

Acesta este un document demonstrativ, cu scopul de a ilustra funcționalitatea și capacitățile unui model lingvistic. Acesta este, în esență, o funcție matematică antrenată cu obiectivul să poată prezice un cuvânt pe baza cuvintelor din jur. Practic, la fiecare pas, modelul lingvistic asignează o probabilitate fiecărui cuvânt din dicționar. Cuvintele “potrivite”, fie că vorbim din punct de vedere tipografic, sintactic ori semantic, vor avea probabilități mari, în timp ce cuvintele nepotrivite vor avea probabilități foarte mici. Astfel, comparând un text dat bucată cu bucată, pot fi identificate eventuale nepotriviri: acelea sunt locuri în care ar trebui să începem să căutăm posibile greșeli. Mai jos ilustrăm câteva situații:

Scenariul greșelii de tipar:

Folosim un text generic, în care introducem greșeli, inclusiv nefolosirea diacriticelor: Primul Război mondial a fost un război global declanșat în europa, la care au participat peste 70 de milioane de militari, inclusiv 60 de milioane de europeni, mobilizați într-unul dintre cele mai mari războaie din istorie.

Scenariul greșelii sintactice

În timp ce primul scenariu este rezolvat și cu un simplu dicționar de verificări, un caz mai subtil îl reprezintă greșelile sintactice, de exemplu, un dezacord. Cei doi prieteni buni merge spre școală. Puțin mai subtil, luăm alt caz. Nu mai știu cât timp a trecut dar Anca și Andrei a fost mereu în gândul meu.

Scenariul greșelii semantice

Cea mai dificilă eroare, pe care am zice că doar un om o poate repera, presupune înțelegerea textului și folosirea cunoștințelor exterioare textului. Primul Război Mondial a început în anul 1912 și a fost un conflict extrem de sângeros. Ziua Națională a României este sărbătorită în data de 7 August. Toate că toate, dar ce se întâmplă când efectiv scriitorul nu reușește să se exprime fluent? Ce avem este că s-a produs o extorsiune a unui degajament ceea ce face că conflagrația să se propage recursiv, în paralel cu curgerea temporală.

REZUMAT

Expresia „Lorem ipsum dolor sit amet” apare în Ajutorul online din Microsoft Word. Această frază are aspectul unei expresii latine inteligibile. În realitate, este un nonsens.

INFORMAȚII SUPLEMENTARE

Deși expresia este un nonsens, are o lungă istorie. Expresia a fost utilizată de-a lungul a câteva secole de către tipografi pentru a expune caracteristicile cele mai pregnante ale fonturilor. Era utilizată deoarece literele conținute și spațiile dintre litere din această combinație prezintă, cât se poate de bine, grosimea, forma și alte caracteristici importante ale tipului de literă.

Ediția din 1994 a magazinului „Before & After” descrie expresia „Lorem ipsum ...” ca fiind o versiune latină reordonată a unui pasaj din de Finibus Bonorum et Malorum, un tratat de teorie etică scris de Cicero în 45 î.Hr. Pasajul „Lorem ipsum ...” este preluat din textul, „Neque porro quisquam est qui dolorem ipsum quia dolor sit amet, consectetur, adipisci velit ...”, ceea ce s-ar traduce astfel: „Nimănui nu-i place durerea în sine, nici nu o caută sau o dorește, deoarece este durere...”

În anii 1500, un tipograf a adaptat textul lui Cicero pentru a dezvolta o pagină cu eșantioane de tipuri de litere. De atunci, textul aparent latin a devenit standard al industriei tipografice pentru mostrele de text. Înainte de apariția publicațiilor electronice, proiectanții de grafică trebuiau să facă o machetă a proiectelor lor desenând linii care imitau scrisul pentru da indicii despre text. Apariția foliilor auto-adezive preimprimare cu „Lorem ipsum” a oferit o modalitate mai realistă de indicare a locului unde va fi text într-o pagină.