



Big Data

Hola, sóc Juan Tomás García, de SPG, i sóc el responsable de big data (dades massives). Fa vint anys, la quantitat d'informació que hi havia al món era la que era: limitada. Ara disposem de moltíssima informació, que és el que entenem per «dades massives». Es pot preveure quan se t'espatllarà el cotxe: tots els cotxes vénen ja equipats amb un connector, que agafa totes les avaries, tots els sensors i és capaç d'emmagatzemar aquesta informació i dir-te que el cotxe se t'espatllarà o que li hauries de canviar l'oli.

El màxim projecte de dades massives que hi ha ara mateix és la quantitat d'informació que ha estat capaç de recollir l'Agència Nacional de Seguretat dels Estats Units i que està tractant. Òbviament hi ha empreses, com Google, que són capaces d'oferir-te millors resultats amb la informació. Temps enrere hi havia molt poca informació i de sobte se n'està generant moltíssima i aquest creixement és exponencial.

La part important de les dades massives que té lloc ara es diu machine learning i és de quina manera les màquines són capaces d'aprendre i trobar patrons que no han estat ensenyats prèviament per un programador. Les dades massives i el machine learning sobretot investigaran tota la part del coneixement automàtic i de les màquines, ajudaran a millorar-lo.

Tot això que sembla una mica ciència-ficció ja es fa i jo crec que és una via que a mi em dóna esperança; que no serveix només per espiar-nos o per saber els nostres hàbits comercials, sinó que aquesta tecnologia té un doble ús i farà que millorem i que millorem força.

De vegades no és tan important tenir tota la informació, sinó saber capturar parts de la informació que tenen sentit. Això és molt pragmàtic i està molt orientat al negoci, perquè era la part que tenia més recursos per explotar-lo, però cada vegada s'aplica més a la recerca. Coses com intentar trobar la correlació que té un ADN amb un determinat tipus de càncer, encara que sembli que no són dades massives, suposen manejar moltíssima informació, requereixen el mateix esforç de càlcul i cada vegada n'aprenem més gràcies a això.

De les coses més interessants que passen amb les dades massives és que no són només les dades estàndards o clàssiques que es generen, sinó coses com el maquinari lliure, com Arduino, les impressores 3D, que et permeten fabricar dispositius... S'està canviant al que s'anomena «la Internet de les coses» i que tindrà com a proper nom «la internet de tot», perquè ja no és tan sols de les coses, sinó de tot el que passa.

Jo crec que és una de les convergències més interessants i més a curt termini que es produiran.