

## SOFTWARE HERITAGE

### ¿Quién es Roberto Di Cosmo?: Breve biografía y descripción de la carrera de Roberto Di Cosmo como científico de la computación.

Roberto Di Cosmo es un científico informático italiano nacido en 1962. Es conocido por su trabajo en el campo de la ingeniería del software, la programación funcional y la investigación en sistemas operativos. Estudió principalmente en la universidad de Pisa y desde entonces ha trabajado en varias universidades europeas.

Di Cosmo es un defensor del software libre y ha estado muy activo en la promoción del uso de herramientas de software libre en la educación y la investigación. También ha sido un defensor de la reforma de la ley de propiedad intelectual para promover el acceso libre a la información científica y tecnológica.

Actualmente es director del Software Heritage Project, una iniciativa para preservar y compartir el patrimonio del software libre y de código abierto en todo el mundo.

Dentro de sus contribuciones destaca la herramienta de compilación para el lenguaje de programación Caml, que es ampliamente utilizado en la comunidad de programación funcional, investigación en sistemas operativos distribuidos, teoría de la programación y semántica formal.

- ➔ Roberto Di Cosmo resume su vida en 3 líneas: 30 años investigando y enseñando las ciencias de la computación y la mitad de su vida investigando a cerca del software libre.
- ➔ En el podcast también se menciona que escribió el libro *Hijacking the World en el año 1998 que fue un best seller*

### ¿Qué es Software Heritage?: Descripción de qué es Software Heritage, cómo funciona y cuál es su importancia para la comunidad científica y tecnológica.

El Software Heritage es una iniciativa internacional para preservar y compartir el patrimonio del software libre y de código abierto en todo el mundo. Fue fundada en 2015 por Roberto Di Cosmo y otros expertos en software libre y tiene como objetivo construir una biblioteca universal de software libre para garantizar su preservación a largo plazo.

El proyecto utiliza técnicas de preservación digital para capturar el código fuente, el historial de desarrollo y la documentación asociada de todos los proyectos de software libre y de código abierto disponibles públicamente. Luego, se almacenan en un repositorio permanente y se proporciona acceso gratuito a través de una interfaz web.

El objetivo final de Software Heritage es ser la referencia mundial para el software libre y de código abierto, permitiendo a los desarrolladores y usuarios acceder, estudiar, compartir y utilizar el software de manera libre y sostenible. Además, el proyecto fomenta la colaboración y la reutilización del software para impulsar la innovación y el avance tecnológico en todo el mundo.

¿Por qué se creó Software Heritage?: Explicación de los motivos detrás de la creación de Software Heritage y su relevancia en la preservación del software. La misión de Software Heritage: Descripción de la misión de Software Heritage, sus objetivos y cómo estos están relacionados con la preservación del patrimonio tecnológico.

Se creó para abordar la creciente preocupación sobre la pérdida del patrimonio del software libre y de código abierto en todo el mundo. A medida que el software libre y de código abierto se vuelve más popular y se utiliza en una amplia variedad de aplicaciones, también aumenta la necesidad de garantizar su preservación a largo plazo.

A diferencia del software propietario, el software libre y de código abierto es creado y mantenido por una comunidad global de desarrolladores y usuarios, y su valor radica en su accesibilidad y capacidad de ser modificado y mejorado por cualquier persona. Sin embargo, a medida que los proyectos de software crecen y evolucionan, es cada vez más difícil rastrear y preservar el código fuente, el historial de desarrollo y la documentación asociada.

El objetivo del Software Heritage es construir una biblioteca universal de software libre y de código abierto para garantizar su preservación a largo plazo y hacer que sea accesible para todos. Al preservar el patrimonio del software libre y de código abierto, el proyecto también fomenta la colaboración y la reutilización del software, lo que puede impulsar la innovación y el avance tecnológico en todo el mundo.

La importancia de la preservación del software: Explicación de por qué la preservación del software es importante y por qué es esencial para el avance de la ciencia y la tecnología.

La preservación del software es importante por varias razones:

- **Patrimonio cultural:** El software es una parte importante de nuestro patrimonio cultural y tecnológico. A través del software, podemos entender cómo han evolucionado las tecnologías de la información y cómo han afectado a nuestra sociedad. La preservación del software nos permite documentar y conservar esta historia para las generaciones futuras.
- **Accesibilidad:** Muchos proyectos de software libre y de código abierto tienen un valor social y educativo significativo. Al preservar el software, se garantiza que siga siendo accesible a cualquier persona que quiera estudiarlo, utilizarlo o mejorarlo.
- **Seguridad:** La preservación del software también tiene implicaciones de seguridad. Si un software crítico deja de estar disponible debido a la pérdida del código fuente o de la documentación, puede haber graves consecuencias para la seguridad pública o la estabilidad económica. La preservación del software garantiza que los sistemas críticos puedan seguir siendo utilizados y manteniendo su integridad a lo largo del tiempo.
- **Innovación:** Al preservar el software, se fomenta la colaboración y la reutilización del código, lo que puede impulsar la innovación y el avance tecnológico. Al tener acceso a un gran corpus de software libre y de código abierto, los desarrolladores pueden aprovechar soluciones ya existentes para construir nuevas tecnologías y sistemas, lo que puede acelerar la innovación y reducir el costo de desarrollo.

¿Cómo funciona Software Heritage?

EL FUNCIONAMIENTO SE BASA PRINCIPALMENTE EN:

- UNA RECOPIACION AUTOMATICA DE SOFTWARE LIBRE Y CODIGO ABIERTO QUE SE ENCUENTRA PUBLICO EN INTERNET, USA TECNICAS AVANZADAS DE RECOPIACION DE DATOS PARA EXTRAERLOS Y ALMACENARLOS.
- ANALISIS DE METADATOS, QUE SON ANALIZADOS PROYECTO POR PROYECTO PARA COMPRENDERLOS Y RELACIONARLOS CON OTROS PROYECTOS DE SOFTWARE. USA TECNICAS DE ANALISIS DE LENGUAJE NATURAL PARA COMPRENDER TAMBIEN SU DOCUMENTACION.
- TODO ESTO SE ENGLOBA DENTRO DE UN REPOSITORIO DE ALTA CAPACIDAD DE CODIGO ABIERTO TAMBIEN QUE ALMACENA Y GESTIONA TODOS LOS DATOS.
- PARA ACCEDER A ESTE REPOSITORIO SE PROPORCIONA UNA HERRAMIENTA DE BUSQUEDA Y NAVEGACION DENTRO DE LA PROPIA PAGINA EN LA QUE SE PUEDE BUSCAR POR PROYECTO, AUTOR, LICENCIA Y ALGUN CRITERIO MAS.
- TAMBIEN PROPORCIONAN UNA API PARA ACCEDER A LOS DATOS DEL RESPOSITORIO, IMPORTARLOS O EXPORTARLOS.

PARA EL USUARIO SE BASA PRINCIPALMENTE EN:

- CREAR SU CODIGO EN UN REPOSITORIO PUBLICO CON ARCHIVOS DE README LICENCIAS Y DEMAS
- SUBIRLO A LA PAGINA MANUALMENTE AÑADIENDO EL LINK AL REPOSITORIO O BIEN USAR UNA EXTENSION DE CHROME PARA ELLO O BIEN MEDIANTE LA API QUE PROPORCIONAN QUE LO HARA DE MANERA AUTOMÁTICA, Y YA ESTARIA GUARDADO Y ACCESIBLE PARA TODO EL MUNDO.
- TAMBIEN RECOMIENDAN QUE REFERENCIAS EL CODIGO UNA VEZ SUBIDO

### ¿Cómo funciona Software Heritage?

EL FUNCIONAMIENTO SE BASA PRINCIPALMENTE EN:

- UNA RECOPIACION AUTOMATICA DE SOFTWARE LIBRE Y CODIGO ABIERTO QUE SE ENCUENTRA PUBLICO EN INTERNET, USA TECNICAS AVANZADAS DE RECOPIACION DE DATOS PARA EXTRAERLOS Y ALMACENARLOS.
- ANALISIS DE METADATOS, QUE SON ANALIZADOS PROYECTO POR PROYECTO PARA COMPRENDERLOS Y RELACIONARLOS CON OTROS PROYECTOS DE SOFTWARE. USA TECNICAS DE ANALISIS DE LENGUAJE NATURAL PARA COMPRENDER TAMBIEN SU DOCUMENTACION.
- TODO ESTO SE ENGLOBA DENTRO DE UN REPOSITORIO DE ALTA CAPACIDAD DE CODIGO ABIERTO TAMBIEN QUE ALMACENA Y GESTIONA TODOS LOS DATOS.
- PARA ACCEDER A ESTE REPOSITORIO SE PROPORCIONA UNA HERRAMIENTA DE BUSQUEDA Y NAVEGACION DENTRO DE LA PROPIA PAGINA EN LA QUE SE PUEDE BUSCAR POR PROYECTO, AUTOR, LICENCIA Y ALGUN CRITERIO MAS.
- TAMBIEN PROPORCIONAN UNA API PARA ACCEDER A LOS DATOS DEL RESPOSITORIO, IMPORTARLOS O EXPORTARLOS.

PARA EL USUARIO SE BASA PRINCIPALMENTE EN:

- CREAR SU CODIGO EN UN REPOSITORIO PUBLICO CON ARCHIVOS DE README LICENCIAS Y DEMAS
- SUBIRLO A LA PAGINA MANUALMENTE AÑADIENDO EL LINK AL REPOSITORIO O BIEN USAR UNA EXTENSION DE CHROME PARA ELLO O BIEN MEDIANTE LA API QUE

PROPORCIONAN QUE LO HARA DE MANERA AUTOMÁTICA, Y YA ESTARIA GUARDADO Y ACCESIBLE PARA TODO EL MUNDO.

- TAMBIEN RECOMIENDAN QUE REFERENCIAS EL CODIGO UNA VEZ SUBIDO
- **Implementación de la preservación a largo plazo:** El Software Heritage ha implementado técnicas avanzadas de preservación a largo plazo, lo que garantiza que el software almacenado en su repositorio esté disponible y accesible para las generaciones futuras.

### ¿Cómo puede ayudar Software Heritage a la comunidad

El Software Heritage puede ser de gran ayuda para la comunidad científica y tecnológica en varias maneras. Aquí te presento algunas de las formas en que la comunidad puede utilizarlo para acceder y utilizar software de manera más eficiente y efectiva:

- **Acceso a versiones antiguas de software:** almacena múltiples versiones de cada objeto de software, lo que permite a los usuarios acceder a versiones antiguas de software y compararlas con las versiones más recientes.
- **Recuperación de software perdido:** puede ser utilizado para recuperar software que ha sido perdido o eliminado de otros repositorios. Al mantener una copia de todo el software disponible públicamente, el Software Heritage puede actuar como una fuente de última instancia para recuperar software perdido.
- **Análisis de la evolución del software:** proporciona una herramienta de análisis de la evolución del software que permite a los desarrolladores y científicos analizar cómo ha evolucionado el software a lo largo del tiempo, identificar patrones y tendencias, y comprender mejor la historia del software.
- **Referencias de software:** proporciona una forma única de referenciar objetos de software mediante el uso de identificadores persistentes. Esto facilita la cita y el enlace a objetos de software específicos, lo que puede ser útil para la investigación científica y la documentación de software.
- **Preservación del patrimonio del software:** Finalmente, se enfoca en preservar el patrimonio del software a largo plazo, lo que garantiza que el software esté disponible y accesible para las generaciones futuras. Esto es particularmente importante para la comunidad científica y tecnológica, que a menudo depende de software de código abierto para llevar a cabo investigaciones y desarrollar nuevas tecnologías.