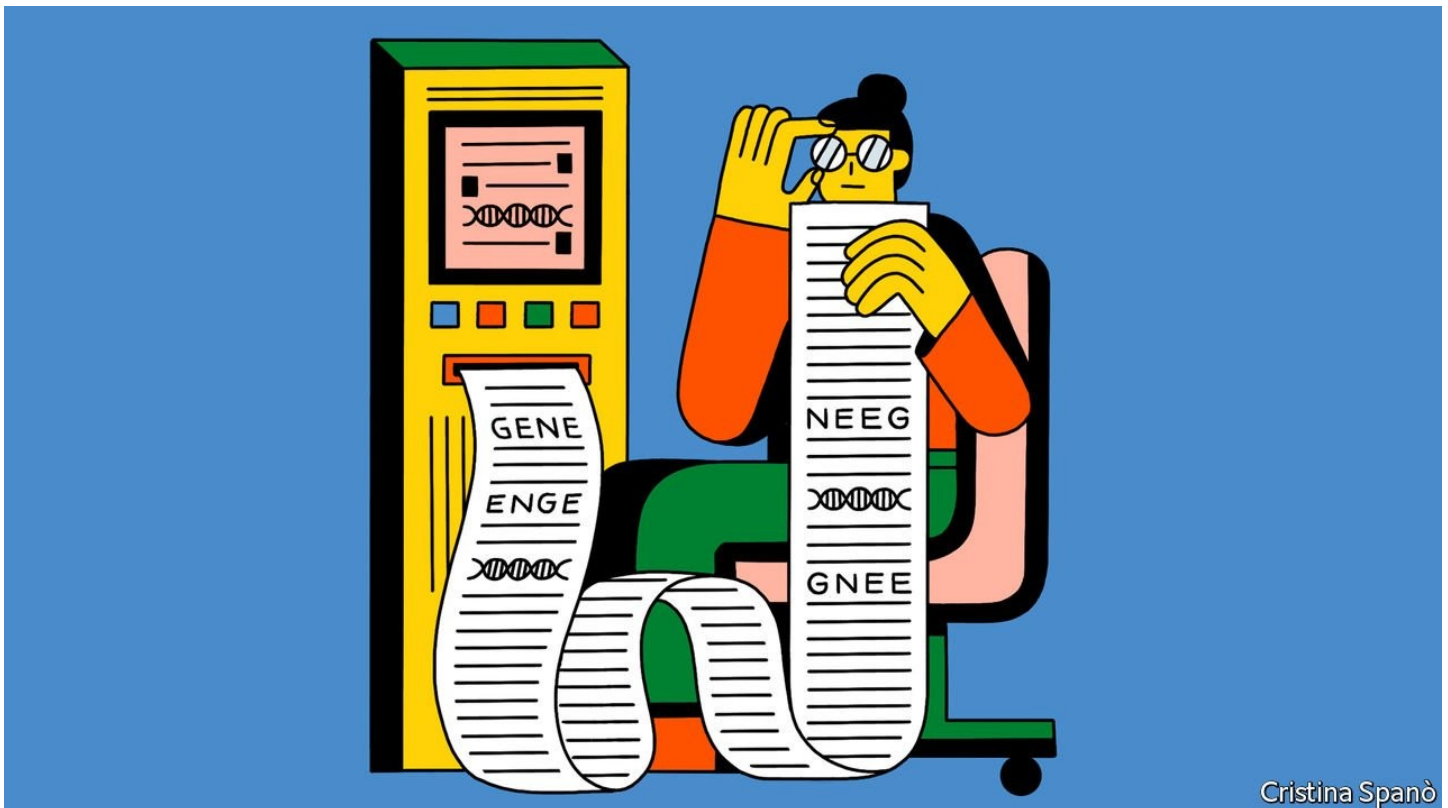


Scientific errors

Software is subbing mistakes into scientific papers

The names of genes are particularly vulnerable



Sep 2nd 2021

Give this article



MICROSOFT EXCEL'S zeal for autocorrection has long irritated casual users. Lists of numbers preceded by dashes are misread as instructions to subtract. Phone numbers lose their leading zeroes. Credit-card numbers get re-expressed in scientific notation

Listen to this story

0:00 / 0:00

Listen on the go

Get *The Economist* app and play articles, wherever you are

 Play in app

Geneticists struggle with a particular version of this problem. A gene called Membrane Associated Ring-CH-type finger 1, commonly known as MARCH1, is, for instance, frequently re-encoded as the date March 1. Something similar happens to genes in the Septin family, of which SEPT1 is a member, and to Basic Helix-Loop-Helix Family Member E41, often known as DEC2.

This problem was first noticed in 2004, but was brought to wider attention in 2016 by Mark Ziemann of Deakin University, in Australia. In July Dr Ziemann followed up with a paper in *PLOS Computational Biology* entitled “Gene name errors: Lessons not learned”. By surveying 166,000 genomics-related papers published between 2014 and 2020, he and his co-authors showed that the number of papers using Excel has steadily increased, and the proportion plagued with autocorrect errors continues to hover at around 30%.

Errors have also been flagged by researchers in other languages. In Portuguese, for instance, AGO2 (Argonaute RISC Catalytic Component 2) rebrands itself as Agosto 2. Dutch users experience problems with MEI1 (Meiotic Double-Stranded Break Formation Protein 1), “Mei” being the Dutch for “May”. And geneticists in Finland, where the first month of the year is called Tammikuu, find TAMM41 (Mitochondrial Translocator Assembly And Maintenance Protein) encoding itself as the 41st of January.

This continuing state of affairs is surprising. In August 2020, aware they were playing David to Microsoft’s Goliath, the committee which standardises gene

names renamed those beginning MARC, MARCH and SEPT to begin MTARC, MARCHF and SEPTIN, and rebranded DEC1 as DELEC1. Other problematic gene names remained, but this was widely seen as a step in the right direction. Dr Ziemann's latest paper, though, suggests that few researchers have taken it.

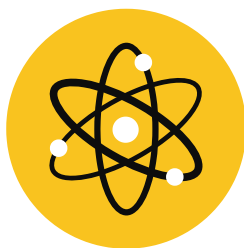
Such errors often seem amusing rather than dangerous. But they reflect a deeper problem, which is that spreadsheets allow such silent errors, too well camouflaged for authors or their editors to spot, to go undetected for years. To stop this, Dr Ziemann recommends researchers abandon software like Excel in favour of bespoke code written with scientific applications in mind. Such programs are less susceptible to erroneous autocorrections and easier to audit. Whether his advice will be heeded this time round remains to be seen. ■

An early version of this article was published online on September 1st 2021

This article appeared in the Science & technology section of the print edition under the headline "Autoincorrect"

Reuse this content

The Trust Project



SIMPLY SCIENCE

Curious about the world? Enjoy a weekly fix of our mind-expanding science coverage

Delivered to you every week

Sign up

More from Science & technology

Everyone's going to the Moon

In 2022 a Moonrush will begin in earnest

The new covid variant

Omicron causes a less severe illness
than earlier variants

A little covid cheer

Omicron covid cases are less severe than Delta, new studies say

Subscribe

Group subscriptions

Reuse our content

The Trust Project

Help and contact us

Keep updated



Published since September 1843 to take part in *“a severe contest between intelligence, which presses forward, and an unworthy, timid ignorance obstructing our progress.”*

[About](#)
[Advertise](#)
[Press centre](#)
[Store](#)

[The Economist Group](#)
[Economist Intelligence](#)
[Economist Impact](#)
[Economist Events](#)
[Working Here](#)
[Which MBA?](#)
[GMAT Tutor](#)
[GRE Tutor](#)
[Executive Jobs](#)
[Executive Education Navigator](#)
[Executive Education: The New Global Order](#)
[Executive Education: Business Writing](#)

[Terms of Use](#) [Privacy](#) [Cookie Policy](#) [Manage Cookies](#) [Accessibility](#) [Modern Slavery Statement](#)

[Do Not Sell My Personal Information](#)

Copyright © The Economist Newspaper Limited 2022. All rights reserved.