

المحاضرة الثانية: لغات أنظمة المعلومات (اللغة الطبيعية واللغة التوثيقية)

أولا اللغات الطبيعية :

1. بداية نخرج إلى مفهوم نظم استرجاع المعلومات Information retrieval هو أحد النظم الفرعية في نظام المعلومات ووظيفته الأساسية هي تخزين واسترجاع المعلومات وفقا لاحتياجات مجتمع المستفيدين. ويمكن تعريفه أيضاً على أنه (علم البحث عن الوثائق وعن المعلومات داخل الوثائق وعن البيانات الوصفية التي تصف الوثائق، بالإضافة إلى البحث في قواعد البيانات وشبكة الانترنت)

- هناك استعمالات مشتركة وشائع بين كل من استرجاع البيانات واسترجاع المعلومات واسترجاع الوثائق واسترجاع النصوص، ولكن لكل منها كيانه العلمي الخاص ونظرياته وتقنياته.

- كما شهدت نظم استرجاع المعلومات تطورات عديدة منذ ظهورها حتى الوقت الحاضر وتزايد استخدام مصطلح نظم المعلومات منذ انتشار واستخدام تقنيات الحاسبات والاتصالات معلومات.

- وتوجد العديد من نظم المعلومات من بينها :

- نظم إدارة قواعد البيانات
- نظم المعلومات الإدارية
- نظم دعم اتخاذ القرار
- نظم الإجابة عن الأسئلة
- نظم المعلومات الجغرافية
- نظم استرجاع المعلومات أو نظام المعلومات الجغرافية.

2. معالجة اللغات الطبيعية NLP

هي مجال علوم الحاسوب واللغويات المعنية بالتفاعلات بين الحاسوب واللغات الطبيعية، والتي بدأت كفرع من الذكاء الاصطناعي والتي بدورها متفرعة من المعلوماتية

حيث أن هنالك جدال حول تقارب وتباين معالجة اللغة الطبيعية عن مجال اللغويات الحاسوبية. وقد قامت جمعية اللسانيات الحاسوبية بتعريف اللغويات الحاسوبية على أنها تركز على الجوانب النظرية لمعالجة اللغة الطبيعية . وترتكز الخوارزميات الحديثة لمعالجة اللغة الطبيعية في تعلم الآلة . كما أن الأبحاث الحديثة في خوارزميات تعلم الآلة الإحصائي تتطلب فهم عدد من المجالات المتباينة ، بما في ذلك اللغويات وعلوم الحاسب والإحصاء .

3. مستويات معالجة اللغة الطبيعية:

بالنسبة للنصوص المكتوبة، فإن تحليها يمر بـعدة مراحل تختلف باختلاف طريقة التحليل وغناء اللغة، فمثلا معالجة اللغة العربية تحتاج إلى تحليل يشمل لكل المستويات . ومستويات التحليل الأكثر انتشارا تتبع المراحل التالية :

التحليل الصرفي:

يقوم بتحليل الكلمة لمعرفة جذرها ووزنها الصرفي وما طرأ عليها من زيادة أو نقصان أو إعلال أو إبدال أو إدغام أو قلب، بالإضافة إلى معرفة ما اتصل بها من سوابق ونوع تلك السوابق، وما لحقها من ضمائر أو لواحق ومكونات تلك الضمائر واللواحق، ومعرفة نوع الكلمة نفسها، اسما أو فعلا أو حرفا، إلى غير ذلك من التوصيفات الصرفية التي تخص كل قسم من أقسام الكلم العربي.

التحليل النحوي

وهو الجزء الذي يهتم بعلاقة الكلمات بعضها مع بعض، هيكلية الجملة، وغيرها من المعلومات النحوية، ويعتمد على المرحلة الصرفية.

التحليل الدلالي

وهو الجزء الذي يهتم بفهم المقصود من الجملة عن طريق الربط المنطقي بين موضوع الحديث في الجملة ومعلومات من العالم الواقعي.

4. المجالات الرئيسية لمعالجة اللغات الطبيعية:

- ✓ تعرف علي الصوتيات
- ✓ تعرف علي أنماط الكلام
- ✓ تحليل صرفي
- ✓ تحليل نحوي
- ✓ تحليل دلالي
- ✓ تصنيف النصوص
- ✓ تدقيق أو تنقيح النصوص
- ✓ الترجمة الآلية
- ✓ فهم الأسئلة والإجابة عليها
- ✓ استرجاع المعلومات
- ✓ استخلاص المعلومات
- ✓ التنقيب في النصوص
- ✓ البحث عن المعلومات
- ✓ نظم التعليم الذكية.

محاضرة الثالثة : اللغات التوثيقية اللاحقة والتركيبية

1. مفهوم اللغة الوثائقية (التصنيف):

لغة: صنف الشيء أي جعله أصنافا ، يعني جمع و ترتيب الأشياء في أقسام تبعا للصفات المتشابهة. أما اصطلاحا: ترتيب مصادر المعلومات المتشابهة موضوعيا في مكان محدد و برمز خاص. كما أن **التصنيف في المكتبات** : هو فن اكتشاف موضوع الكتاب و الدلالة عليه برمز من رموز التصنيف الذي تستخدمه المكتبة.⁵

⁵ - البدوي, أبو زيد أحمد, فن التصنيف. القاهرة: دار الفكر العربي, 1993. ص. 15.