

التلخيص التلقائي التلخيص التلقائي هو خلق نسخة مختصرة من النص بواسطة برنامج كمبيوتر. [1][2][3] نتاج هذا الإجراء لا يزال يحتوي على أهم النقاط من النص الأصلي. إن ظاهرة الإغراق المعلوماتي باتت تعني أن الحصول على ملخصات متماسكة ومتطورة بشكل صحيح هو أمر حيوي. وكما أن الوصول إلى البيانات زاد كذلك الاهتمام بالتلخيص التلقائي. على نطاق واسع، الأشخاص يميزون بين نهجين : الاستخراج والتجريد. (والجمل أو الفقرات)، بشكل عام، يمكن للتجريد تلخيص النص بقوة وكثافة أكبر من الاستخراج، ولكن البرامج التي يمكنها القيام بذلك من الصعب تطويرها لأنها تتطلب استخدام تقنيات مولدات اللغات الطبيعية، أنواع الملخصات [عدل] هناك أنواع مختلفة من الملخصات على حسب ما يركز عليه برنامج التلخيص لتقديم موجز النص، على سبيل المثال ملخصات عامة أو ملخصات ذات الصلة بالاستعلام (التي تسمى أحيانا ملخصات منحازة للاستعلام). نظم التلخيص قادرة على خلق كل من نصوص ملائمة للاستعلام وملخصات عامة مؤلدة آليا اعتمادا على ما يحتاجه المستخدم. كما أنه من الممكن أيضا تلخيص وثائق الوسائط المتعددة، على سبيل المثال الصور أو الأفلام. بعض النظم ستقوم بإنشاء ملخص على أساس وثيقة مصدر واحد، والبعض الآخر يمكنها استخدام وثائق مصادر متعددة (على سبيل المثال، مجموعة من القصص الإخبارية عن الموضوع نفسه). التطبيقات ونظم تلخيص [عدل] الأول هو تلخيص عام، والذي يركز على الحصول على ملخص عام لمجموعة من الوثائق (سواء وثائق مكتوبه، الأخبار وغيرها). والثاني هو تلخيص ذات الصلة بالاستعلام، وتسمى أحيانا التلخيص القائم على الاستعلام، الذي يلخص بناء على الاستفسار. نظم التلخيص قادرة على إنشاء كل من ملخصات الاستعلام والملخصات العامة اعتمادا على ما يحتاج المستخدم. بعض الطرق المعاونة التلخيص [عدل] أن تقنيات تعلم الآلة من المجالات وثيقة الصلة مثل استخراج المعلومات أو تعدين النصوص تم استخدامها وتطويرها لمساعدة في التلخيص التلقائي. وبصرف النظر عن المُلخِصَات كاملة التلقائية (اف ا اس)، هناك نظم تساعد المستخدمين بمهمة التلخيص (التلخيص البشري بمساعدة الآلة)، على سبيل المثال من خلال تسليط الضوء على مقاطع لتكون موجودة ضمن الملخص، التقييم [عدل] من القضايا المستمرة في هذا المجال هو عملية التقييم. الحكم البشري لديه تنوع كبير على ما يعتبر جيد من الملخصات، مما يعني أن عملية التقييم التلقائي صعبة بشكل خاص. ويمكن استخدام التقييم اليدوي، ولكن هذا مكلف للوقت وللعمالة لأنه يتطلب البشر لقراءة ليس فقط الملخصات ولكن أيضا وثائق المصدر. اما عن القضايا الأخرى فتتعلق بالتماسك والتغطية. إحدى المقاييس المستخدمة في مؤتمرات فهم الوثائق السنوية التابعة لمنظمة نيسست، والتي تقدم المجموعات البحثية انظمتها لكل من التلخيص ومهام الترجمة، هو المقياس روج (بديل للانحياز الرجوعي لتقييم جيستينج) [1]. تحسب في الأساس من تداخلات ال(ن-غرام) بين الملخصات التلقائية وملخصات الإنسان المكتوبة من قبل. التداخل العالي المستوى من الضروري ان يدل على مستوى عالٍ من المفاهيم المشتركة بين الملخصين. علما أن مثل هذا المقاييس غير قادر على تقديم أي معلومات عن التماسك