

Генератор неологизмов – это прикладной проект по компьютерной генерации новых лексических единиц (неологизмов). Он представляет собой базу данных, созданную на основе комбинаторного сочетания трех групп словообразовательных элементов (префиксы, корни и суффиксы), с интерфейсом, позволяющим осуществлять поиск неологизмов по заданным параметрам (в том числе по укрупненным семантическим группам).

Пилотный набор словообразовательных элементов был аналитически выделен из «Проективного словаря» М. Эпштейна. На пилотном этапе было принято решение использовать словообразовательные элементы преимущественно латинского и греческого происхождения. Комбинирование производилось по шести моделям, или способам словообразования: суффиксальный, префиксальный, префиксально-суффиксальный, безаффиксный, сложение основ, сложение основ плюс суффиксация. Для выделения из получившегося массива неологизмов существующих слов и наиболее перспективных неологизмов в интерфейс генератора был также встроен такой показатель, как частота встречаемости сгенерированного неологизма в поисковой строке Google. В докладе пойдет речь о сложностях при реализации проекта и его перспективах.

Авторы: М. Эпштейн (автор идеи, идеолог и методолог проекта), Ю. Троицкий (руководитель разработки: концепция и архитектура), А. Павлов (компьютерная лингвистика), А. Плигузов (программная оболочка, интерфейс продукта).