

Cada vez son más habituales los grandes reportajes que se apoyan en ingentes cantidades de datos: casos de corrupción y malversación mundiales, resultados electorales, eventos deportivos de atención planetaria, cambios climáticos y migratorios. Para cubrir estas historias, los periodistas se rodean de programadores, a veces ingenieros con talento periodístico, a veces periodistas con conocimientos de código.

[Imagen de cabecera: [Proyecto cartográfico sobre Abasto Market Supplier](#) de [Julieta Bernstein](#).]

Hoy tenemos acceso a tan gigantesca cantidad de datos que para referirnos a ellos ya no son suficientes los números y recurrimos a las letras. Y aunque buena parte de esos petabytes son *cookies*, comunicaciones M2M solo comprensibles entre máquinas, contenidos duplicados y/o copiados y/o requetecopiados, vídeos ilegales que YouTube tarda demasiado tiempo en borrar..., en definitiva, morralla digital que ocupa lugar, sí hay datos que son relevantes. Y accesibles. La transparencia tomada en serio ha hecho emerger desde los archivos de instituciones públicas, organizaciones y algunas empresas todo tipo de datos que sí importan porque nos dicen cómo somos. Y es en este océano de datos, cuando los libros de Excel se vuelven torpes e incapaces, donde los programadores se convierten en imprescindibles compañeros de los periodistas.

Y no sólo hablamos de las imaginativas infografías interactivas de los Juegos Olímpicos o del detalle del voto barrio a barrio de las elecciones. El trabajo de los programadores va mucho más allá de la visualización de datos, que es cierto que es la parte más reconocible y viral de sus funciones, pero ésta solo se sostiene si detrás hay una base de datos suficientemente depurada y trabajada.

La transparencia tomada en serio ha hecho emerger desde los archivos de instituciones públicas, organizaciones y algunas empresas todo tipo de datos que sí importan porque nos dicen cómo somos

La combinación de programación y periodismo llegó hace más de un lustro a España, y ya son varios los cursos y másteres que se imparten sobre esta especialidad, casi siempre bajo el paraguas del periodismo de datos. Su implantación en la prensa española es general, al menos entre los medios más grandes y los nativos digitales. Normalmente suelen ser equipos de dos o tres personas a los que se recurre para buscar o plasmar datos una vez está decidido el enfoque de la información. Los programadores no suelen participar a la hora de elegir los temas, salvo en unos pocos casos: cuando quien escribe el código no es un ingeniero, sino un periodista con conocimientos de programación.

Periodistas que saben programación

En [El Confidencial](#) trabajan cinco perfiles relacionados con la programación: uno en la unidad de datos, otro en automatización de contenidos y tres más en formato que comparten su tiempo con el departamento de *branded content* (contenidos vinculados a una marca). En total, cinco programadores para una redacción de noventa redactores.

Jesús Escudero es quien trabaja en la unidad de datos. Terminó Periodismo, cursó un máster de periodismo de datos y, después, hizo un curso MOOC (siglas en inglés de Cursos Online Gratis) de R, un lenguaje de programación muy útil para el análisis estadístico y gráfico de ingentes cantidades de datos. “Cuando entré [al máster], no sabía casi ni Excel. Había trabajado con informes de sueldos de alcaldes, pero siempre a mano, con Word y algo de Excel. Entré en la programación por necesidad, porque si te encuentras con un proyecto que necesite una base de datos extensa, entonces Excel no te vale porque solo te deja trabajar con un millón de filas y no vas a trocear la base de datos”, explica Escudero mientras recuerda que para [un reportaje sobre el éxodo urbano](#) analizaron 53 millones de filas de datos procedentes de 31 archivos.

Aunque se define como un noventa por ciento periodista y diez por ciento programador, Escudero emplea casi todo su tiempo en combinar datos, limpiarlos y homogeneizarlos. “Raro es que me llegue una base de datos limpia. Como con las elecciones. Lo suyo sería que Interior te mandara todos los datos en un CSV (archivo de Excel), pero la realidad es que no vienen así y tú tienes que limpiar todo eso para que luego se pueda usar”.

Y limpiar no es solo separar el grano de la paja, sino adecuar los formatos —a veces pueden llegar varios formatos distintos en un mismo documento o ser ilegibles para los lenguajes de programación—, reordenar las cifras en nuevas columnas buscando la sencillez o crear fórmulas que crucen datos para que los periodistas puedan obtener números que ayuden a explicar lo que están contando en el reportaje.

En el caso de Escudero, su relación con los periodistas es fluida. “En *El Confidencial* no hay una jerarquía rígida, así que hay un diálogo donde cualquiera puede proponer cosas. A veces viene un periodista con una pista o una hipótesis y pide al programador que lo confirme con los datos. Y al revés: nosotros podemos encontrar un dato que nos de pie a proponer algo a los periodistas.”

La tirantez viene cuando, explica Escudero, el periodista pide al programador que los datos avalen su hipótesis y resulta que estos no siguen esa línea. “Hay gente que tiene la mente más abierta para decir ‘Oye, esta hipótesis mía no se ha validado’ y no pasa nada, y otros que dicen ‘Pues habrá algún dato que la avale’ y siguen adelante aunque los datos digan lo contrario”. No pasa todos los días, reconoce Escudero, pero pasa.

Programadores que saben periodismo

En el otro extremo están los programadores de carrera que trabajan en las redacciones. También su entorno de trabajo son los datos, pero llegaron a estos desde una ingeniería. Como le ocurrió a David Cabo, codirector de [Civio](#).

Civio es una fundación sin ánimo de lucro dedicada a la investigación de los poderes públicos a través del periodismo y la innovación. Su equipo está formado por diez personas, de las que dos se dedican a la programación a tiempo completo. Una se encarga de la parte de diseño y visualización de datos, y la otra, de conseguir y limpiar los datos y también del mantenimiento de los sistemas y del desarrollo de la aplicación. A estas dos personas hay que sumar a David Cabo, quien combina sus responsabilidades como codirector con las de seguidor de datos procedentes de contrataciones públicas y del Gobierno.

En total, dos personas y media dedicadas a la programación en una redacción de diez integrantes. “Es posible que, al compararnos con un medio, la parte de informática esté sobredimensionada”, dice Cabo, y explica que

en *Civio*, además de la parte informativa, también ofrecen otros servicios que requieren de desarrollo y mantenimiento informáticos.

Para Cabo, que es ingeniero superior en Informática, la combinación de periodismo y tecnología permite desarrollar trabajos que de otra forma no sería posible y pone como ejemplo el [Indultómetro](#), una clasificación de todos los indultos concedidos desde 1996. “Nosotros ya estábamos trabajando en sacar la base de datos completa de los indultos de los últimos 20 años antes de que llegara Eva [Belmonte, codirectora], pero cuando llegó ella, empezó a analizar tendencias y a encontrar patrones y problemas de fondo en el uso de los indultos, en vez de centrarse en casos particulares. Esa combinación de periodismo y tecnología es lo que hace que puedas sacar más temas de fondo, especialmente ahora que la contratación pública te permite sacar todas estas cosas que igual un periodista solo no podría hacer.”

Muchos medios publican no solo los datos en bruto con los que han hecho los trabajos periodísticos sino también el código de los programas que han usado para el tratamiento y visualización de las cifras

Preguntado por la parte menos buena de trabajar mano a mano con periodistas, Cabo responde que no la hay pero sí destaca una diferencia importante: la cantidad de información que publicas mientras estás trabajando. “Los periodistas trabajan en secreto, con una paranoia de no puedo contar nada porque piensan que va a venir alguien a apropiarse de tu trabajo. Y esto en informática no es así. Aunque hay distintos tipos de gente, existe la costumbre de trabajar en abierto y en público. Y esto en periodismo es imposible. El periodista está obsesionado con llegar el primero y no puede contar nada”. “En nuestro caso”, prosigue Cabo, “siempre ha habido cierta tensión ahí porque queremos publicar todo. Al final hemos conseguido cierto equilibrio: publicamos cosas antes de lanzar, pero no todo”.

Se publique todo o no, lo cierto es que el periodismo de datos ha traído nuevas cotas de transparencia en el trabajo periodístico. Muchos medios publican no solo los datos en bruto con los que han hecho los trabajos periodísticos sino también el código de los programas que han usado para el tratamiento y visualización de las cifras. “Si nosotros demandamos transparencia a los políticos y empresas, pues también tendremos que hacer lo mismo con los ciudadanos y lectores”, explica Jesús Escudero, de *El Confidencial*.

“Con los datos pasa lo mismo que con las fuentes: los reportajes con fuentes anónimas tiendes a no creértelos del todo. Si muestras en cuatro o cinco párrafos metodológicos de dónde salen los datos y cómo los has analizado, es mucho mejor. Para profundizar y también para detectar errores. Si alguien detecta que me he equivocado en los datos, mejor que sea a la media hora de publicarlo que cuando lleve una semana y el impacto de la noticia haya sido importante”, concluye Escudero.

Un futuro prometedor

El periodismo realizado mano a mano entre periodistas y programadores vive aparentemente un momento

dulce y cada vez son más los trabajos con fuerte carga de datos que despiertan atención. Sin embargo, algunos periodistas especializados denuncian que sufren su propia versión del síndrome del informático al que sus compañeros llaman para que arregle impresoras o desbloquee equipos. “Creo que las unidades de datos van a seguir existiendo, pero deberían especializarse en grandes proyectos”, explica Escudero. “En el día a día, los programadores estamos satisfaciendo necesidades básicas de estadística porque la gente no sabe descargarse un Excel y trabajarlo. Las técnicas de recuperación de datos o saber cómo solicitar el acceso a informes de transparencia deberían ser una herramienta básica más, como saber hacer una entrevista, no una especialidad solo de los que se dedican a los datos. De esta forma, nos quitaría trabajo a las unidades de datos y podríamos centrarnos en los grandes proyectos que requieren de programación y de millones de registros”, concluye Escudero.

Para David Cabo, de Civio, “quiero creer que el número de informáticos en las redacciones crecerá. Son muy caros, pero el periodismo es una industria que trabaja con datos, y los informáticos pueden aportar mucho, no solo haciendo gráficos sino abriendo nuevas vías a la hora de analizar información. Esa es la promesa del periodismo de datos. Si cuajará o no, ya veremos, porque ahora mismo está todo un poco turbulento y la gente es escéptica con el futuro del periodismo, pero quiero creer que crecerá”.