

# Los desarrolladores de software libre

**Gregorio Robles**  
TU Berlín

**grex@scouts-es.org**

El desarrollo del software libre se ha consolidado en los últimos tiempos como una nueva forma de generar software de gran calidad. Sin embargo, mientras que su proceso de desarrollo se conoce con cierta exactitud y por ello ha podido ser refinando y mejorado en los últimos años, la información acerca de los desarrolladores que lo llevan a cabo es ciertamente escasa. En esta ponencia se mostrará la metodología y los resultados obtenidos en un estudio realizado sobre los desarrolladores de software libre en la Universidad Técnica de Berlín.

Copyright Gregorio Robles. Se otorga permiso para copiar, distribuir y/o modificar este documento bajo los términos de la Licencia de Documentación Libre GNU 1.1 o cualquier otra versión posterior publicada por la Free Software Foundation. Se considerará como Secciones Invariantes todo el documento, no habiendo Textos de Portada ni Textos de Contraportada. Puede consultar una copia de la licencia en: <http://www.gnu.org/copyleft/fdl.html>.

## 1. Introducción

En los últimos meses un grupo de investigación de la Universidad Técnica de Berlín ha estado realizando un estudio sobre aspectos demográficos, sociales, laborales y técnicos de los desarrolladores de software libre. Como fruto de ese trabajo, se publicó un documento del cual este trabajo es una adaptación para el IV Congreso HispaLinux.

Esta ponencia no sólo pretende mostrar el método y los resultados de este estudio, sino que además pretende dar una panorámica completa de las diferentes aproximaciones que se han hecho hasta el momento a la hora de abordar investigaciones acerca de los desarrolladores. Se presentarán varios métodos, poniendo especial énfasis en las cuatro fuentes y aplicaciones de las que se mostrarán los resultados. Estas cuatro fuentes serán, además, analizadas desde un punto de vista crítico. En todo caso, se remite al documento original (en inglés) a todo aquel que desee un tratamiento exhaustivo y pormenorizado de todas las cuestiones que se van a presentar.

## **2. Objetivos**

A la hora de abordar este estudio, debemos tener en cuenta que el desarrollo del software libre se ha venido llevando a cabo durante mucho tiempo sin los datos que estamos buscando. Si se quiere participar en un proyecto de software libre, nunca (o casi nunca) serán requeridos. En el desarrollo del software libre no es importante la edad, ni la nacionalidad, ni la profesión, ni siquiera se requiere una cualificación académica específica; si se quiere colaborar y se tienen los conocimientos necesarios (o se está en vías de adquirirlos), se está gratamente invitado a hacerlo sin más. Esto es bueno, y así debe seguir.

Sin embargo, aunque a primera vista no lo parezca, la obtención de los datos y su consiguiente procesamiento puede suponer importantes beneficios tanto para el desarrollo del propio software libre, como para el estudio de fenómenos sociales, económicos y políticos ligados al mismo. Veremos en los siguientes párrafos por qué.

Desde el punto de vista de ingeniería del software, cuanto más sepamos acerca de los desarrolladores, mejores modelos de desarrollo podremos hacer. Un modelo teórico cercano a la realidad permitirá la elaboración de métricas precisas que permitan predecir con relativa exactitud la velocidad de desarrollo y la calidad de las aplicaciones. Incluso podrían crearse métodos que indiquen los problemas existentes y las soluciones que se deben tomar. Saber más sobre los que desarrollan software libre, su entorno, sus condiciones laborales y su preparación es un primer paso para mejorar la calidad del mismo.

Nadie duda de que el software libre es hoy en día un bien de alto valor económico. En un estudio sobre la distribución RedHat 7.1 se ha calculado que una empresa privada necesitaría más de mil millones de euros y ocho mil personas-año para realizarla completamente desde el principio. Además, se observó cómo cada año casi se duplicaba el número de líneas de código (*More Than a Gigabuck: Estimating GNU/Linux's Size* (<http://www.dwheeler.com/sloc/>)). Por si esto fuera poco, si

tenemos en cuenta que en la actualidad el modelo de negocio alrededor del software libre se encuentra todavía en un estado más bien difuso, informaciones como las que proporciona este y otros estudios pueden suponer una buena base para afrontar esta problemática.

Además, hay mucha expectación sociológica. El paradigma de desarrollo del software libre, generalmente distribuido, con unos patrones y reglas ciertamente novedosos y sin una aparente complejidad organizativa ni burocrática, está siendo abordado con mucho interés por grupos de sociólogos y psicólogos para intentar, en primer lugar, comprenderlo y, a continuación, extrapolarlo a otros campos productivos y de las relaciones sociales.

Los resultados también pueden tener una lectura política: estamos hablando de naciones, de regiones, de culturas, de educación. A pesar de que no es el objetivo principal de este estudio, es inevitable hacer valoraciones y sacar conclusiones de este tipo.

Otro de los aspectos interesantes es acabar con tantas frases hechas y clichés sin fundamentos. Los desarrolladores han sido calificados desde niños (teenagers) pasando por comunistas, anarquistas o trabajadores frustrados. Un rápido repaso a los resultados es suficiente para ver cuantas afirmaciones han sido hechas sin ningún fundamento.

Por último, hay que decir que a los mismos desarrolladores de software libre nos interesa. Queremos saber cómo es la comunidad alrededor del software libre de la que formamos parte, qué tipo de personas la componemos, cuáles son las herramientas y entornos preferidos y un largo etcétera. Ejemplo de ello es la amplia aceptación que tuvo nuestro cuestionario.

Otro de los objetivos, aunque secundario, era poder realizar el estudio sobre desarrolladores de software libre con software libre o fuentes de libre acceso. Esto se debe a varias razones. La primera es obvia y se puede resumir como que hacemos lo que predicamos. La segunda es que permite que cualquiera pueda corroborar los resultados, aspecto muy importante, cuando menos desde el punto de vista investigador. Y la tercera es que tomando como base nuestro estudio, mejorando el software y viendo los problemas que reportamos, se pueden realizar estudios más exhaustivos y que alcancen una mayor precisión sin tener que reinventar la rueda a cada paso. Vemos, por tanto, que el software libre tiene una serie de características que lo hacen idóneo para este tipo de investigaciones.

### **3. Metodología**

El gran problema con el que nos encontramos al principio de este estudio es que no había ningún método conocido que nos llevara a resultados certeros. Es verdad que algunos estudios anteriores han intentado indagar un poco, cada uno a su manera, en los desarrollos del software libre, pero ninguno abarcaba el tema de forma global, tal y como nosotros queríamos.

Al no haber precedentes muy claros, decidimos tomar tantas fuentes como fuera posible para, a posteriori, entrar a valorar los diferentes resultados. La comparación y correlación entre los mismos nos diría qué grado de credibilidad debía tener cada fuente. A la vista de los resultados, todas las fuentes, aunque tengan desviaciones más o menos pronunciadas, dan unos resultados que convergen, lo que nos lleva a pensar que los resultados son muy próximos a la realidad.

El estudio se basó en cuatro fuentes diferentes que a su vez siguen cuatro métodos de extracción de datos diferentes:

1. el catálogo de aplicaciones de software libre SourceWell (<http://sourcewell.berlios.de>), al estilo Freshmeat, que hay en BerliOS (<http://www.berlios.de>),
2. la base de datos (<http://db.debian.org>) de los desarrolladores de una conocida distribución de GNU/Linux (Debian (<http://www.debian.org>))
3. una herramienta llamada CODD (<http://codd.berlios.de>), que analiza el código fuente automáticamente y extrae los autores del mismo,
4. y un formulario en línea conocido con el nombre de Widi (<http://widi.berlios.de>) (que viene de "Who Is Doing It?", "¿Quién lo está haciendo?").

Todas estas fuentes serán presentadas y analizadas pormenorizadamente en el siguiente apartado. Antes, vamos a ver otros métodos y fuentes que han sido utilizados en estudios anteriores.

En 1999 el UNC Open Source Research Team (<http://metalab.unc.edu/osrt>) estudió el repositorio de aplicaciones de software libre en el MetaLab (<http://metalab.unc.edu/>). Este estudio se basa en el uso del Linux Software Map (LSM, mapa del software Linux), una especie de estándar para dar información sobre los paquetes que se depositan en un repositorio. Uno de los apartados que es de obligado cumplimiento en el LSM es indicar el nombre y dirección de correo-e del desarrollador principal. Los autores del estudio se dedicaron a analizar los archivos LSM de cada paquete, extrayendo el dominio de correo electrónico de los autores y tratando estadísticamente los resultados. En este estudio no hemos incluido ningún método que utilice el LSM, ya

que es un método poco arraigado. Su uso en el repositorio MetaLab se debe más bien a razones históricas: hace una década, cuando el ancho de banda era escaso, el LSM proporcionaba importante información acerca de la funcionalidad de los paquetes antes de descargárselos. El auge de las distribuciones de GNU/Linux y del CD-ROM han hecho que el LSM haya pasado a un segundo plano.

Una forma más completa de abordar el tema es mediante la aplicación de la Trinity Participation Metric (<http://www.cse.ucsc.edu/~alison/projects/cmpe276/index.html>). Se trata de una métrica que mide la participación en proyectos de software libre que se basa en estudiar no sólo el código fuente, sino también mensajes a la lista de correo principal del proyecto y los parches enviados. A la hora de sumar las tres fuentes, cada una es multiplicada por un peso según su importancia relativa. Esta métrica implicaría un gran esfuerzo si se quieren estudiar muchos proyectos, por lo que es más bien aplicable a proyectos específicos. En el caso del documento original, estudian la participación en The GIMP.

Desde hace pocos meses, la Fundación del Software Libre ha comenzado un proyecto (<http://france.fsfeurope.org/coposys/index.html>) que representa de manera visual la procedencia de los desarrolladores de software libre. Los desarrolladores tienen que registrarse antes de rellenar un formulario con sus coordenadas. Los resultados se muestran con pequeños retardos en un mapa del mundo en varios tamaños.

Por último, y ya más bien desde el punto de vista sociológico, nos tenemos que hacer eco del Linux Study (<http://www.psychologie.uni-kiel.de/linux-study/>), un análisis sociotécnico realizado sobre los desarrolladores del núcleo Linux. Este estudio se basaba en un cuestionario que fue rellenado por cerca de 150 desarrolladores.

## 4. SourceWell

SourceWell (<http://sourcewell.berlios.de>) es un índice de aplicaciones de software libre al estilo del popular Freshmeat (<http://www.freshmeat.net>). Tiene una base de datos con más de mil aplicaciones en las que además del nombre de la aplicación, su número de versión, una breve descripción y las referencias para descargárselo, contiene el nombre y la dirección de correo del autor principal de la aplicación (o en su defecto la persona que la mantiene).

Los puntos flacos de SourceWell se pueden resumir en los siguientes: al tratar solamente con un autor por aplicación, no se tiene en cuenta el grado de accesibilidad del software libre que permite participar al que así lo desee en el desarrollo de una aplicación.

En esta fuente se clasifica la nacionalidad según el dominio de correo electrónico, una afirmación que si de por sí no tiene que ser cierta, empeora al constatar que la existencia de dominios genéricos como .com, .org, .net, etc., constituye más del 60% de los casos.

Debemos también tener en cuenta que la cantidad de aplicaciones registradas superan ligeramente el millar, cifra que en comparación con los cuarenta mil proyectos que se calcula que hay es bastante pequeña.

Veremos que aún con todos estos inconvenientes, los resultados ofrecidos por SourceWell no difieren en demasía de los de las otras fuentes, pudiéndose observar ciertas peculiaridades que se comentarán más adelante. Podemos entender que la estadística, al fin y al cabo, lo que hace es contrarrestar los diferentes errores compensándolos unos con otros.

## **5. Debian**

Debian es una de las distribuciones GNU/Linux más populares. Además, como se puede ver en los resultados de otra de las fuentes de este estudio, es la preferida entre los propios desarrolladores de software libre. Una de las características más notorias de Debian es que se trata de una distribución no comercial que es gestionada, desarrollada y mantenida por un nutrido grupo de voluntarios. En la actualidad, Debian cuenta con más de 700 desarrolladores registrados en esta base de datos distribuidos a lo largo y ancho de todo el planeta.

En realidad, la base de datos de Debian no contiene la nacionalidad de los desarrolladores, sino su país de residencia (esto es así para permitir reuniones locales o poder localizar desarrolladores cuando se viaja). Cabe comentar que el proceso de admisión que hay que llevar a cabo para ser desarrollador Debian puede provocar efectos en red locales, de forma que la probabilidad de ser desarrollador Debian en una región con muchos desarrolladores per cápita sea superior al de otra región donde Debian sea menos popular. Este efecto no se tiene al participar en otros proyectos de software libre, sobre todo si partimos de la idea de que una forma de colaborar en un proyecto es enviando un parche o realizando una traducción.

## **6. Codd**

Codd es una aplicación que rastrea el código fuente de los programas y trata de asignar cuotas de código a los programadores que figuran en las cabeceras de los ficheros. Codd fue desarrollado por Vipul Ved Prakash después tras leer un artículo de Rishab

Aiyer Ghosh (Cooking pot markets: an economic model for the trade in free goods and services on the Internet ([http://www.firstmonday/issues/issue3\\_3/ghosh/index.html](http://www.firstmonday/issues/issue3_3/ghosh/index.html))) en el cual se afirma que el software libre forma parte de una economía virtual en la que el intercambio de ideas pasa a un primer plano en detrimento de las transacciones monetarias.

El gran problema de Codd es que no hay ningún procedimiento ampliamente aceptado para denotar la autoría de código. También aparece el problema que teníamos con SourceWell al intentar extraer la nacionalidad de los desarrolladores por medio del dominio de su correo electrónico. Para solventar este problema, se creó una base de datos con la nacionalidad de los desarrolladores más participativos y con dominio de correo-e genérico. Aún así, por desgracia, la proporción de código que no ha podido ser asignado es grande.

## **7. Widi**

Widi es la parte del estudio en la que pedimos directamente a los desarrolladores que nos cuenten más sobre ellos. Se basa en un cuestionario en Internet con preguntas de tres ámbitos diferentes: datos personales, datos profesionales y cuestiones relacionadas con la experiencia y habilidades informáticas.

En el apartado de datos personales preguntamos por el apodo, la nacionalidad, el país de residencia, la edad, el dominio de su correo-e y los idiomas que dominan los desarrolladores.

El siguiente apartado corresponde al de los datos profesionales. En él se incluyen la profesión, la cualificación académica, si reciben contraprestación económica por desarrollar software libre y una valoración de las perspectivas sociales, económicas y políticas en su país acerca del software libre. Además, se tratan aspectos ligados con la satisfacción laboral, como son la relación con sus superiores y la afinidad al puesto laboral que desempeñan en la actualidad. Se termina preguntando el poder adquisitivo y las horas dedicadas al desarrollo de software libre a la semana.

El tercer bloque comprende cuestiones relacionadas con la experiencia y las habilidades informáticas. Aparte del número de proyectos en los que se está involucrado y de los lenguajes de programación y herramientas conocidas, se indaga en aspectos más puntuales como editor, escritorio o sistema operativo favorito. Se termina escuetamente preguntando si se prefiere el término "Open Source" o "Free Software", una de las preguntas que, lejos de nuestros principales propósitos, ha generado la mayor polémica hasta el punto de que en el documento final del estudio nos hayamos decantado por el término Libre Software (<http://libre.act-europe.fr/>).

Para promocionar Widi, el equipo de investigación realizó notas de prensa en varios idiomas con el fin de poder anunciar la encuesta en las bitácoras sobre software libre más populares del planeta y conseguir una participación lo más amplia posible. Por desgracia, en algunos casos, como con la baja participación japonesa, podemos ver que nuestro propósito no fue muy exitoso. Además, Widi plantea la problemática de no saber hasta qué punto podemos considerar los datos recogidos por medio de una encuesta como válidos. A la vista de los resultados y de las correlaciones con las otras fuentes, podemos estar bastante satisfechos.

## **8. Resultados**

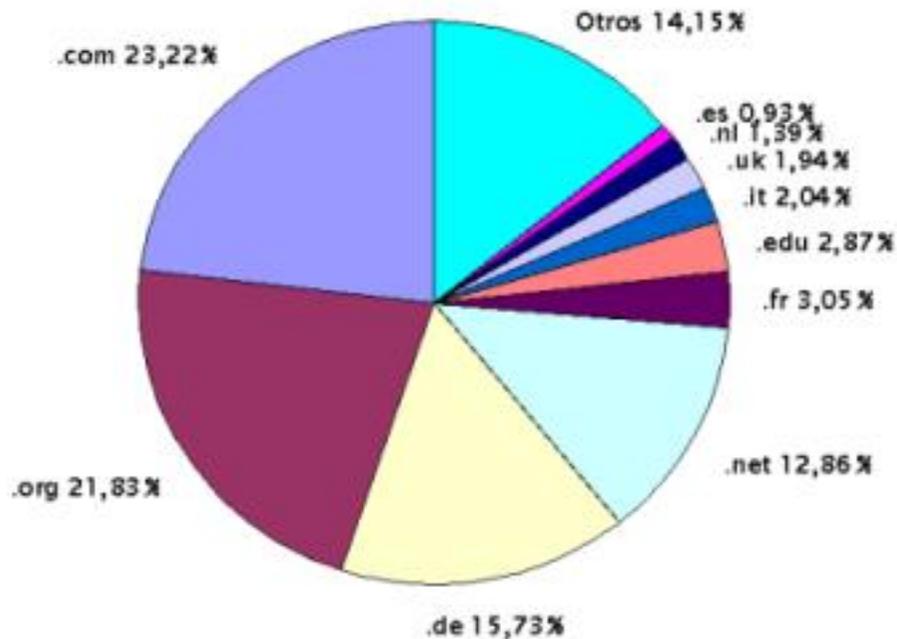
En el siguiente capítulo se tratarán los resultados obtenidos por las diferentes fuentes analizadas en apartados anteriores. Las gráficas serán comentadas de manera escueta, haciendo especial hincapié en los datos más relevantes o que difieran entre las diferentes fuentes.

## **9. SourceWell**

Los datos que se van a mostrar a en el siguiente apartado corresponden a una muestra de las aplicaciones de software libre contenidas en SourceWell el 20 de agosto de 2001. Ese día la base de datos de SourceWell contenía 1136 aplicaciones. En las estadísticas en tiempo real (<http://sourcewell.berlios.de/stats.php3>) se pueden ver estadísticas de SourceWell actualizadas.

La primera gráfica es la que se obtiene al extraer los dominios de correo electrónico de los desarrolladores las aplicaciones (o personas que las mantienen). Se puede observar que predominan los dominios denominados como genéricos (gTLD, generic Top Level Domain) como el .com, el .org, el .net y el .edu. Hemos supuesto que los dominios vinculados a los países (denominados en la jerga ccTLD por country code Top Level Domains) suelen ser mayoritariamente usados por personas de dicha nacionalidad.

Figura 1. [Gráfica: Dominios más frecuentes en SourceWell]



Es interesante destacar que esta gráfica tiene una distribución muy similar a la que nos dará más adelante el cuestionario en línea. Hay algunas variaciones que se pueden achacar a que en los índices de aplicaciones aparecen frecuentemente dominios del tipo sourceforge.net (54 veces) o del clon de SourceForge que hay en BerliOS (berlios.de aparece 24 veces). También hay que tener en cuenta que el dominio .org aparece en multitud de aplicaciones, básicamente porque con frecuencia se dan direcciones genéricas al proyecto o de listas de correo. Tal es el caso para kde.org (39 veces), gnu.org (30 veces) o gnome.org (10 veces).

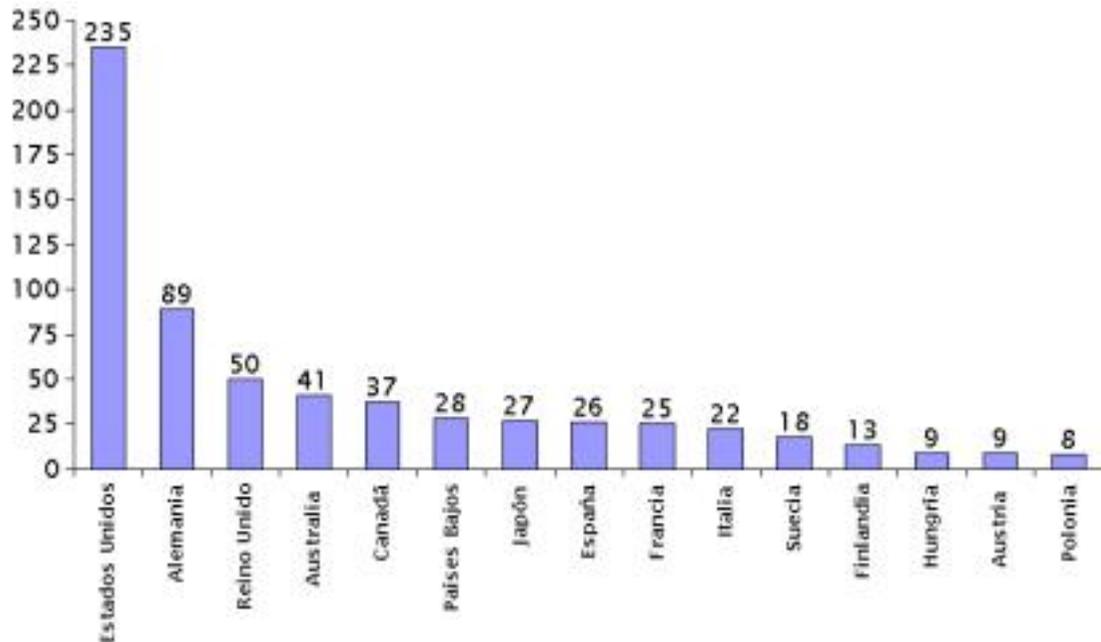
## 10. Debian

Las datos que vamos a mostrar a continuación fueron extraídos de la base de datos de Debian el 29 de junio de 2001. Entonces, estaban registrados 706 desarrolladores de 41 países diferentes.

La primera gráfica, casi obligada, es la distribución de desarrolladores por países de residencia. Podemos observar (no sin cierta satisfacción) que España ocupa el séptimo

lugar, un puesto que nos coloca a la cabeza de los países involucrados en el desarrollo de Debian.

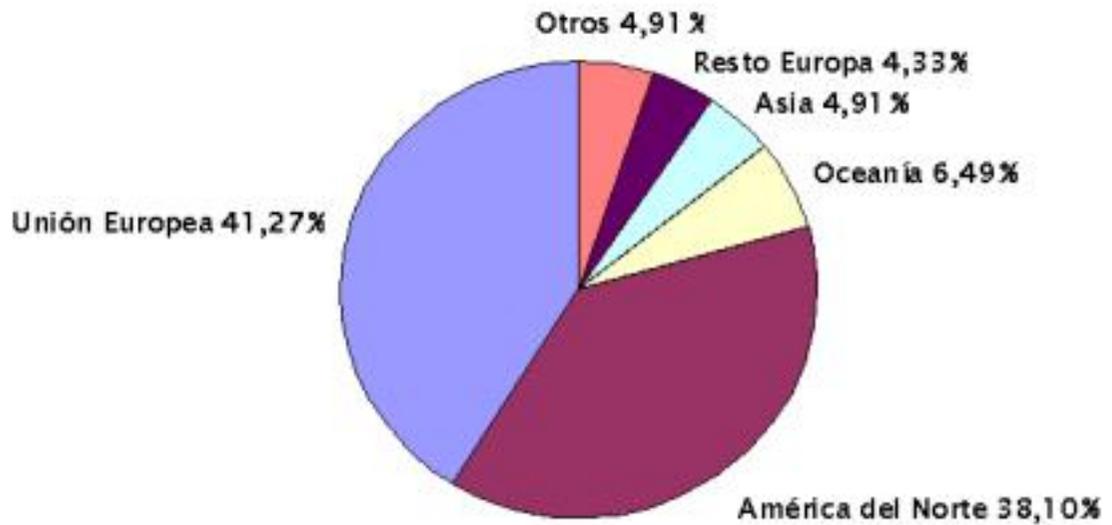
**Figura 2. [Gráfica: Distribución de desarrolladores Debian por país de residencia. 706 desarrolladores Debian]**



Mientras que en el capítulo por países de residencia, los Estados Unidos llevan claramente la delantera, si tenemos en cuenta la distribución por continentes, tendremos una imagen totalmente distinta. Cabe destacar que hemos dividido el continente europeo en naciones pertenecientes a la Unión Europea y naciones que no pertenecen a la misma para resaltar todavía más este aspecto. Gracias a esta distinción, observamos que no sólo Europa tiene un mayor número de desarrolladores Debian que Norteamérica. La Unión Europea es la que contribuye con un mayor número de desarrolladores al proyecto.

**Figura 3. [Gráfica: Distribución de desarrolladores Debian por continentes (de**

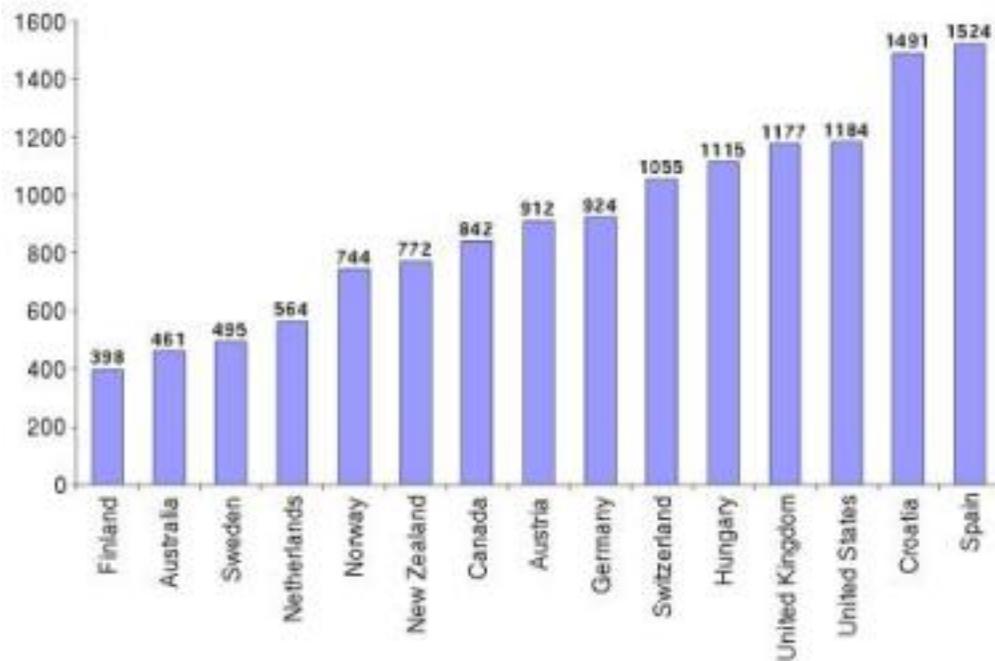
residencia). 706 desarrolladores Debian]



No sólo es interesante el número de desarrolladores que residen en cada país, sino que estaría bien ver la proporción de desarrolladores Debian que hay por habitante, algo así como una tasa de desarrolladores Debian per cápita. Para ello, se ha tomado el número de habitantes de cada país de los datos que ofrece la ONU (<http://www.un.org/popin/wdtrends/pop1999-00.pdf>) y se ha dividido entre el número de desarrolladores Debian. Vemos que allí donde esta tasa es mayor (y, por tanto, el número de habitantes por cada desarrollador Debian es menor) es en la zona nor- y centroeuropea, así como en Australia y Nueva Zelanda. España, aunque no mal situada, ya se va quedando muy atrás, incluso por debajo de la media de la Unión Europea (1 desarrollador cada 1.226.000 ciudadanos). Para alcanzar el mismo número de desarrolladores per cápita que hay en Finlandia, en España deberíamos llegar al centenar.

**Figura 4. [Gráfica: Países con menor número de habitantes por cada**

desarrollador Debian. 706 desarrolladores Debian]



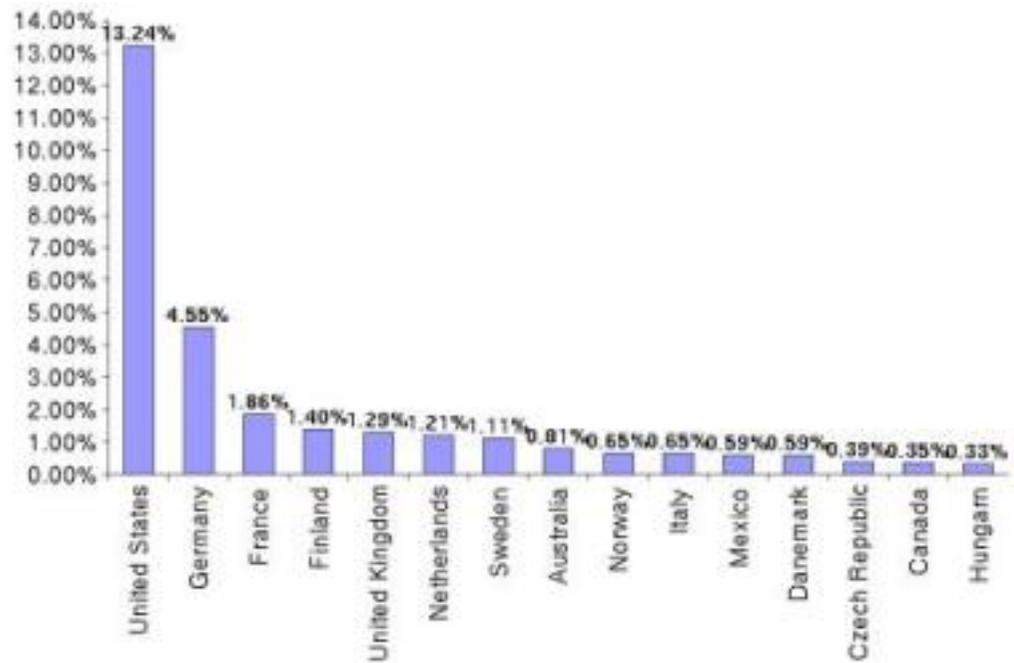
Para los más interesados, me gustaría recomendar la visita del mapa (<http://www.debian.org/devel/developers.loc>) en las páginas de Debian que muestra el lugar de residencia de los desarrolladores. Basándose en estos datos, un miembro de Debian ha calculado los centros de masa (<http://people.debian.org/~edward/average/>) mundial (que viene a estar al sur de Groenlandia), europeo (por Luxemburgo) y norteamericano (por Kansas) de los desarrolladores Debian.

## 11. Codd

Codd ha sido utilizado como herramienta para rastrear código fuente de una distribución GNU/Linux (en este caso Mandrake). En total, el código fuente analizado asciende a más de un gigabyte, un volumen nada despreciable.

En la primera gráfica, se puede ver la contribución por nacionalidades. El factor desconocido (código que no puede ser asignado o que ha sido asignado a una organización plurinacional, como es el caso de la Free Software Foundation) es muy grande, llegando a ser ligeramente superior al 60%.

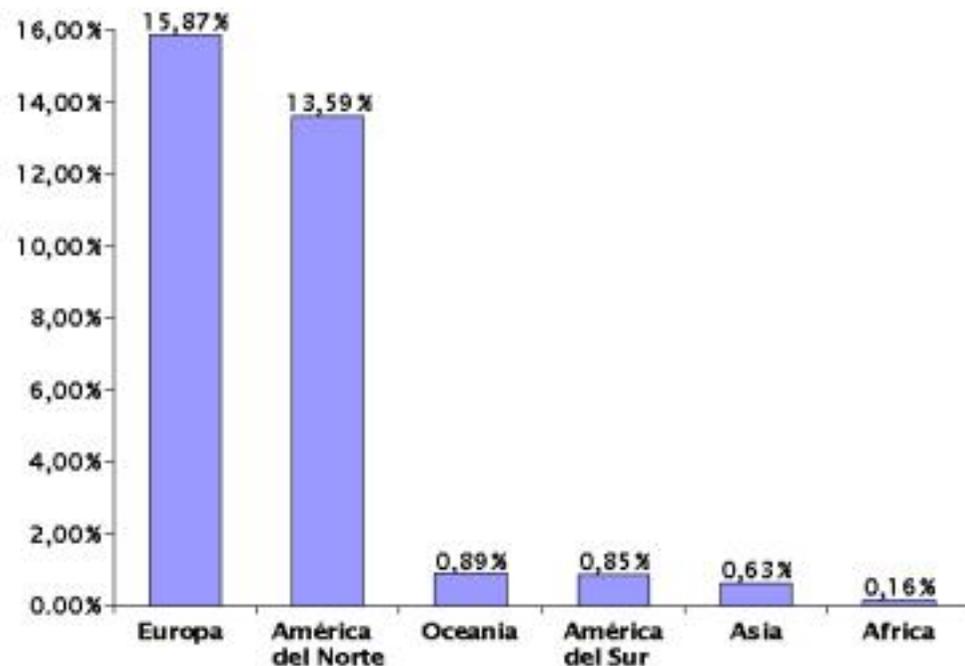
**Figura 5. [Gráfica: Contribución según nacionalidad según Codd. Volumen: +1067619717 bytes]**



Si sumamos la contribución por continentes, vemos que Europa se vuelve a situar por delante de Norteamérica.

**Figura 6. [Gráfica: Contribución según continentes según Codd. Volumen: +1067619717 bytes]**

+1067619717 bytes]



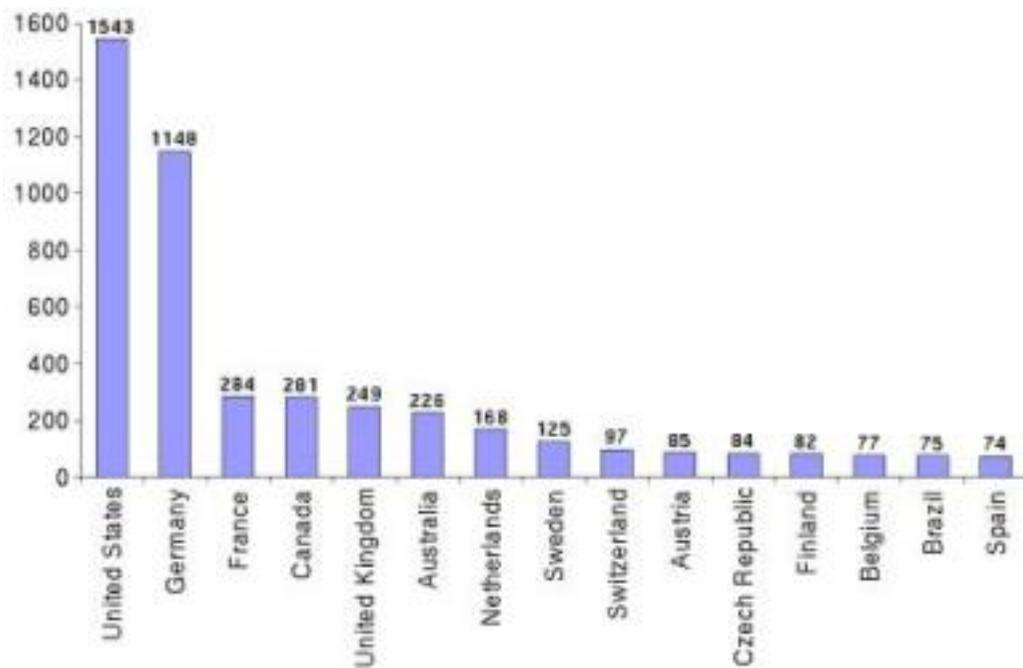
## 12. Widi

Widi fue abierto al público el 24 de junio de 2001 y se cerró el 14 de agosto tras contar con una participación superior a los cinco mil desarrolladores. Entre estas dos fechas, padecimos dos "efectos Slashdot", en particular al anunciarlo en heise.de, probablemente la bitácora más popular del ámbito germano, y en Slashdot, con seguridad la más visitada a nivel mundial.

Nota: al pie de cada gráfica, se indicará en primer lugar el número de desarrolladores que ha respondido a la pregunta sobre el número total de desarrolladores que ha rellenado el formulario. Los dos números difieren, ya que todas las respuestas eran voluntarias.

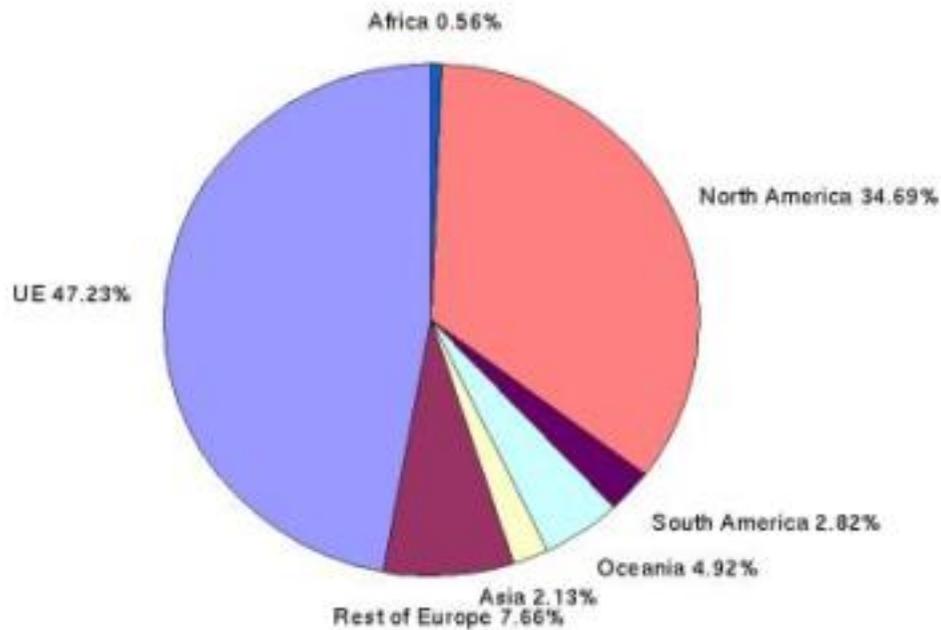
En la primera gráfica se puede ver la distribución de nacionalidades en Widi. En comparación con las otras fuentes del estudio sólo cabe resaltar la buena posición de Canadá, muy por encima de los otros resultados. España ocupa la decimoquinta posición.

Figura 7. [Gráfica: Nacionalidades en Widi. 5391 respuestas de 5478 desarrolladores]



Como en los casos anteriores, vamos a agrupar los países por continentes para ver el peso de cada uno. Se puede observar una gran participación europea, muy superior a la norteamericana. Widi, al igual que las otras fuentes, también certifica que la participación de desarrolladores provenientes de los continentes africano, asiático y sudamericano es bastante pobre.

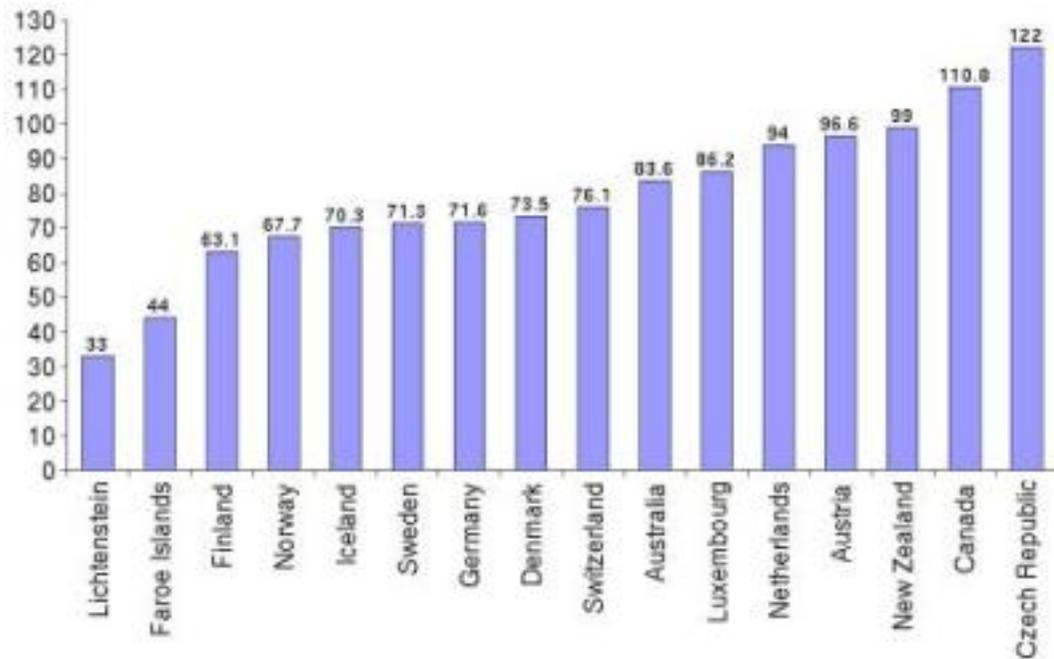
Figura 8. [Gráfica: Continentes en Widi. 5391 respuestas de 5478 desarrolladores]



Al igual que con Debian, con los datos de Widi hemos intentado ver la penetración del software libre en los diferentes países tomando como medida la tasa de desarrolladores per cápita. Es importante destacar que mientras en el caso de Debian manejábamos valores absolutos (el número total de desarrolladores Debian nos era conocido), en este caso vamos a tratar con valores relativos. Se puede ver que aún así los países y regiones con mayor número de desarrolladores per cápita sigue siendo prácticamente la misma.

Figura 9. [Gráfica: Desarrolladores (relativos) per cápita en Widi. 5391 respuestas

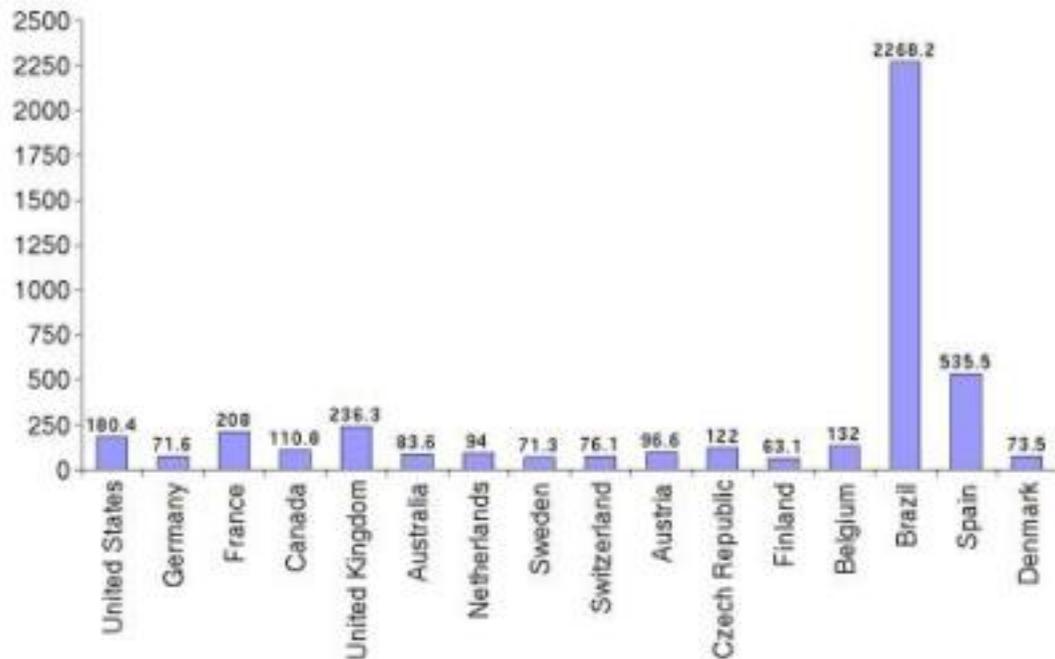
de 5478 desarrolladores]



En la gráfica anterior, disponemos de un ranking de países según la tasa de desarrolladores per cápita. Sin embargo, es interesante ver esa tasa para los países con mayor número de desarrolladores. Como resultado se puede observar que en España, a pesar de estar entre las naciones con mayor número de desarrolladores, la tasa de desarrolladores per cápita es bastante pobre en comparación con los países punteros.

**Figura 10. [Gráfica: Desarrolladores per cápita - Países con mayor número de**

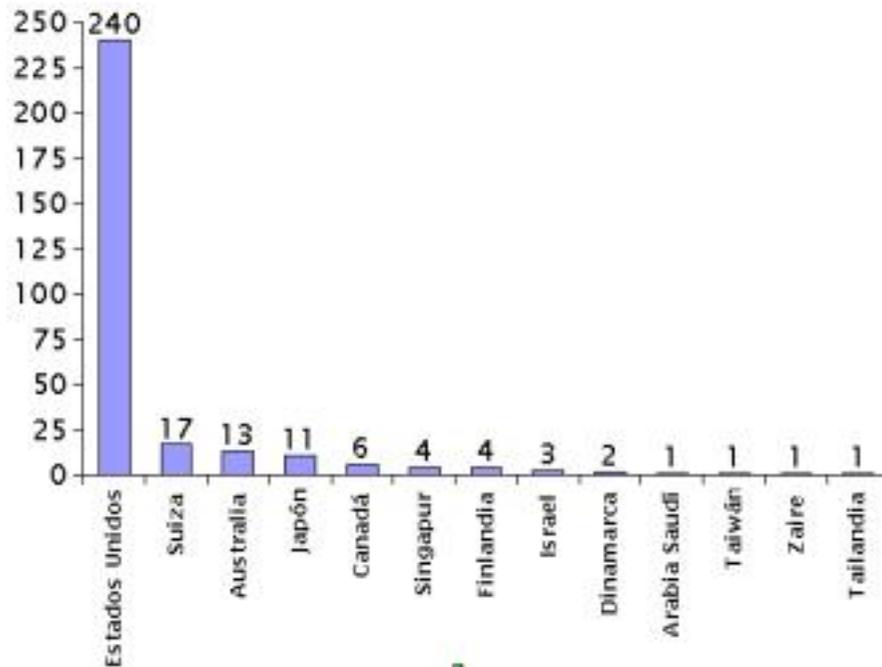
desarrolladores en Widi. 5391 respuestas de 5478 desarrolladores]



A partir de los datos de nacionalidad y país de residencia podemos obtener datos acerca de los flujos migratorios. Los flujos migratorios los definimos como el resultado de los inmigrantes (desarrolladores de otros países que residen en el país) menos los emigrantes (desarrolladores del país que residen en el extranjero). En la siguiente gráfica, la de los países con un balance de migración positivo, podemos ver que este flujo es claramente favorable para los Estados Unidos y unos pocos países más (aunque es cierto que muchos de estos países están por debajo del umbral de incertidumbre; téngase en cuenta que estamos hablando de un total de más de cinco mil desarrolladores y que el margen de error puede considerarse con creces superior al de un par de desarrolladores).

**Figura 11. [Gráfica: Países con flujo migratorio positivo en Widi. 5342 respuestas**

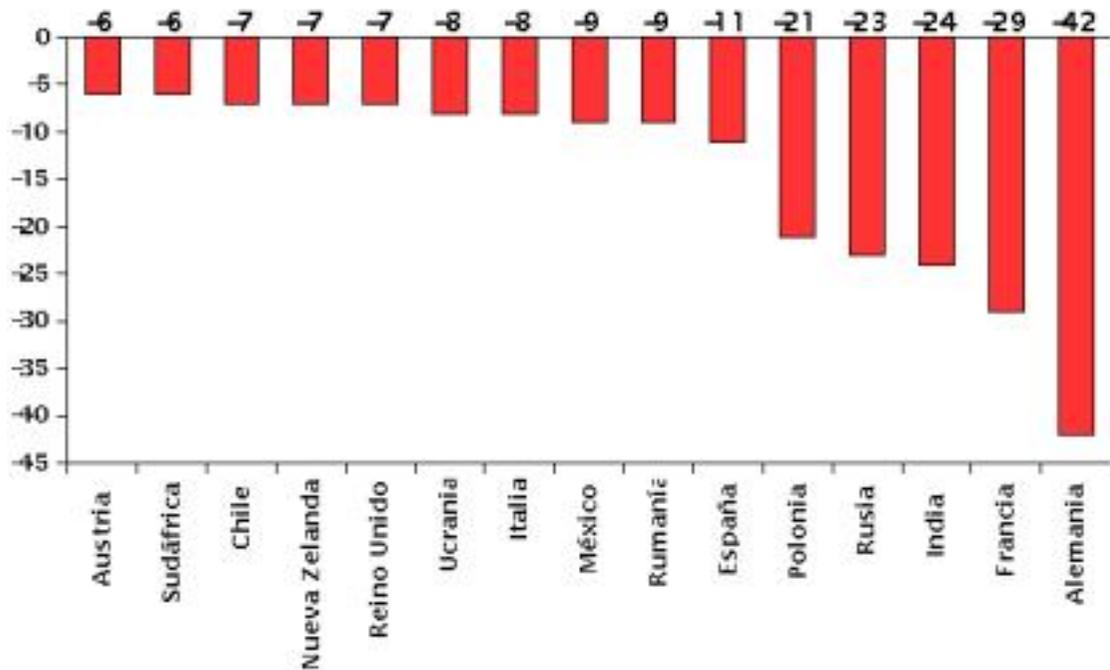
de 5478 desarrolladores]



En el lado opuesto, los países que acogen en menor medida a desarrolladores que los que ven partir, se encuentran principalmente países europeos y la India. España alcanza en esta clasificación (más bien negativa, todo sea dicho de paso) su mejor posición a lo largo del todo el estudio, la sexta plaza.

**Figura 12. [Gráfica: Países con flujo migratorio negativo en Widi. 5342 respuestas**

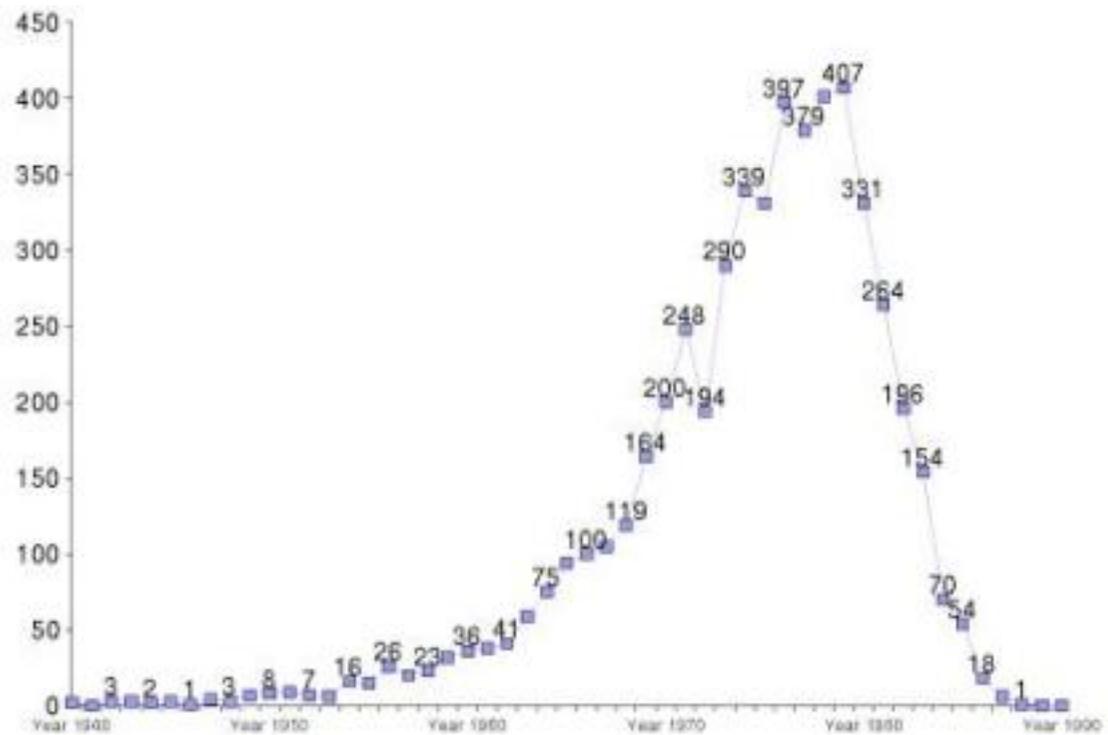
de 5478 desarrolladores]



El año de nacimiento que aparece con mayor frecuencia es el correspondiente a los desarrolladores nacidos en 1979. La media, sin embargo, se sitúa en 1974.1, por lo que el desarrollador medio a día de hoy tiene 27 años de edad y va camino de los 28. Puede verse cómo uno de los clichés más arraigados, el que dice que el software libre lo realizan niños, no concuerda mucho con la realidad. Aún así, hay más desarrolladores nacidos en la década de los 80 (un 20%) que en la de los 60 (un 16%).

**Figura 13. [Gráfica: Año de nacimiento en Widi. 5326 respuestas de 5478**

desarrolladores]



La distribución de sexos, que se muestra a continuación, es probablemente uno de los aspectos del software libre menos desconocidos. Existen desarrolladoras, pero su número es, en comparación con el de desarrolladores de sexo masculino, muy pequeño.

**Figura 14. [Gráfica: Distribución de sexo en Widi. 5272 respuestas de 5478**

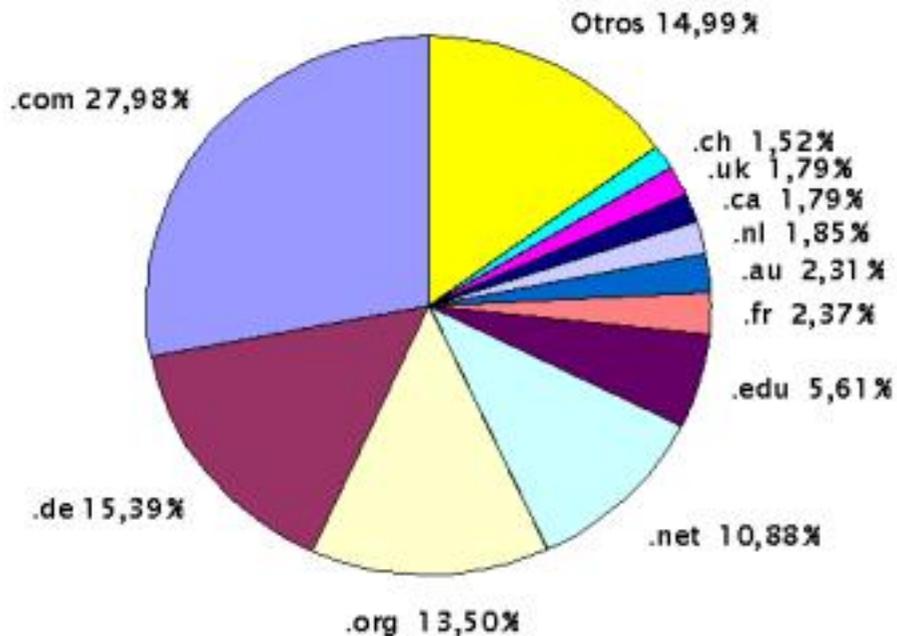
desarrolladores]



En Widi también preguntábamos por el dominio de correo electrónico de los desarrolladores. Nos era interesante obtener estos datos con la finalidad de poder compararlo con los obtenidos con anterioridad en SourceWell y en Codd.

**Figura 15. [Gráfica: Distribución de dominios de correo-e en Widi. 5185**

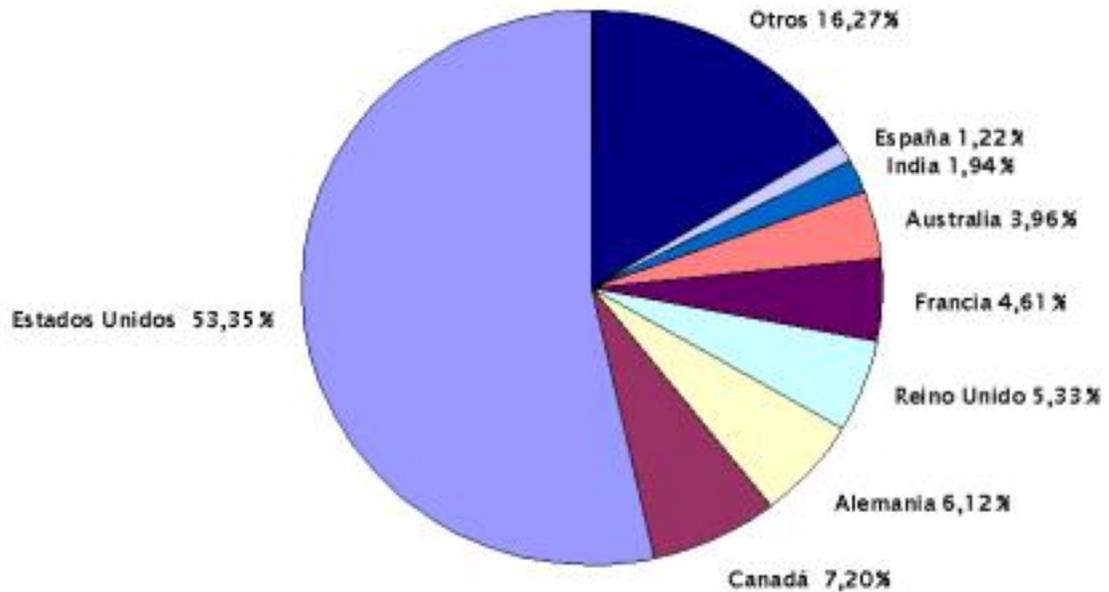
respuestas de 5478 desarrolladores]



Un aspecto que podemos evaluar gracias a la correlación de diferentes preguntas es ver la distribución de los dominios en cuanto a la nacionalidad de sus propietarios. Así, se puede ver que el dominio .com es de ciudadanos estadounidenses en igual número que de ciudadanos del resto de países del mundo. En el estudio original, se ha realizado esta correlación con todos los dominios genéricos y algunos específicos de país. Los datos de los dominios genéricos se pueden utilizar en SourceWell y Codd para asignar de manera estadística nacionalidades a estos dominios, mientras que los específicos de país permiten corroborar que la suposición de corresponder el dominio del país con una persona de esa nacionalidad es bastante exacta (superior al 95%).

**Figura 16. [Gráfica: Distribución por nacionalidades del dominio .com en Widi.**

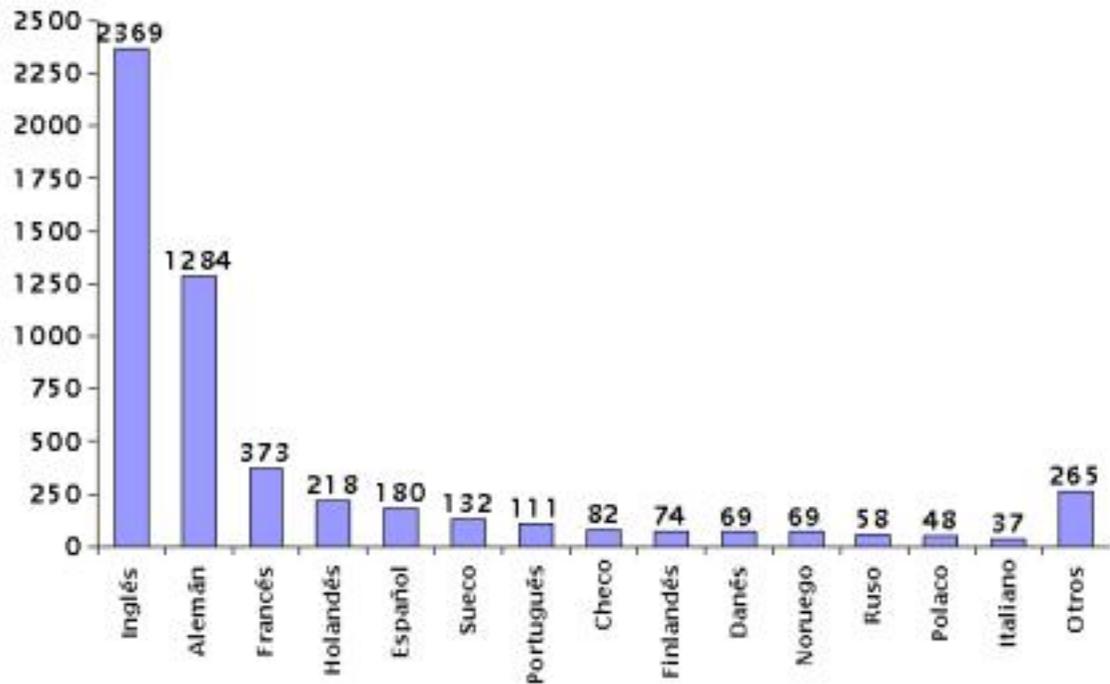
1433 respuestas de 5185 desarrolladores]



A continuación mostramos la distribución de las lenguas que los desarrolladores han indicado como sus lenguas maternas. El inglés es la lengua materna de alrededor del 50% de los desarrolladores de software libre. Además, se puede ver que gracias a la numerosa comunidad holandesa, el holandés se sitúa por delante del español.

**Figura 17. [Gráfica: Lengua materna en Widi. 5369 respuestas de 5478**

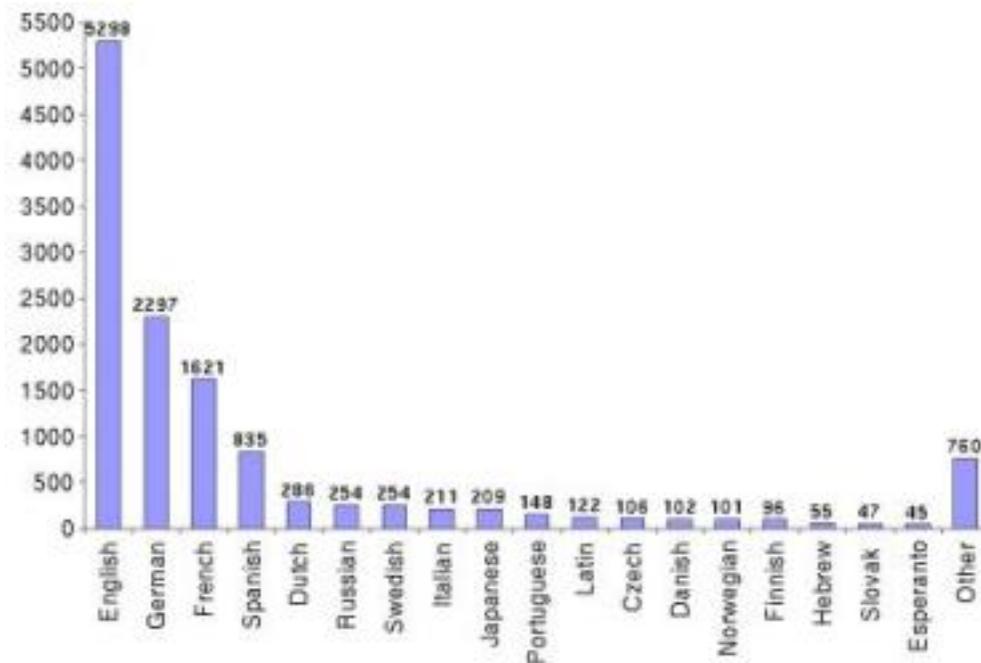
desarrolladores]



En la siguiente gráfica cabe resaltar que el dominio del inglés está muy arraigado, más que nada por que es casi una necesidad. El 97% de los encuestados declaran tener conocimiento de dicha lengua, situándose las demás a gran distancia. La comunidad de habla hispana es la cuarta más numerosa después de la inglesa, la alemana y la francesa.

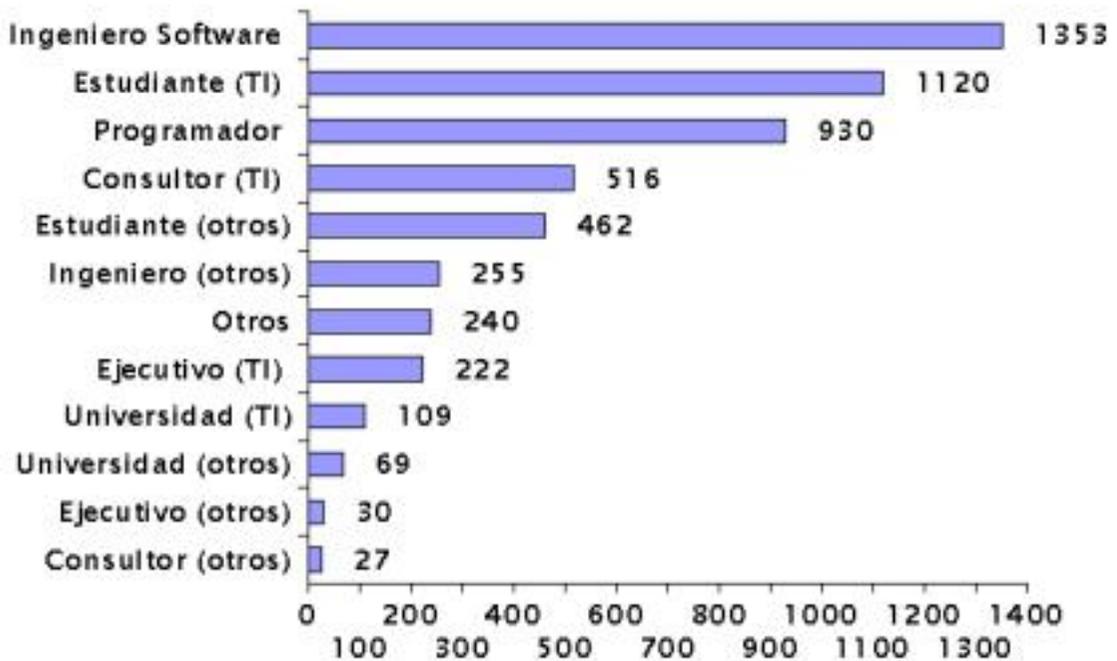
**Figura 18. [Gráfica: Lenguas que hablan los desarrolladores en Widi. 12847**

respuestas de 5478 desarrolladores]



Entramos ya en el campo laboral. La primera cuestión trata con la profesión. Nótese que hemos dividido las profesiones en dos áreas diferentes: el de las tecnologías de la información y el resto. Es interesante ver que la profesión más frecuente es la de ingeniero software. Los programadores, a su vez, se sitúan por detrás de los estudiantes de carreras de relacionadas con las tecnologías de la información, pero también muy por delante del resto de profesiones.

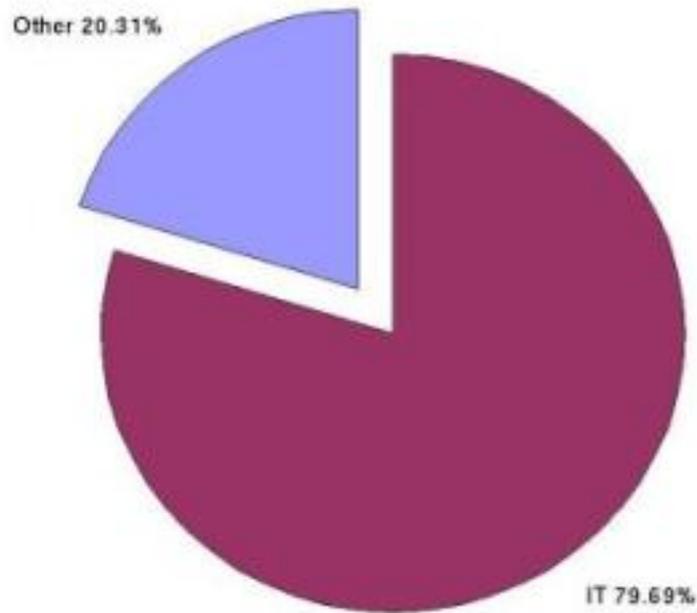
Figura 19. [Gráfica: Profesión en Widi. 5333 respuestas de 5478 desarrolladores]



Al haber dividido las profesiones en dos tipos de áreas diferentes, podemos ver el grado de participación de sectores ajenos al de las tecnologías de la información. Las estadísticas nos cuentan que uno de cada cinco no tiene una profesión relacionada con las tecnologías de la información.

Figura 20. [Gráfica: Distribución por sectores en Widi. 5333 respuestas de 5478]

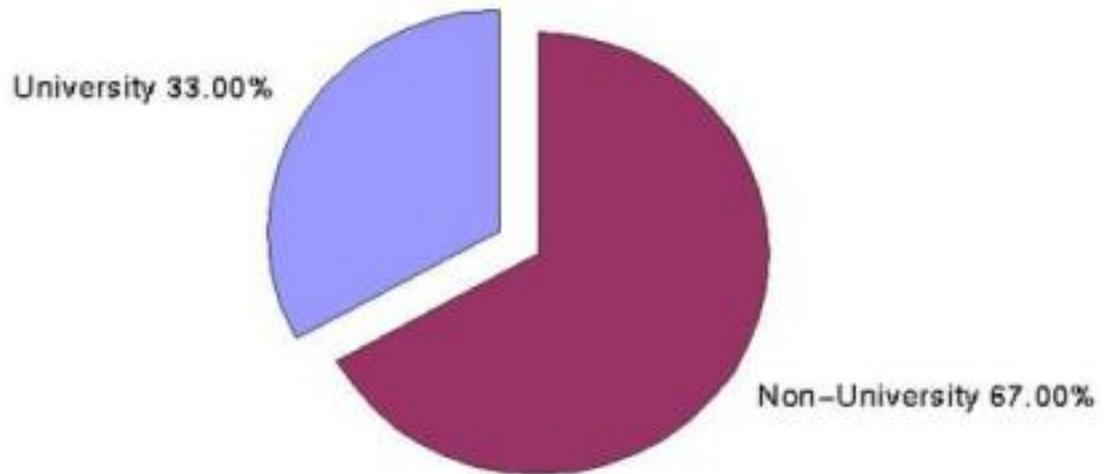
desarrolladores]



Otro aspecto de interés es ver la vinculación de las universidades con el software libre. Históricamente, muchas universidades han aportado de manera palpable al desarrollo de herramientas y proyectos. Podemos ver que la proporción de estudiantes y profesores universitarios involucrados es relativamente alta.

**Figura 21. [Gráfica: Profesión relacionada con la universidad en Widi. 5333**

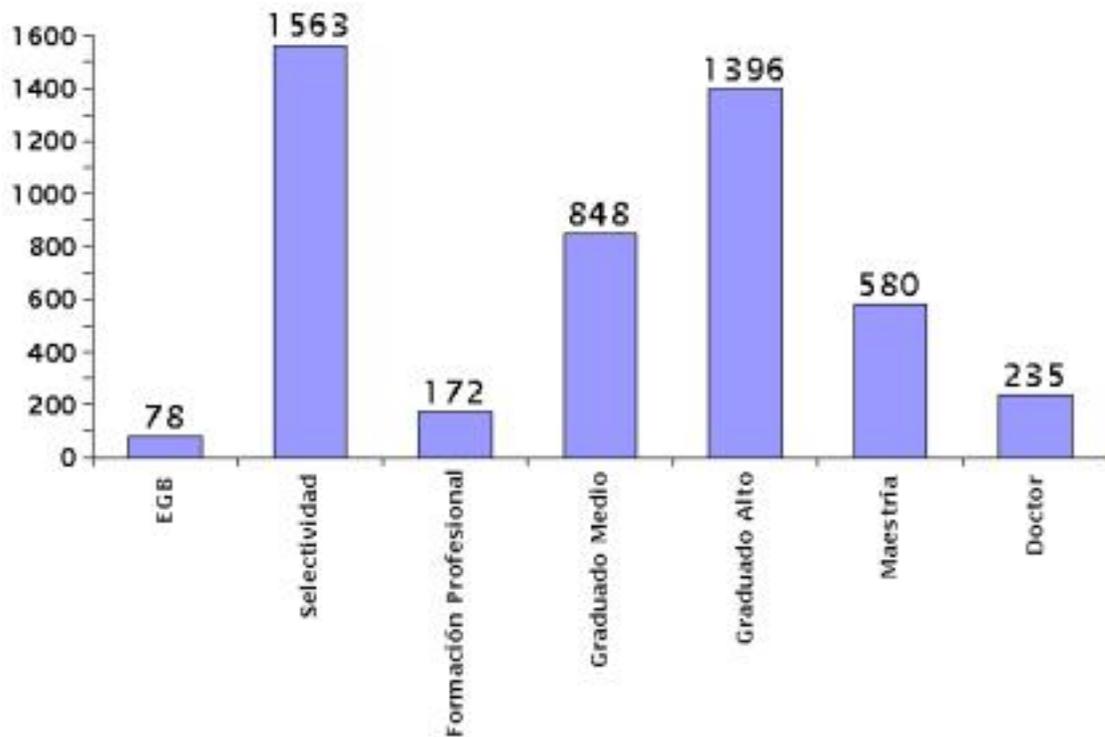
respuestas de 5478 desarrolladores]



La distribución de la cualificación académica que poseen los desarrolladores arroja resultados dispares. De todas formas, si tenemos en cuenta que más de un tercio de los encuestados es menor de 25 años y que en su mayoría son estudiantes, podemos interpretar que una gran mayoría de desarrolladores o tienen un título universitario o están en vías de conseguirlo.

**Figura 22. [Gráfica: Cualificación académica en Widi. 4872 respuestas de 5478**

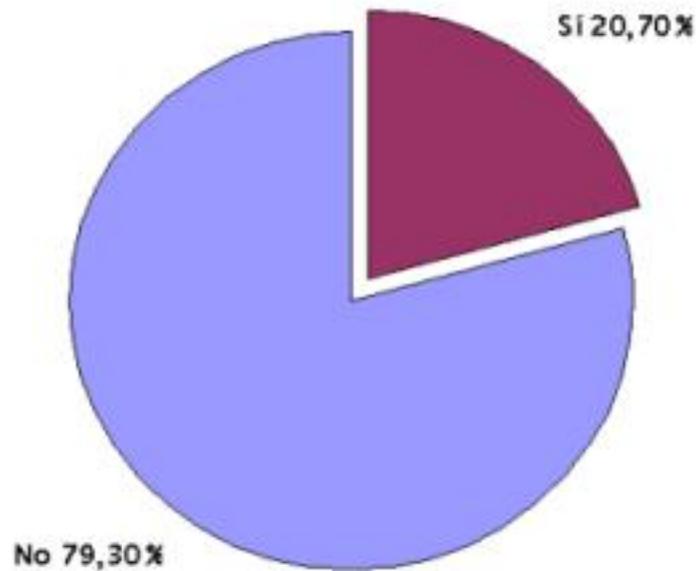
desarrolladores]



El porcentaje de desarrolladores de software libre que respondieron afirmativamente a si han sido pagados o les pagan por desarrollar software libre supera ligeramente el 20%. Los beneficios profesionales y económicos que los desarrolladores obtienen de manera directa gracias a su implicación en el software libre están, por tanto, limitados a una minoría de ellos.

**Figura 23. [Gráfica: ¿Te pagan o han pagado por desarrollar software libre?**

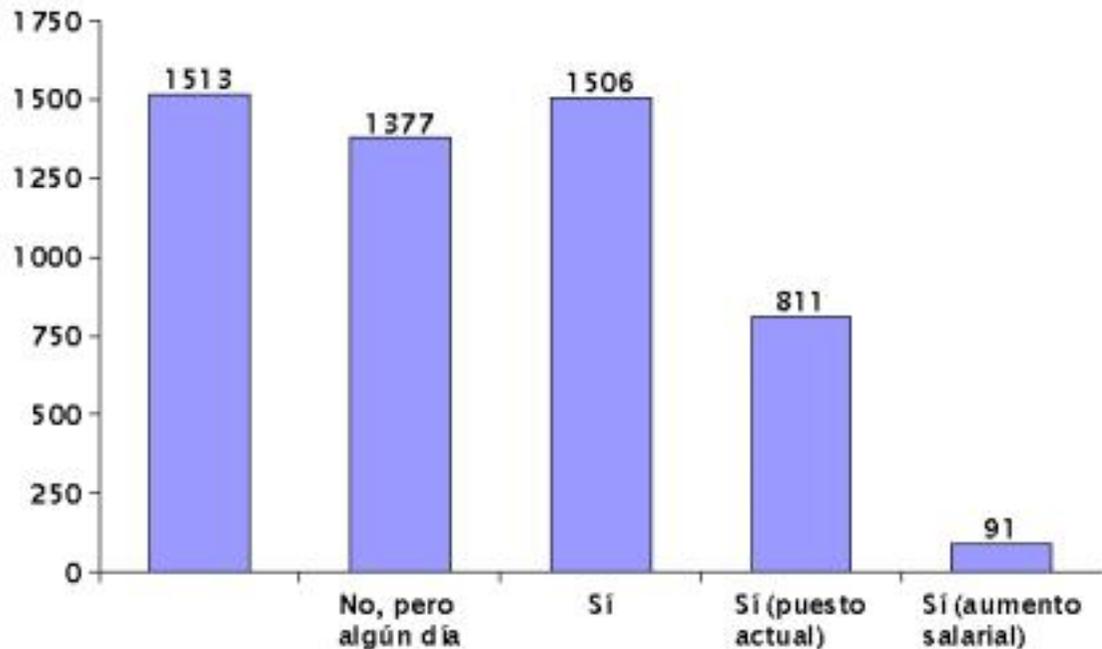
(Widi). 5015 respuestas de 5478 desarrolladores]



Ante la pregunta de si se han beneficiado indirectamente profesional o económicamente por colaborar en proyectos de software libre, más de la mitad dice no haber recibido ningún tipo de recompensa por su labor. En el lado contrario, podemos ver que un número casi idéntico afirma haberse beneficiado del mismo. Eso sí, mientras la cifra de desarrolladores que dicen haber conseguido su actual puesto gracias a su relación con el software libre es apreciable, aquéllos que dicen haber recibido un aumento salarial por los mismos motivos es bastante pequeño. Asimismo, mirando el diagrama desde otro punto de vista, podemos decir que tres de cada cuatro desarrolladores de software libre se han beneficiado o pretenden beneficiarse en un futuro de la labor que desempeñan.

**Figura 24. [Gráfica: ¿Te has beneficiado profesionalmente por participar en**

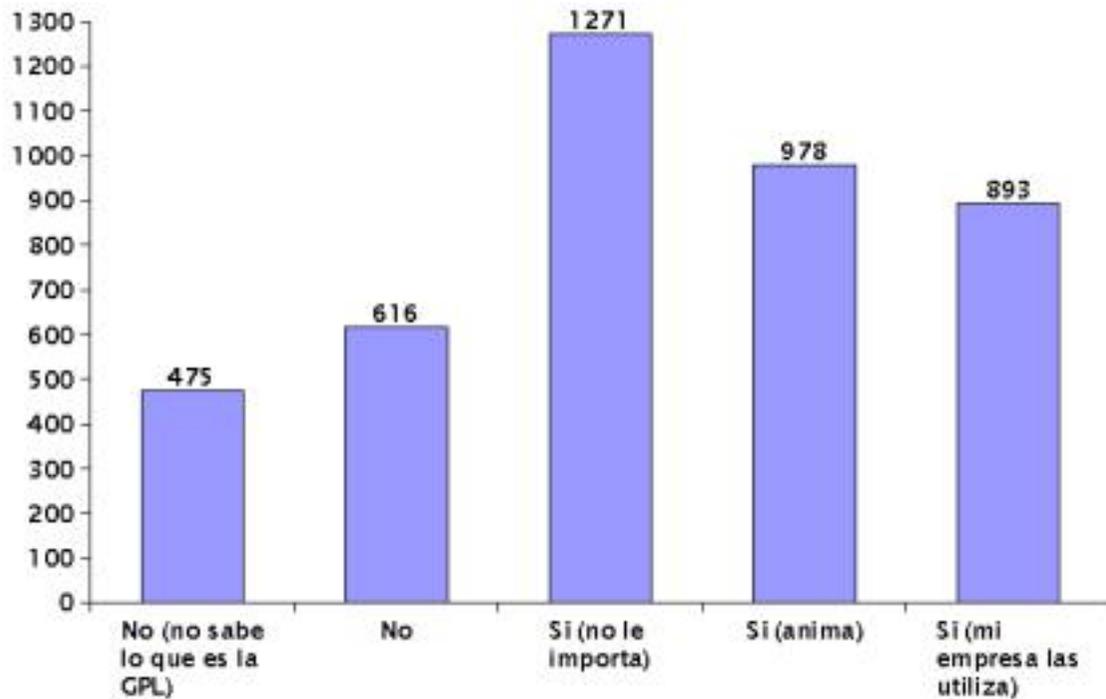
proyectos de software libre? (Widi). 5292 respuestas de 5478 desarrolladores]



Una de las razones más frecuentes a la hora de explicar por qué los desarrolladores dedican tiempo a la creación de software libre es que sus condiciones laborales no son satisfactorias. Por eso incluimos dos preguntas para ver qué es lo que nos podían contar los propios desarrolladores al respecto. La primera es la relación con su superior y qué tal lleva éste que se participe en desarrollos de software libre. Como se puede ver, en un tercio de los casos, los jefes ven con buenos ojos que se colabore. Aproximadamente otro tercio lo sabe, pero no le importa, y el último tercio se lo reparten los que no lo saben o incluso los jefes que no lo saben, por que no tienen ni idea de lo que es la GPL, la licencia más conocida dentro del mundo del software libre.

**Figura 25.** [Gráfica: ¿Sabe tu jefe que participas en proyectos de software libre?

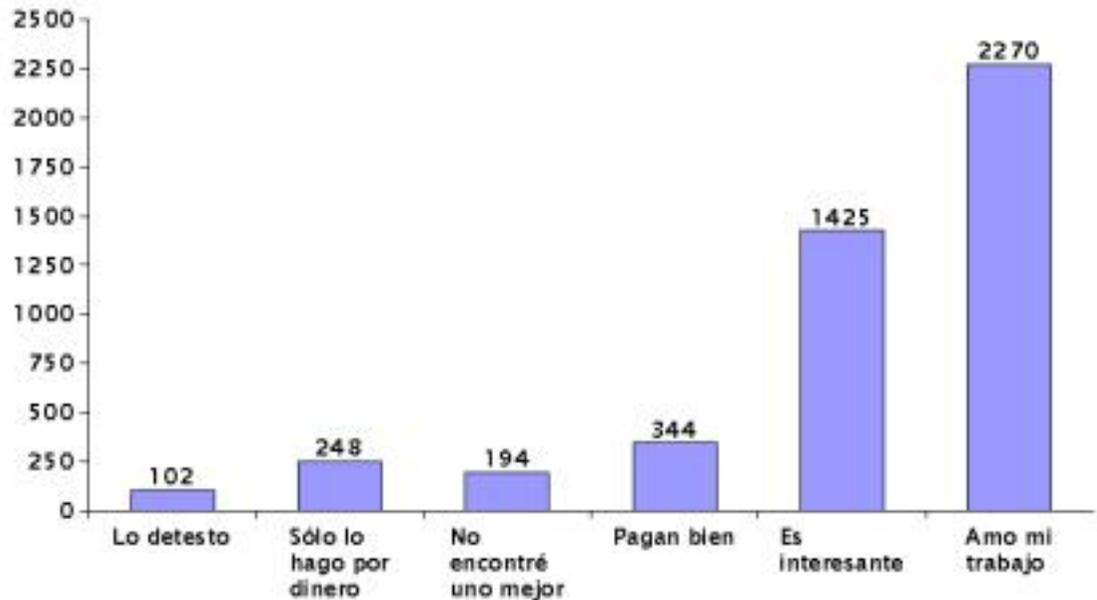
(Widi). 5194 respuestas de 5478 desarrolladores]



Sin embargo, nuestra curiosidad no se limitó a conocer la relación con sus superiores. También queríamos saber lo que piensan los desarrolladores de su puesto de trabajo. De los resultados podemos ver que a más de un 40% le gusta su actual puesto de trabajo, mientras que las cifras van disminuyendo paulatinamente con el grado de descontento. En resumen, por una u otra razón, más del 80% de los desarrolladores se encuentran a gusto con su actual puesto.

**Figura 26. [Gráfica: ¿Te gusta tu trabajo? (Widi). 5153 respuestas de 5478**

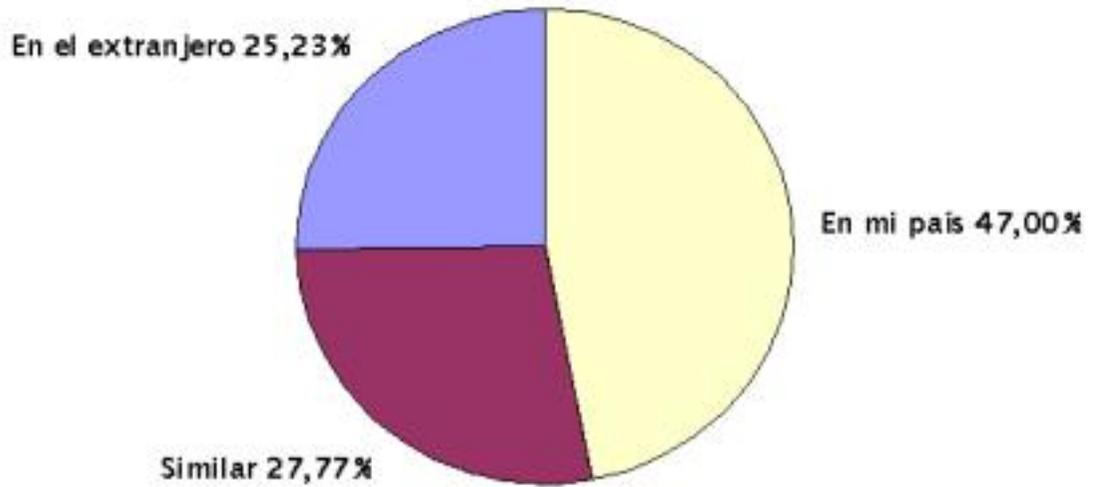
desarrolladores]



Ya que estamos preguntando a los desarrolladores por su situación laboral y el software libre, incluimos una pregunta para ver qué pensaban sobre las posibilidades de futuro del software libre en su país. La mitad de los desarrolladores encuestados opina que en su país hay más oportunidades que en el extranjero, mientras que un cuarto cree respectivamente que la situación está mejor en el extranjero o que es similar. La distribución de respuestas que los desarrolladores españoles nos dieron es muy parecida.

**Figura 27. [Gráfica: ¿Hay suficientes oportunidades en tu país? (Widi). 4333**

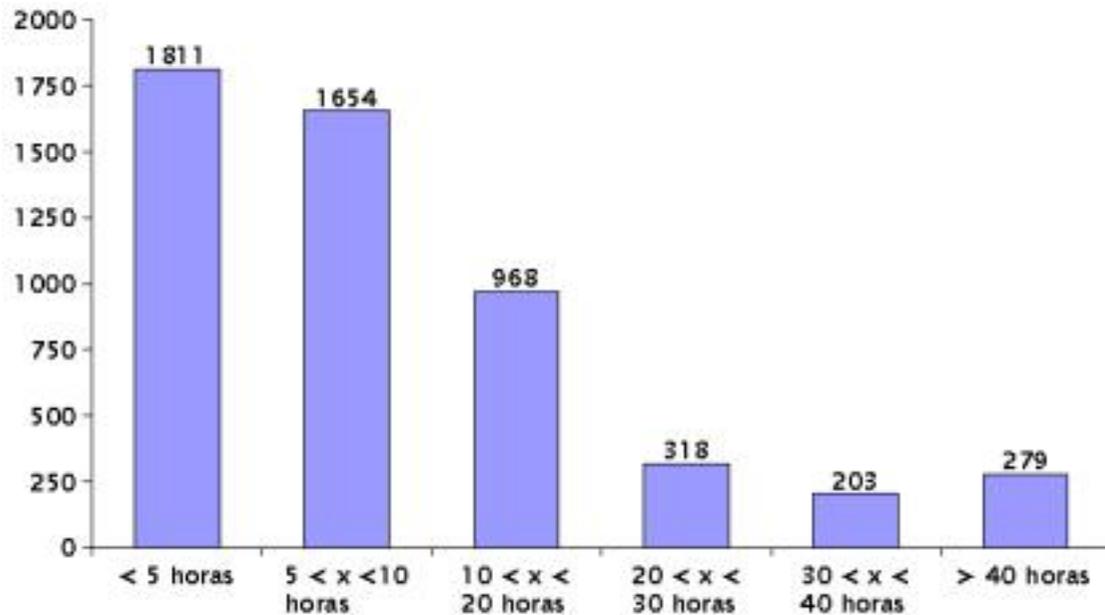
respuestas de 5478 desarrolladores]



De la siguiente gráfica se desprende que más de un 60% de los desarrolladores dedican menos de 10 horas a la semana al software libre. Aquéllos que le dedican más de 20 horas representan algo menos de la sexta parte del total.

**Figura 28. [Gráfica: Horas de desarrollo a la semana >(Widi). 5233 respuestas de**

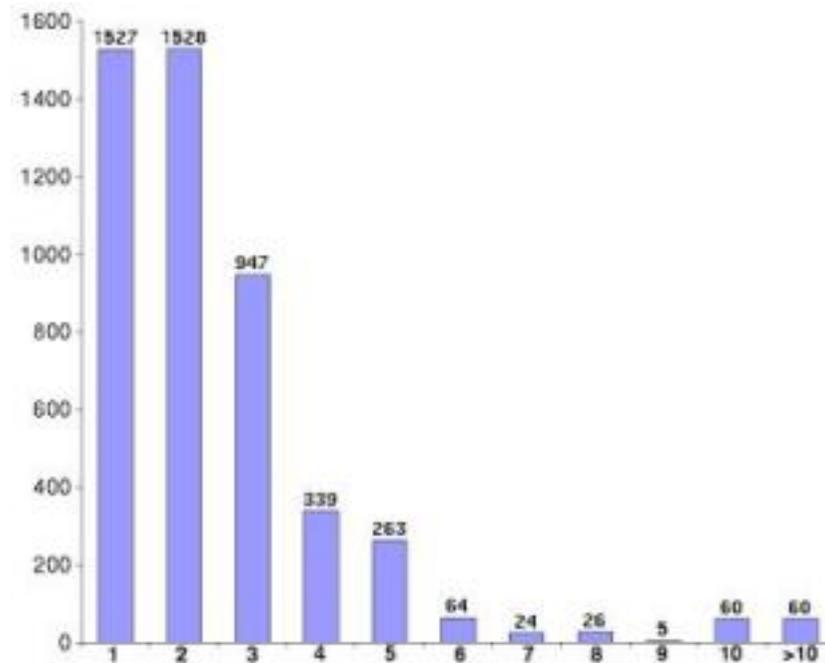
5478 desarrolladores]



Al igual que en el párrafo anterior en el que nos ocupamos de la dedicación temporal, podemos ver que el número de proyectos en los que un desarrollador participa, sigue una distribución similar. El 60% está vinculado a uno o dos proyectos, mientras que el número de desarrolladores que participan en cuatro o más no llega a la sexta parte.

**Figura 29. [Gráfica: Número de proyectos a los que estás vinculado (en Widi).**

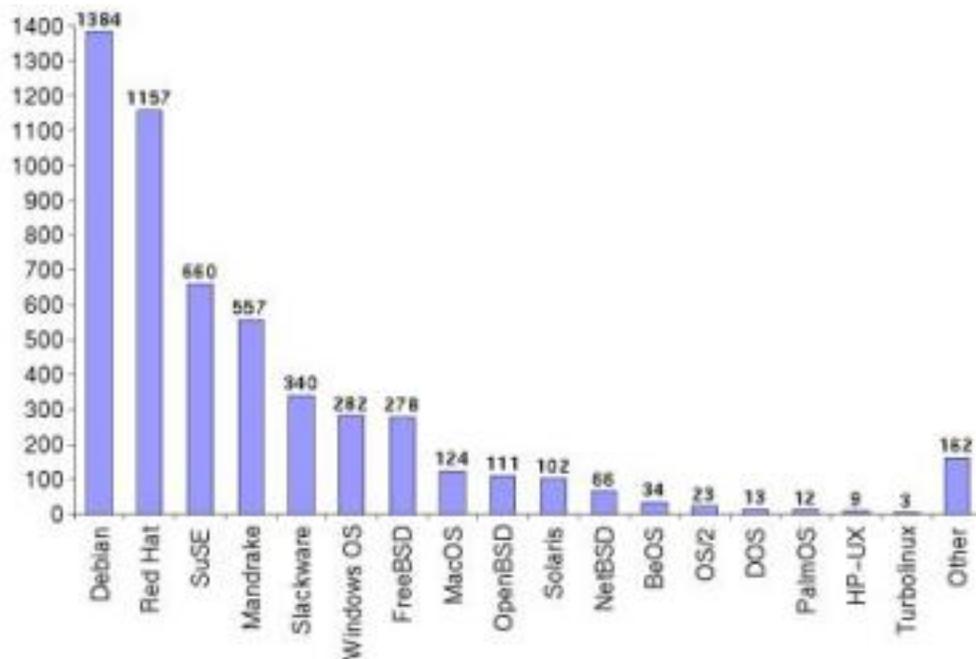
4843 respuestas de 5478 desarrolladores]



Entramos ya en la parte en la que cubrimos las preferencias de los desarrolladores. La primera pregunta con la que nos vamos a ocupar es con el sistema operativo / distribución preferido/a de los desarrolladores. Se puede ver que una mayoría prefiere las distribuciones de GNU/Linux, situándose Debian a la cabeza. Los adeptos a Windows son igual de numerosos que los que respondieron que su sistema operativo favorito era FreeBSD.

**Figura 30. [Gráfica: Distribución / sistema operativo favorita/o en Widi. 5317**

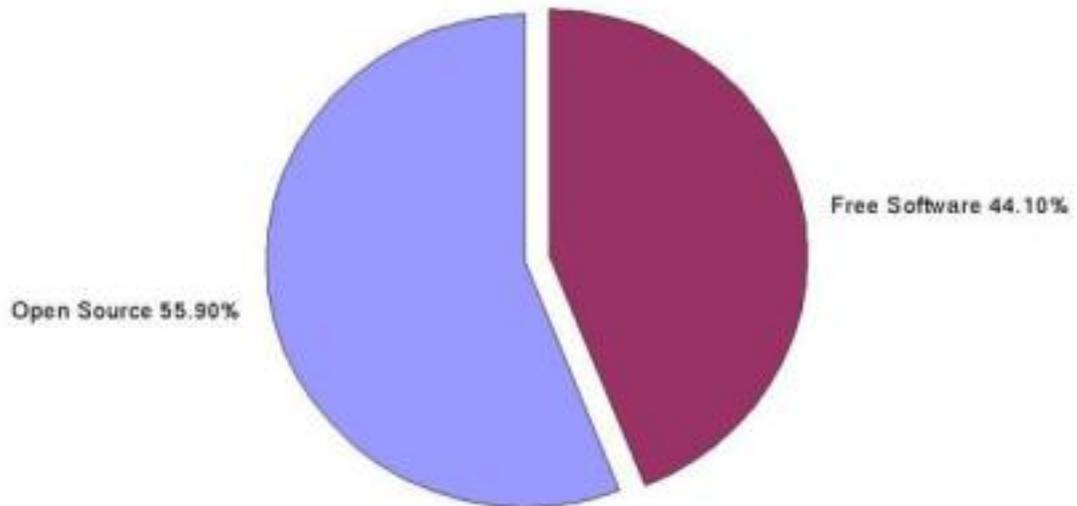
respuestas de 5478 desarrolladores]



En cuanto a la pregunta de si prefieren "Open Source" o "Free Software", se puede ver que en términos globales, hubo más desarrolladores que eligieron Open Source en vez de Free Software a la hora de responder esta pregunta. Hubo algunos que se quejaron, indicando que la pregunta no era suficientemente clara. Nuestra intención no era, ni mucho menos, aclarar conceptos; sólo queríamos saber de la manera más intuitiva posible, qué término escogían los desarrolladores. También hemos podido comprobar al correlacionar las respuestas de este apartado y las nacionalidades que mientras la dualidad de significados en lengua inglesa ha llevado a esta división de opiniones, en países como Francia o España, el término software libre (o logiciel libre) supera en aceptación al de Open Source.

**Figura 31. [Gráfica: "Open Source" o "Free Software" en Widi. 5104 respuestas**

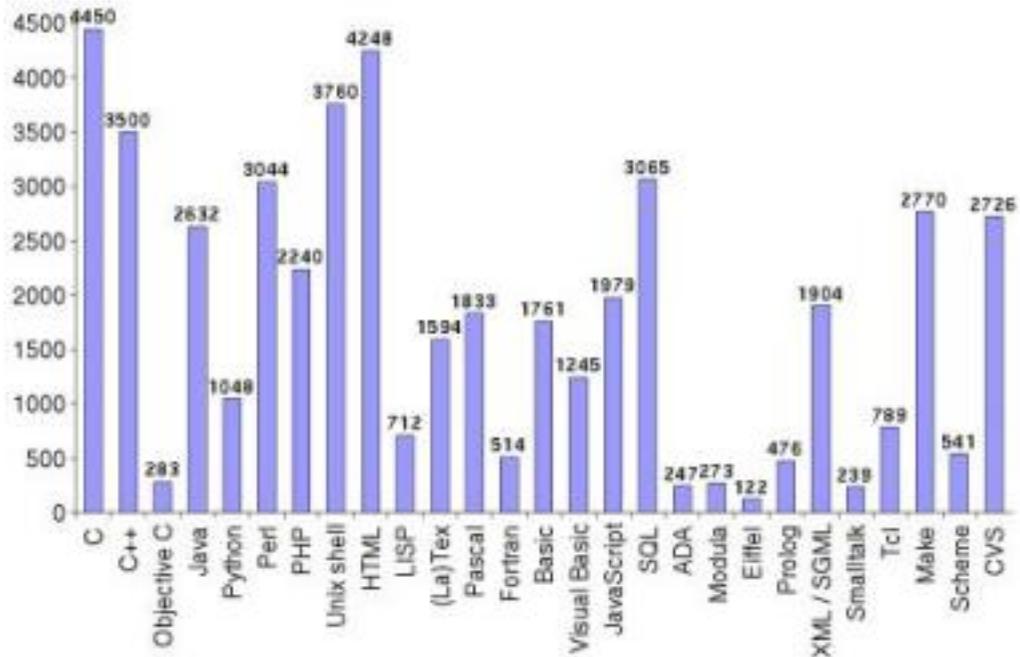
de 5478 desarrolladores]



Para terminar, y de manera escueta, mostramos la gráfica de las herramientas y lenguajes que los desarrolladores dicen conocer. Una nota curiosa al respecto es que esta distribución se mantiene con mínimas diferencias también de modo nacional y por continentes, lo que supone que los desarrolladores libres suelen tener unos conocimientos similares independientemente de su emplazamiento geográfico y cultural.

**Figura 32. [Gráfica: Herramientas y lenguajes de programación con los que los desarrolladores están familiarizados (en Widi). 5392 respuestas de 5478**

desarrolladores]



## 13. Conclusiones

En esta ponencia hemos podido acercarnos un poco más a la realidad de los desarrolladores de software libre. Cabe destacar que a pesar de que la mayoría de las grandes empresas del sector del software libre son americanas, las cuatro fuentes estudiadas muestran un claro predominio europeo en cuanto a número de desarrolladores. Europa no debe dejar pasar uno de los pocos trenes tecnológicos en el campo de las tecnologías de la información donde Europa muestra la delantera de forma nítida: políticamente se debe apoyar en mayor medida al software libre. No es comprensible que mientras gran parte del desarrollo se lleve en Europa, la industria se encuentre en los Estados Unidos. La "fuga de cerebros" es una realidad que se debe atajar con medidas oportunas. Las últimas decisiones encaminadas en aceptar las patentes del software y otras restricciones adicionales basadas en el copyright son más bien pasos en dirección contraria.

Sin embargo, aún dentro de Europa existen grandes diferencias entre los países del

norte y centro de Europa, donde el desarrollo del software libre está muy arraigado, y países del sur. España no es una excepción a esta regla y se encuentra en casi todas las estadísticas muy lejos de los países punteros. La situación es todavía peor en países del llamado Tercer Mundo. Esto es especialmente alarmante, ya que el software libre tiene unas características económicas (bajo coste de adquisición), sociales (fácil redistribución) y tecnológico-académicas (sistemas avanzados) muy interesantes para que no se sigan agrandando las diferencias existentes entre los países industrializados y los países en vías de desarrollo.

En cuanto a la preparación, hemos podido ver cómo los desarrolladores de software libre tienen, en general, una buena formación, tanto académica como relacionada con las ciencias informáticas. En la mayoría de los casos se trata de estudiantes o de jóvenes trabajadores altamente cualificados. El carácter abierto y multidisciplinar que muestra el software libre no es una sorpresa a tenor del grado de implicación de personas no relacionadas profesionalmente con el sector de las tecnologías de la información (casi un 20%) y de los desarrolladores que están estrechamente ligados a universidades, ya sea como profesores o como alumnos (casi un tercio).

Es interesante ver a las universidades como el motor para el desarrollo del software libre, ya que en su seno no sólo se gestan grandes aplicaciones y proyectos de investigación, sino un nutrido grupo de interesados que pueden adquirir a bajo coste nuevos conocimientos sobre tecnologías software punteras. En los últimos tiempos, cada vez que una empresa relacionada con el software libre quebraba muchas voces alarmistas han llenado columnas enteras en los sitios de noticias más visitados augurando un negro futuro al software libre. Este estudio muestra de manera concisa que el grado de implicación de desarrolladores adscritos a alguna universidad casi dobla el número de desarrolladores que dicen ser pagados por crear software libre. La buena salud del software libre no depende tanto de las tormentas bursátiles, sino más bien de su fomento, expansión y apoyo en las universidades y centros de formación.

Otro aspecto a tener en cuenta es la distribución de la dedicación personal. Si tenemos en cuenta que menos de una quinta parte de los desarrolladores dedican más de veinte horas semanales al desarrollo de software libre, se puede entender cómo según el Orbiten Survey (<http://www.orbiten.org/ofss/01.html>) un 10% de los autores había realizado más de un 70% del volumen de código investigado. Su conclusión sigue siendo válida en este estudio: el desarrollo del software libre puede que sea distribuido, pero la aportación individual es muy desigual.

De los resultados podemos leer también un alto grado de satisfacción profesional entre los desarrolladores, aún cuando más bien pocos se dedican profesionalmente al desarrollo del software libre y que la situación económica internacional mostraba claros síntomas de desaceleración durante las fechas en las que fueron encuestados.

Incluso la mayoría se muestra optimista cuando se les pregunta acerca de la situación del software libre en su país. Es curioso ver como incluso en España los desarrolladores se mostraban ciertamente esperanzados con las perspectivas que ofrece el software libre. A mi entender se debe básicamente a que esperan una explosión en nuestro país de este fenómeno en los próximos tiempos.

Para terminar, me gustaría enumerar de manera rápida cómo es el desarrollador de software libre medio: es europeo, tiene 27 años, varón, políglota y con conocimientos de inglés, estudiante de los últimos cursos de universidad o ingeniero software en uno de sus primeros empleos, tiene una alta graduación académica, no le pagan por desarrollar software libre, está contento con su actual situación laboral, desarrolla menos de 10 horas a la semana, usa Debian y tiene experiencia en varias (más bien, multitud de) herramientas y lenguajes de programación.