DOPPIOZERO

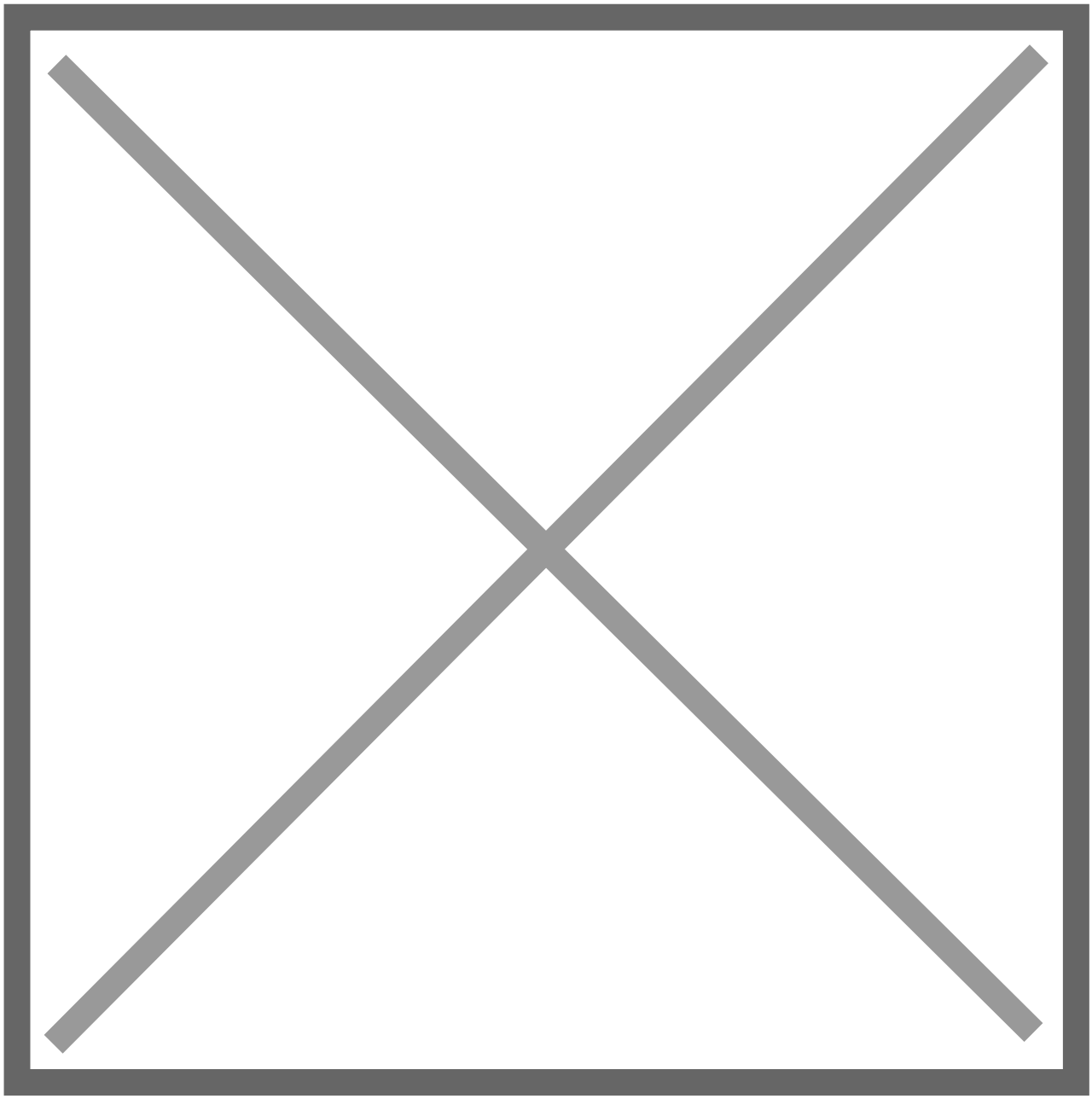
Il cuculo rivelato

Antonio Sgobba

11 Novembre 2013

Il richiamo del cuculo di Robert Gailbraith è uscito in Italia (pubblicato da Salani). Si tratta del primo giallo di J.K. Rowling, pubblicato sotto falso nome. A svelare la reale identità dell'autore è stato il *Sunday Times*. Ma come ha fatto il quotidiano inglese a scoprire che il vero autore di *The Cuckoo's Calling* non era uno sconosciuto esordiente ma l'autrice di Harry Potter?

Un ruolo decisivo lo hanno avuto due studiosi di *digital humanities*: Peter Millicam, docente di filosofia della Oxford University, e Patrick Juola, professore di informatica alla Duquesne University di Pittsburgh. Il primo si è servito di Signature, un programma di stilometria sviluppato nel 2003, il secondo di Jgaap (Java Graphical Authorship Attribution Program), un altro software molto usato per l'analisi testuale.



Certo, tutto era partito dall'intuito di un critico letterario. Leggendo il libro aveva notato alcune somiglianze con lo stile dell'autrice di Harry Potter. Solo un sospetto, gli indizi prendono corpo quando arriva una soffiata via twitter da un profilo poi rimosso. A quel punto il giornalista freelance Cal Flyn e il caporedattore del quotidiano londinese Richard Brooks hanno pensato di corroborare le loro intuizioni affidandosi all'analisi stilistica computazionale.

La procedura è stata spiegata da Juola nel blog Language Log. I due software hanno analizzato il testo del giallo comparandolo con altri quattro romanzi: uno

era *Il seggio vacante* della Rowling, il suo unico libro non fantasy. Gli altri erano opere di P.D. James, Ruth Rendell, Val McDermid. Tra i quattro candidati il più simile era senza dubbio il libro della Rowling. Per raggiungere questo risultato i romanzi sono stati sottoposti a quattro test. Il primo calcola la lunghezza media delle parole che compongono un testo.

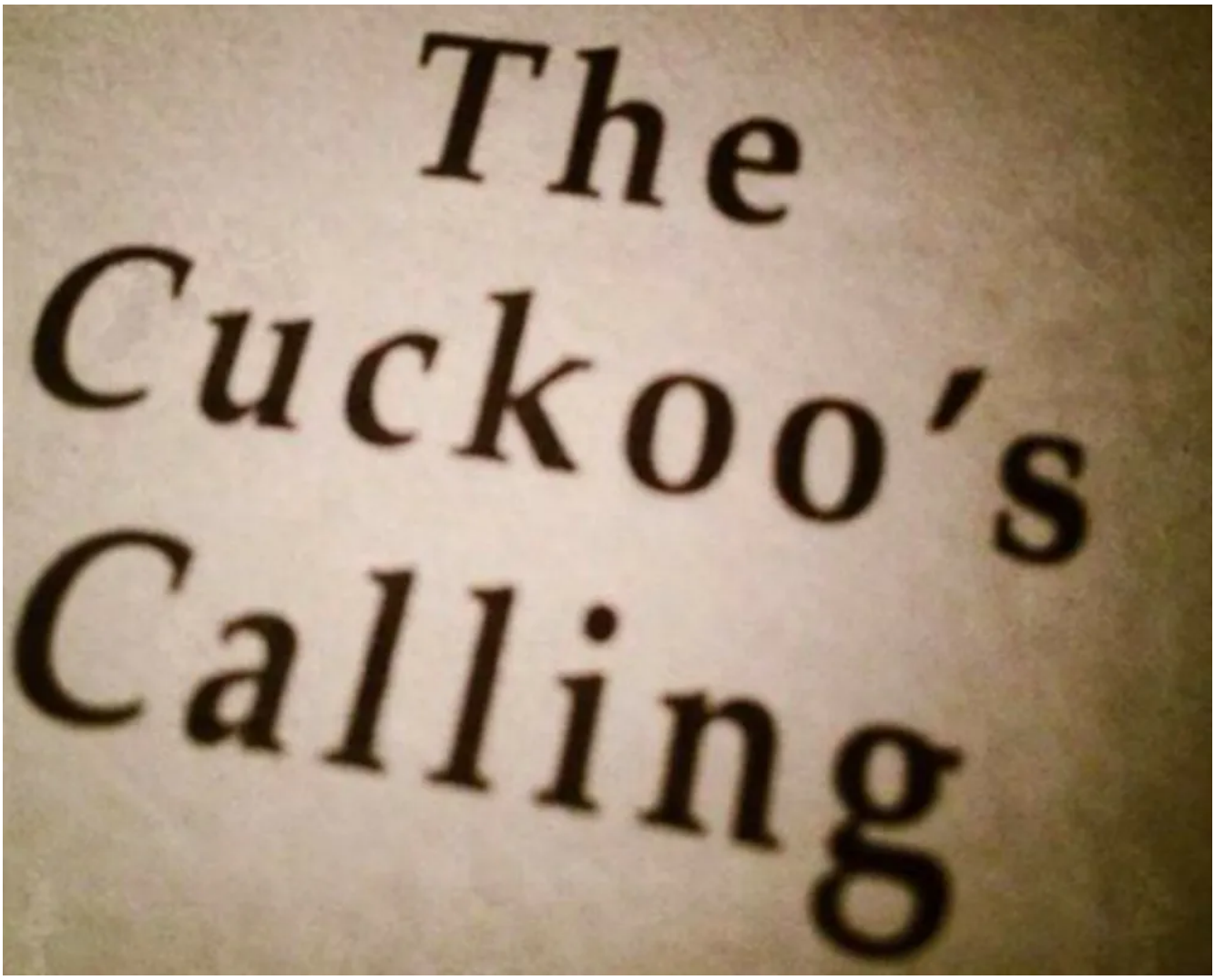
Un sistema considerato non del tutto attendibile se privo di altri riscontri. Per questo Millicam e Juola hanno anche analizzato le stringhe di parole consecutive, prese a coppie e a gruppi di quattro. L'esame decisivo è stato il calcolo della distribuzione delle cento parole più frequenti nel testo. Comprese proposizioni, congiunzioni e articoli. È una della caratteristiche che meglio definiscono uno stile. Un autore che cercasse di nascondersi dietro un vocabolario diverso non riuscirebbe mai a falsificare il modo in cui si serve delle particelle che costituiscono l'impalcatura del suo linguaggio.

Non è un caso infatti che le applicazioni più frequenti di programmi come Signature e Jgaap siano soprattutto in ambito legale. Gli atti di un recente seminario della Brooklyn Law School, pubblicati nell'ultimo numero del *Journal of Law and Policy*, mostrano come le storiche contrapposizioni tra linguisti tradizionali e informatici possano essere superate. Scrive l'organizzatore, il giurista Lawrence M. Solan: «Chi lavora con gli algoritmi può migliorare i propri modelli incorporando i suggerimenti di chi studia da sempre gli stili di scrittura, e chi è abituato ai metodi tradizionali può servirsi delle analisi quantitative per dare più forza alle proprie conclusioni».

Una prima versione di questo articolo è uscita su La Lettura *del* Corriere della sera

---

Se continuiamo a tenere vivo questo spazio è grazie a te. Anche un solo euro per noi significa molto.
Torna presto a leggerci e SOSTIENI DOPPIOZERO


The
Cuckoo's
Calling