

**المصطلحات الشامية ضمن إطار**

**لغوي قابل للاستخدام في**

**تطبيقات الذكاء الاصطناعي**

**ومعالجة اللغة العربية.**

02



## لماذا هذا المشروع مهم؟

بعد السنوات الأخيرة، ازداد اهتمام العالم باللهجة

1

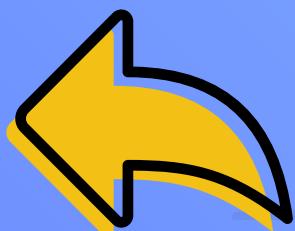
والثقافة السورية.

2

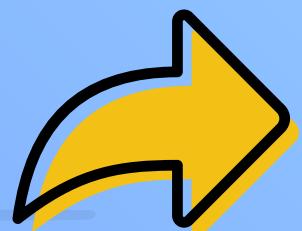
الحاجة إلى بيانات لغوية دقيقة للهجات العربية  
لاستخدامها في البحث، التعليم، والذكاء الاصطناعي.

3

معظم أدوات NLP لا تتعامل مع اللهجات جيداً بسبب  
نقص البيانات المنظمة.



03

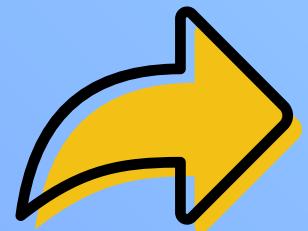
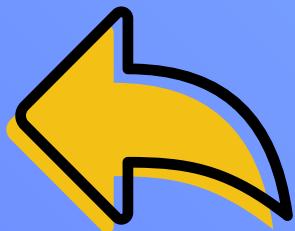


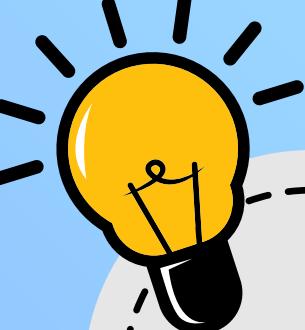


## الهدف النهائي

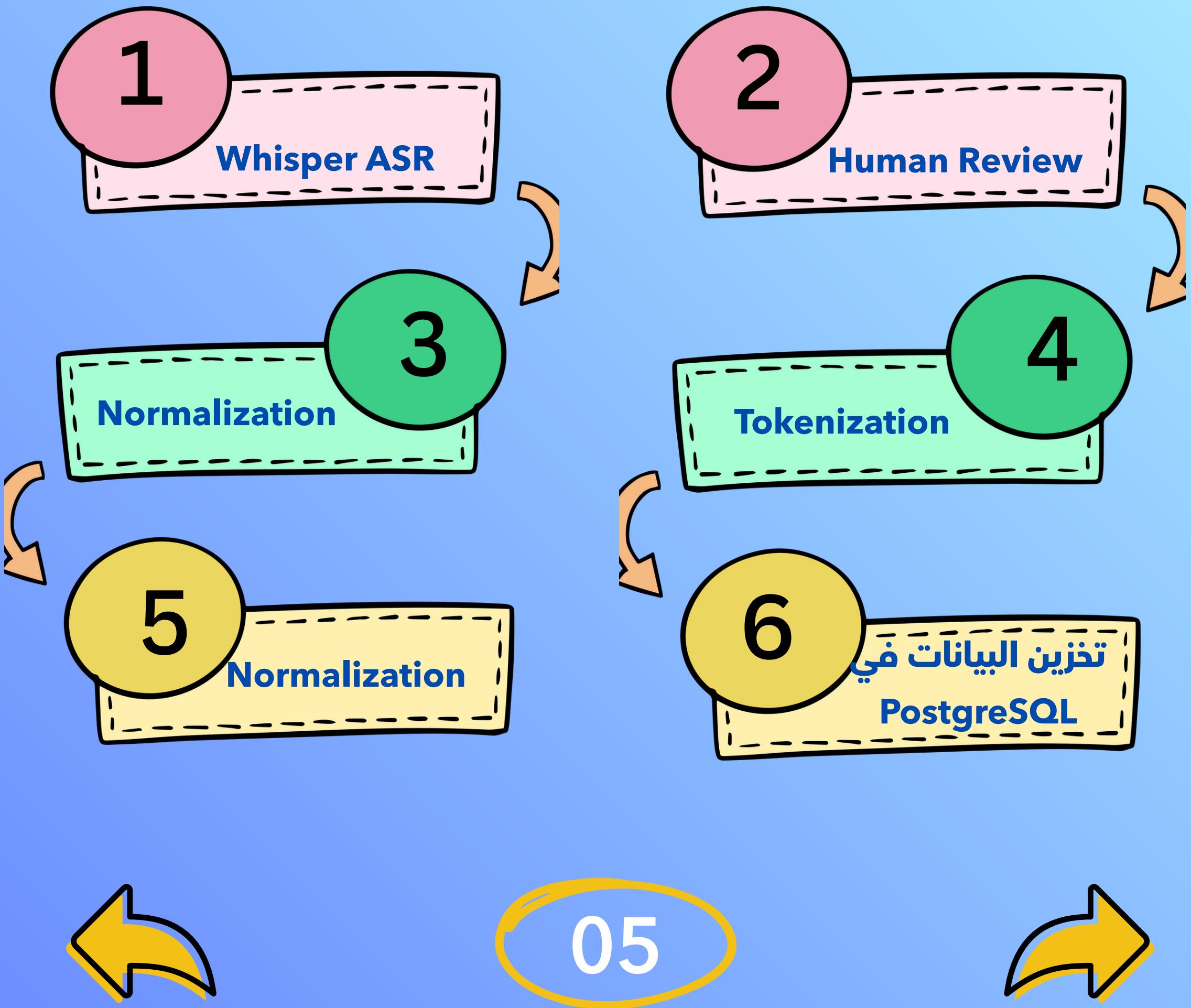
1. بناء قاعدة بيانات لغوية منظمة تشمل:
2. المصطلح الشامي
3. كتابته الموددة (Normalized)
4. معناه العربي الفصيح
5. دوره النحوی (POS)
6. السياق الذي يظهر فيه
7. تخزينه في قاعدة PostgreSQL للاستخدام المستقبلي

04





# مخطط العمل (Pipeline)



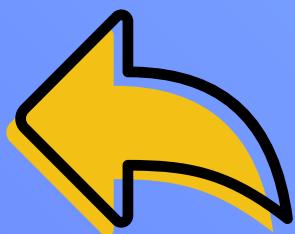


## من الصوت إلى نص

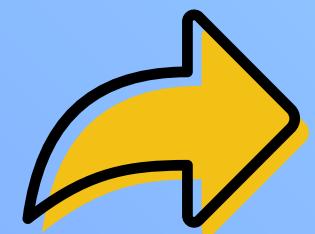
استخدام Whisper لتحويل التسجيلات إلى نص

مراجعة بشرية لتصحيح الأخطاء

توحيد الكتابة (Normalization)



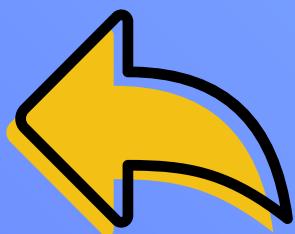
06



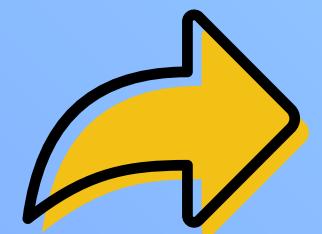


# معالجة النص

- 1 CAMeL باستخدام Tokenization Tools
- 2 إزالة التكرار الحرفي والضبابي
- 3 إضافة المعاني
- 4 تحليل نحوبي باستخدام Stanza



07





الناتج النهائي

**قاعدة بيانات لغوية منظمة للهجة الشامية**

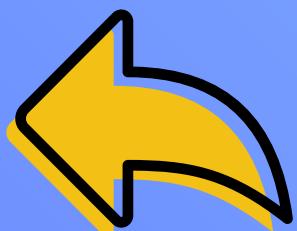
تدعم:

1. تدريب نماذج NLP

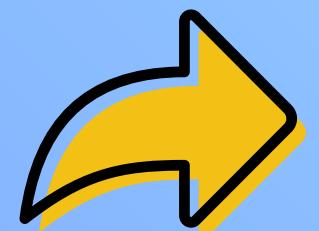
2. بناء Chatbots

3. تطوير أدوات تعليمية

4. دراسات لغوية للهجات السورية



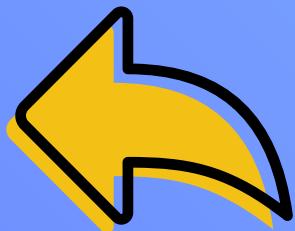
08



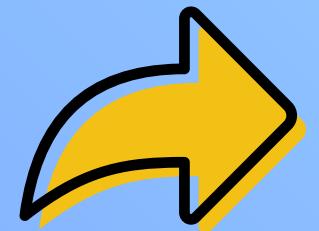


برأيك

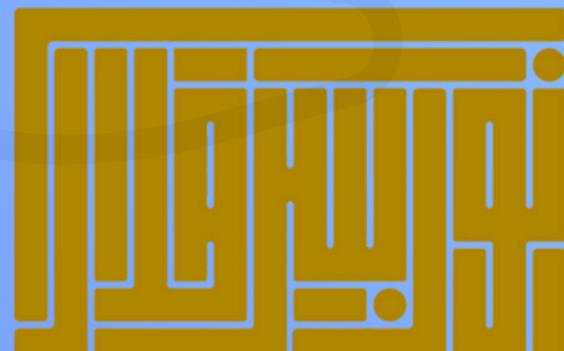
ما هي المصطلحات الشامية التي تعتقد  
أنها منتشرة بكثرة، ولكن لم يتم تضمينها  
بعد في أدوات الذكاء الاصطناعي؟



09



**تابعني للتتساعد على نشر المعلومات  
التقنية باللغة العربية**



Eng. Nour Berkdar

