CULTURA I COMUNICACIÓ | | Actualitzat el 06/11/2019 a les 10:15

Conèixer el «big data»: el periodisme com una forma de coneixement i el que això implica

Incorporar acríticament el periodisme de dades té conseqüències que no es poden ignorar si el periodisme s'entén com una forma de produir coneixement | Els debats contemporanis pel que fa al disseny de models causals poden donar respostes



L'augment de consum de gelats pot tenir alguna correlació amb l'augment de la inseguretat a Barcelona? | The Blue Flower Suivre/Flickr.

Imagini el lector una ciutat qualsevol -Barcelona, per exemple- en què les xifres pel que fa a crims violents s'han disparat durant l'estiu. Imaginem, de passada, que aquest augment de la criminalitat s'ha donat en paral·lel a un augment del consum de gelats. Una pregunta podria sorgir d'aquestes dues dades: **hem de culpar els gelats per la crisi de seguretat de la capital catalana? Aquesta fal·làcia, que ja és un clàssic en el debat sobre la confusió entre correlació i causa**, és una bona manera d'introduir algunes de les preocupacions que enuncia l'autor d'un dels textos més rellevants que s'han fet mai sobre la naturalesa de la professió periodística. Em refereixo a la tesi doctoral del professor de periodisme a la Facultat de Comunicació i Relacions Internacionals Blanquerna Pablo Capilla: una aportació que es compromet amb una idea de la professió en tant que forma de produir coneixement. Un compromís gens trivial, pel que explicaré a continuació.

A l'apartat de la tesi de possibles vies d'investigació posteriors implícites en la seva contribució, Capilla cita una afirmació prou controvertida del llibre de Viktor Mayer-Schönberger i Kenneth Cukier *Big Data: A Revolution That Will Transform How We Live, Work and Think*. Mayer-Schönberger i Cukier escriuen que, **quan es treballa amb una mostra completa d'informació, s'ha d'abandonar la pretensió d'establir relacions de causa-efecte**, com es feia quan es treballava amb una mostra representativa, i acontentar-se amb poder establir un ventall de

possibles correlacions i prou. Aquesta asseveració té un efecte potencialment devastador pels que, com Capilla, creuen que el periodisme o qualsevol altra disciplina que treballa amb grans conjunts de dades aspira a conèixer l'objecte que estudia. Pels que, en definitiva, voldrien establir la relació causal adequada que expliqui l'increment dels crims a Barcelona. **Hi pot haver coneixement sense determinar les relacions causals entre esdeveniments?** Aquesta és una pregunta i un repte que no poden menystenir aquells que creuen que l'aportació del *big data* en periodisme és massa rellevant com per incorporar-la acríticament.

@pensem_cat

Mayer-Schönberger, professor de governança a Internet de la Universitat d'Oxford, sosté que bona part de l'anàlisi del *big data* ?incloent-hi les aportacions de la intel·ligència artificial o l'anomenat *machine learning*? és correlacional. Això vol dir que **els sistemes que s'entrenen avui per processar dades ho fan sense conèixer les relacions causals que hi subjauen**. Però recentment, apuntava Mayer-Schönberger, el científic computacional Judea Pearl ha desenvolupat una via amb què representar la causalitat matemàticament: un pas crucial amb què "calcular" la causalitat. La de Pearl fou una contribució mereixedora del prestigiós premi Turing i que hom pot trobar en el seu llibre més divulgatiu, *The Book of Why*. El professor d'Oxford considera la proposta de Judea Pearl una de les més prometedores per superar els riscos que assenyalaven des que van publicar la seva pròpia reflexió enguany ja fa sis anys.

Dur Judea Pearl al periodisme

Com bé apunta Capilla a la seva tesi, **que implícitament l'horitzó del periodisme sigui l'objectivitat no fa sinó crear confusió epistemològica entre els professionals**. La impossibilitat manifesta d'assolir aquesta tan anhelada objectivitat ha acabat traduint-se en la introducció de succedanis com 'neutralitat' o 'imparcialitat' que hom no sap ben bé a què refereixen, però que a efectes pràctics deixen als periodistes al mateix carreró sense sortida. En altres camps de les ciències socials com la historiografia o l'antropologia es considera exòtic, com a mínim, que algú asseguri que disposa d'un mètode per historiar objectivament. L'ortodòxia anglosaxona estableix que, en periodisme, allò més semblant a la imparcialitat és donar veu a tothom, aportar totes les dades i, d'aquesta manera, deixar que sigui el lector qui faci la mitjana. El que, com sosté Capilla, és epistemològicament absurd.

Dir que el periodisme és una forma de coneixement té conseqüències transcendentals. Implica que quan un periodista demana una fulla de càlcul amb grans conjunts de dades a una administració ha de ser capaç de discernir entre correlacions espúries i relacions causals. Ha de poder comprometre's amb allò que és més plausible, no mostrar-ho tot de forma acrítica. I això ens situa en una disjunció: o bé s'alcen els estàndards d'exigència professional o bé es creen aplicatius que puguin acompanyar el periodista en els processos inferencials a la llum del progrés pel que fa a l'automatització de la presa de decisions. I tenint en compte les dificultats de canviar certs mals vicis com el del periodisme declaratiu, en sintonia amb la citada ortodòxia, el més factible sembla ser encomanar-nos a les noves tecnologies. En una entrevista recent al professor de periodisme de l'Autònoma de Barcelona i de Blanquerna, Francesc Burguet, l'autor de *Les trampes dels periodistes* emprava una metàfora que ens pot ajudar a explicar ben bé com l'automatització ens podria ajudar.

Burguet il·lustrava el que ell creu que és una bona interpretació periodística amb l'analogia dels passatemps pels infants en què aquests han de resseguir uns punts numerats. "Com més puntets, més clar, més indiscutible resultarà el dibuix final", afirma. "Exactament el mateix passa amb la informació: com més dades pertinents, rellevants, que abonen la interpretació i valoració exposades de manera explícita o implícita, més legítima serà la nostra interpretació", rebla Burguet. Quelcom que podem dur al debat que ens ocupa. Pearl mostra al seu llibre els anomenats "factors de confusió", una variable que acaba distorsionant la mesura que obtenim quan associem dues variables. En l'exemple inicial, aquesta variable seria l'augment del consum de gelats. **Un periodista competent a la manera com ho entén Capilla però, entendria la relació causal existent a partir de dos factors més que caldria dur a la situació avaluada: la calor i la**

presència del turisme, dues variables la rellevància de les quals augmenta amb l'estiu.

Si sap discernir entre les fonts adequades, el professional trobarà la literatura científica rellevant que constata la plausibilitat d'establir una relació de causa-efecte entre l'augment de les temperatures i la disposició a actuar amb violència, d'una banda, i la sobrerepresentació dels turistes en les estadístiques pel que fa a víctimes de certs crims violents que impliquen sostracció amb força. D'aquesta manera, la calor explicaria els dos elements que mostràvem al principi: l'increment del consum de gelat i del crim violent, i el factor de l'augment del turisme reforçaria el vector explicatiu de la segona dada. I sent conscient de quin és el volum de producció que exigeixen les redaccions dels mitjans, no seria una mala idea automatitzar la modelització que permet corregir factors de confusió. **Així, en plena història global per la desinformació, seria més difícil que s'acusi als professionals d'interpretacions il·legítimes**.

Bibliografia

- Francesc Burguet: *Les trampes dels periodistes* (Grup 62, 2004)
- Juan Pablo Capilla: *El debate epistemológico en el periodismo informativo. Realidad y verdad en la información* (tesi doctoral; Universitat Ramon Llull, 2015)
- Judea Pearl: *The Book of Why* (Penguin, 2018)
- Viktor Mayer-Schönberger i Kenneth Cukier: *Big Data: A Revolution That Will Transform How We Live, Work and Think* (Houghton Mifflin Harcourt, 2013)