

База представляет собрание электронных текстов на татарском языке. Вошедшие в корпус тексты в большинстве своем относятся к трем стилям – публицистике (60%), художественной прозе (35%) и научной литературе гуманитарного профиля (5%). Объем корпуса составляет более 45 млн. слов, число различных словоформ близко к 400 000..