

# *de la cocina a la mesa*



ver notas al  
final de la  
presentación



5 abril 2023 →  Meta

# Segment Anything | Meta AI

COMPUTER VISION

## Segment Anything

April 05, 2023

 Cornell University

arXiv > cs > arXiv:2304.02643

Computer Science > Computer Vision and Pattern Recognition

[Submitted on 5 Apr 2023]

### Segment Anything

 Product ▾ Solutions ▾ Open Source ▾ Pricing

 facebookresearch / segment-anything Public

<> Code   Issues 387   Pull requests 37   Actions







## Jornadas sobre Información de Biodiversidad y Administraciones Ambientales 2022

Decíamos 'ayer'

23/11/2022

OpenAI

DALL·E 2

GPT-3

Texto  
predictivo

Playground

Load a preset...

Jose: No es costumbre poner tu propio nombre a una planta que describes tú mismo. De t  
¿puedes ampliar la reseña con el lugar de localización, la descripción del biotopo y alguna  
cómo supiste que estabas ante una nueva especie?

Salvador: La especie *Sideritis Rivas-Godayi* se localiza en la Península Ibérica, concretamente en la  
comunidad autónoma de Andalucía. El biotopo en el que se encuentra es un matorral esclerófilo  
mediterráneo.

La anécdota que cuenta el descubrimiento de esta especie es que un día, mi hijo y yo estábamos  
caminando por un sendero en Andalucía y, de repente, nos tropezamos con una planta que no  
habíamos visto antes. Nos dimos cuenta de que se trataba de una especie nueva y, desde entonces, la  
hemos estado estudiando.

Jose: Gracias Salvador, ¿puedes escribir la reseña completa (unas 10-15 líneas) con todos los datos que  
me has dado y terminar con algún latinajo o comentario en latín que le de nivel al texto?

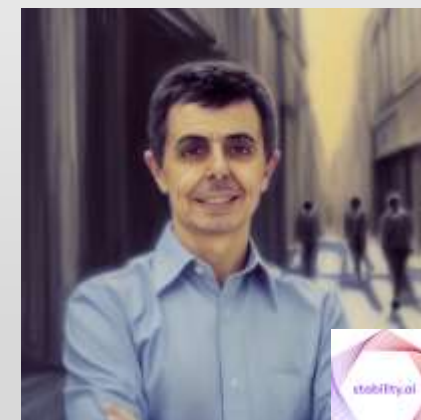
Salvador:

La especie *Sideritis Rivas-Godayi* se encuentra en la Península Ibérica, concretamente en la  
comunidad autónoma de Andalucía. El biotopo en el que se encuentra es un matorral esclerófilo  
mediterráneo.

La especie se distingue por sus flores amarillas y sus hojas oblongas. El descubrimiento de esta  
especie se debió a una casualidad, ya que mi hijo y yo nos encontramos con ella mientras  
caminábamos por un sendero en Andalucía.



Texto  
+  
Imagen



Stable Diffusion



El tren de la IA  
no para

bing chat

30/11/2022



14/3/2023

GPT-4V(ision) System Card

OpenAI

25/9/2023

September 25, 2023



- GPT4-turbo
- GPTs
- DALL·E-3
- Whisper-3
- ...

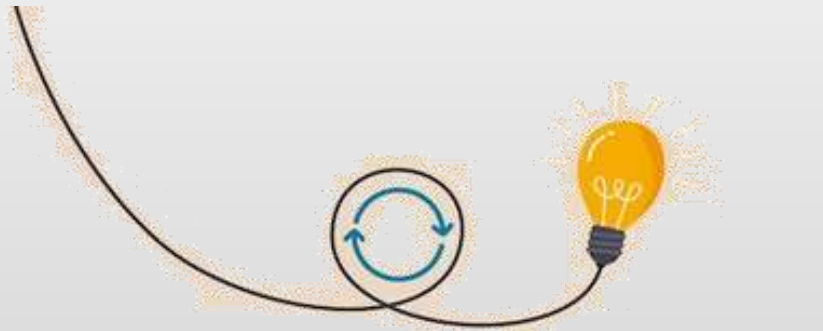
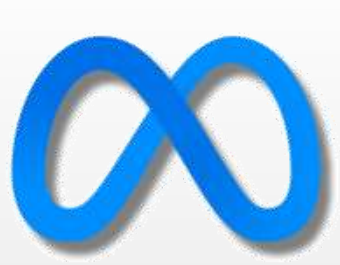


6/11/2023



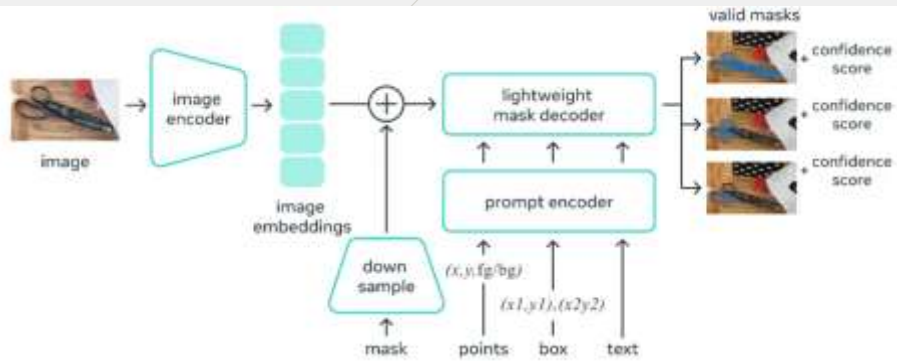
Aplicaciones de ChatGPT en el sector forestal

14/1/2024



# SAM: A Image Segmentation Foundation Model

Segment Anything Model by Facebook's Meta AI Research



Cornell University

arXiv > cs > arXiv:2304.02643

Computer Science > Computer Vision and Pattern Recognition

[Submitted on 5 Apr 2023]

**Segment Anything**

---

Product Solutions Open Source Pricing

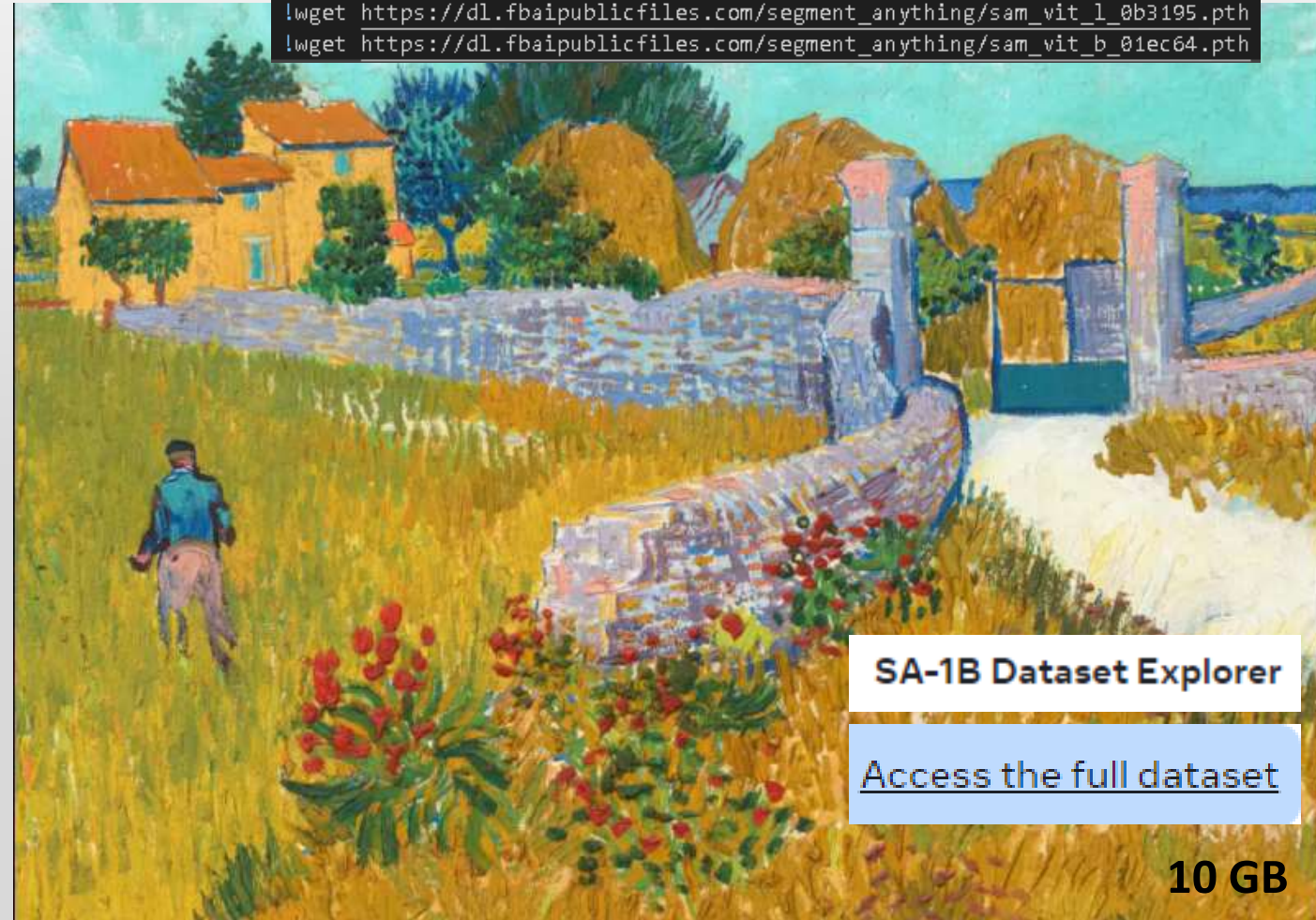
facebookresearch / segment-anything Public

<> Code Issues 387 Pull requests 37 Actions

¿Qué pasa cuando se publica un modelo?  
5/4/2023 → 15/11/2023

# Segment Anything | Meta AI

```
!wget https://dl.fbaipublicfiles.com/segment_anything/sam_vit_h_4b8939.pth  
!wget https://dl.fbaipublicfiles.com/segment_anything/sam_vit_l_0b3195.pth  
!wget https://dl.fbaipublicfiles.com/segment_anything/sam_vit_b_01ec64.pth
```





# Meta publica SAM -> ¿Que pasa después?



Papers 



Huggingface



github



tools





PROMISE: PROMPT-L

## Prototype-oriented Unsupervised Change Detection for Disaster Management

ATTACK ON SEGMENT D)

Youngtack Oh, Minseok Seo, Doyi Kim, Junghoon Seo

SI Analytics

70, Yuseong-daero 1689beon-gil, Yuseong-gu, Daejeon, Republic of Korea  
{ytoh96,minseok.seo,doyikim,jhseo}@si-analytics.ai

Chaoning Zhang\*  
Kyung Hee University

### Abstract

Climate change has led to an increased frequency of natural disasters such as floods and cyclones. This emphasizes the importance of effective disaster monitoring. In response, the remote sensing community has explored change detection methods. These methods are primarily categorized into supervised techniques

examples, which change the model input. Recently, Segment Anything model in computer vision due to its viable attacks on SAM is beneficial and Prasad

## ANYTHING MEETS UNIVERSAL ADVERSARIAL PERTURBATION

A PREPRINT

Ian  
versity

Sheng Zheng  
Beijing Institute of Technology

Chaoning Zhang\*  
Kyung Hee University

October 20, 2023

### ABSTRACT

As Segment Anything Model (SAM) becomes a popular foundation model in computer vision, its adversarial robustness has become a concern that cannot be ignored. This works investigates whether

Towards Training-free Prompt

Article

## Segment Anything Model Inventory: Assessing Zero-shot Performance on Geospatial Data

Jiahao Xia<sup>1</sup>, Gavin Gong<sup>2,†</sup>, Jiawei Liu<sup>3,†</sup>, Zhigang Zhu<sup>4</sup> and Hao Tang<sup>5\*</sup>

<sup>1</sup> Department of Civil and Environmental Engineering, Rutgers, The State University of New Jersey; jx198@soe.rutgers.edu

<sup>2</sup> East Brunswick High School; gonggavineb@gmail.com

<sup>3</sup> Data Science and Engineering Program, The CUNY City College; jliu021@citymail.cuny.edu

<sup>4</sup> Department of Computer Science, The CUNY City College, The CUNY Graduate Center; zzhu@ccuny.cuny.edu

<sup>5</sup> Department of Computer Information Systems, The CUNY BMCC, The CUNY Graduate Center; htang@bmcc.cuny.edu

\* Correspondence: htang@bmcc.cuny.edu

† These authors contributed equally to this work.

**Abstract:** In this paper, a Segment Anything Model (SAM)-based pedestrian infrastructure segmentation workflow is designed and optimized, which is capable of efficiently processing multi-sourced

tional model for segmenting ob  
Fine-tuned exclusively on a ric  
ultrasound image-mask pairs, &  
performance on 8 unseen ultras  
ing methods by a significant margin on all metric  
achieves average dice similarity score of > 90% c

signifi  
segmen  
mantic

t 2023

17 Oct 2023

24 Oct

7 Oct 23

19 Oct


19 Oct 2

2



race



# The AI community building the future.

The platform where the machine learning community collaborates on models, datasets, and applications.

Computer Vision

Datasets 77,782

Models 4,272

- Depth Estimation
- Image Classification
- Object Detection
- Image Segmentation
- Image-to-Image
- Unconditional Image Generation
- Video Classification
- Zero-Shot Image Classification

## Model Card for Segment Anything Model (SAM)

- SAM
- Overview
- SamConfig
- SamVisionConfig
- SamMaskDecoderConfig
- SamPromptEncoderConfig
- SamProcessor
- SamImageProcessor
- SamModel
- TFSamModel

### Resources:

- [Demo notebook](#) for using the model.
- [Demo notebook](#) for using the automatic mask generation pipeline.
- [Demo notebook](#) for inference with MedSAM, a fine-tuned version of SAM on the medical domain.
- [Demo notebook](#) for fine-tuning the model on custom data.







facebookresearch / segment-anything

<> Code Issues 389 Pull requests 37 Actions Projects Security Insights

### Running SAM on Large Satellite Images and Implementing Encoder-Decoder Workflow #497

Open Ankit-Vohra opened this issue on Jul 26 · 3 comments

opengeos / segment-geospatial

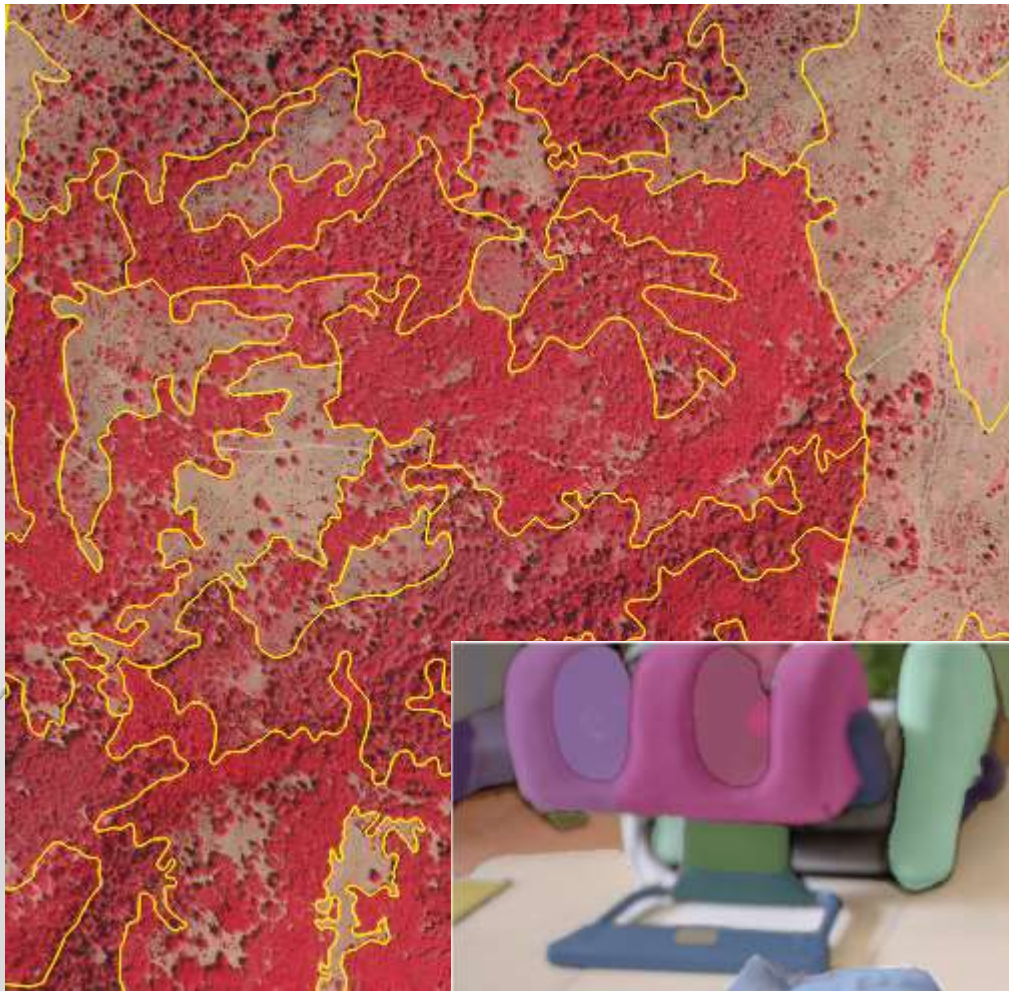
<> Code Issues 17 Pull requests Discussions Actions Projects Security Insights

segment-geospatial Public

facebookresearch / segment-anything

<> Code Issues 389 Pull requests 37 Actions Projects Security Insights

segment-anything Public



**segment-anything**

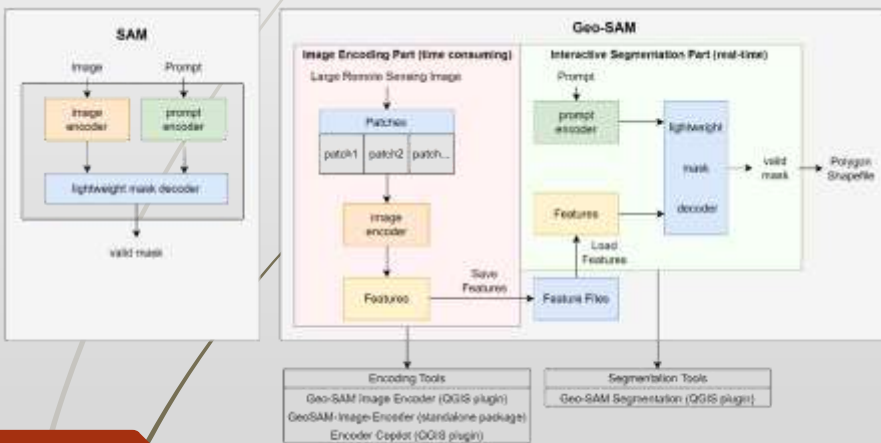


# Geo SAM

DOI: [10.5281/zenodo.8191039](https://doi.org/10.5281/zenodo.8191039) docs passing

By Joey and [Fancy](#) from [Cryosphere Lab](#), ESSC, CUHK.

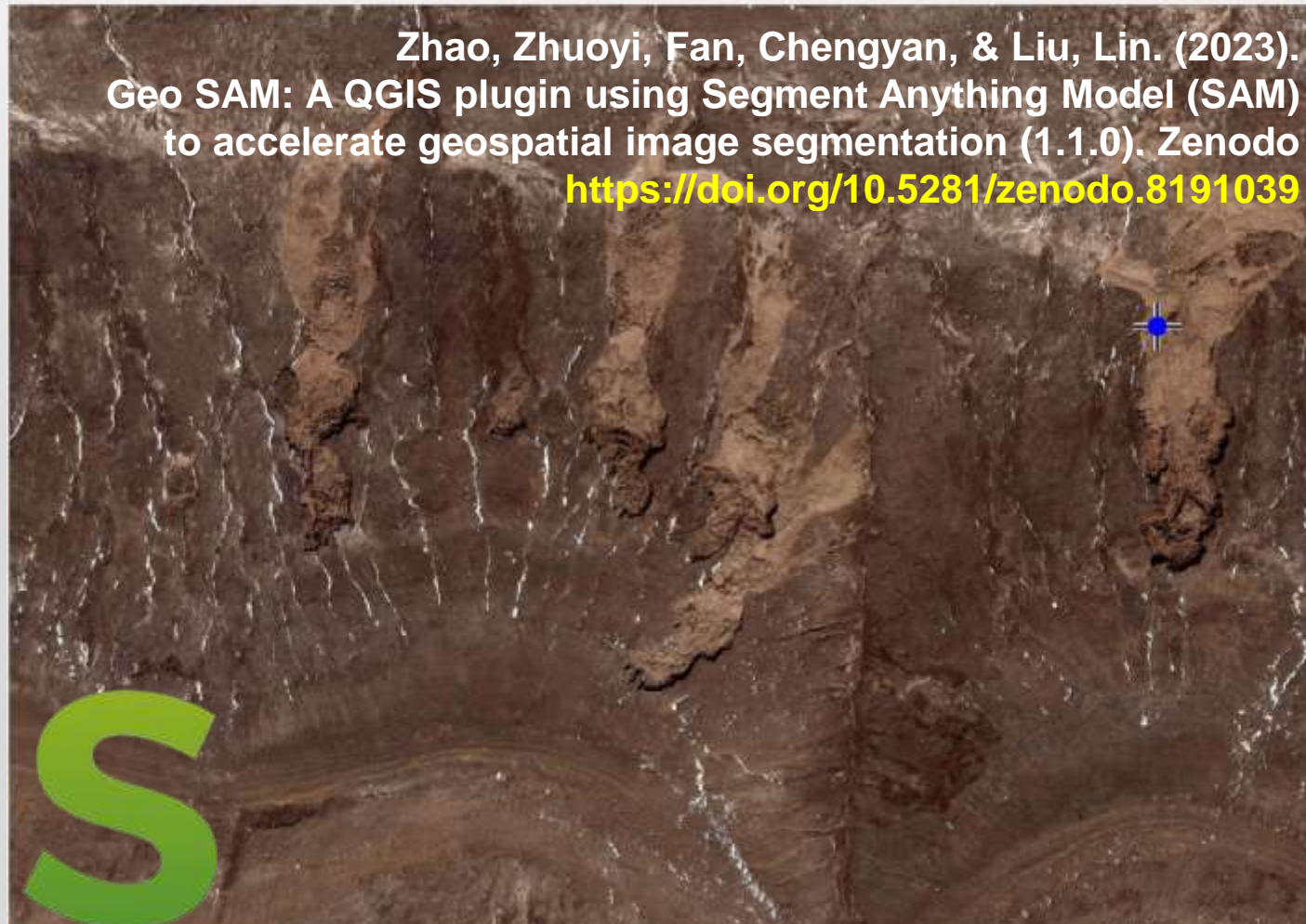
- [Installation and Usage](#)
- [Introduction](#)
- [Reasons for choosing Geo-SAM](#)
- [Demos](#)
- [Citation](#)
- [Acknowledgement](#)



## Geo-SAM Tool



Zhao, Zhuoyi, Fan, Chengyan, & Liu, Lin. (2023). Geo SAM: A QGIS plugin using Segment Anything Model (SAM) to accelerate geospatial image segmentation (1.1.0). Zenodo <https://doi.org/10.5281/zenodo.8191039>





# Geo SAM [↗](#)

DOI [10.5281/zenodo.8191039](https://doi.org/10.5281/zenodo.8191039) docs [passing](#)

By Joey and [Fancy](#) from [Cryosphere Lab](#), ESSC, CUHK.

- [Installation and Usage](#)
- [Introduction](#)
- [Reasons for choosing Geo-SAM](#)
- [Demos](#)
- [Citation](#)
- [Acknowledgement](#)

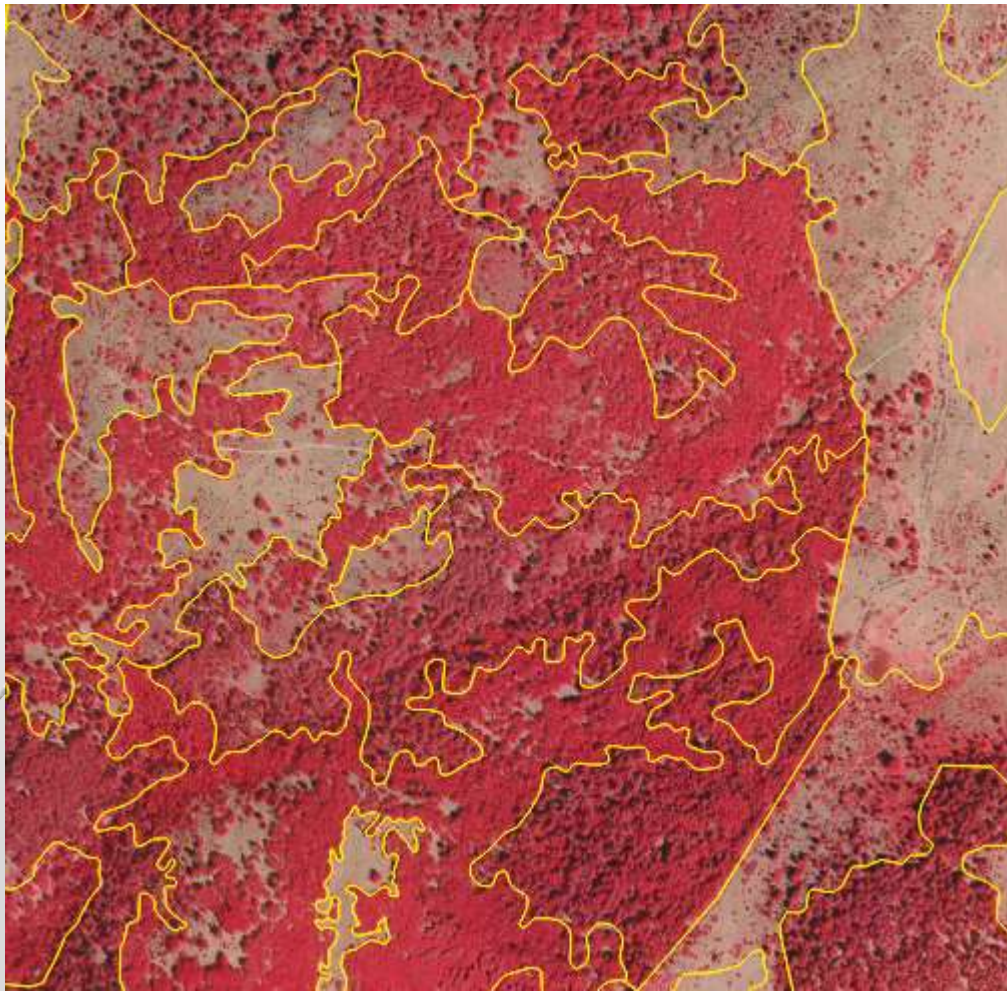
Geo-SAM Tool 📄 🗑️

Prompt **Settings**

<b>Add Points</b>	<b>Add BBox</b>	<b>Clear/Undo</b>	<b>Minimum Pixels</b>	<b>Save Results</b>	
FG BG	BBox	Clear Undo	0,000	Save	
<b>SAM Output Feature</b>		<b>Feature Folder</b>			
<b>From Layer</b>	<b>From file</b>	i25_0483_3\sam_feat_vit_l_bands_1_2_3_53d9' ... Load		<b>Preview mode</b>	<b>Editable</b>
polygon_sam	...			<input checked="" type="radio"/> Enable	<input checked="" type="radio"/> Enable







¿anything?



segment-anything



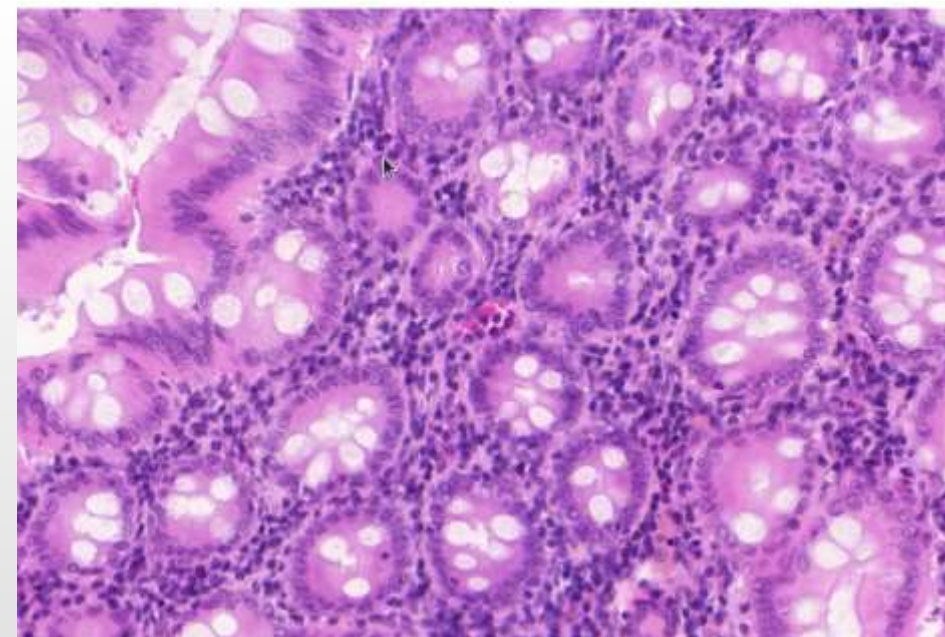








GitHub repository header for **MedSAM** by **bowang-lab**. It includes navigation links for **Code**, **Issues**, and **Pull requests** (4). The repository is marked as **Public**.



Colab interface for the notebook **Fine\_tune\_SAM\_(segment\_anything)\_on\_a\_custom\_dataset.ipynb**. The interface includes a menu with options: **Archivo**, **Editar**, **Ver**, **Insertar**, **Entorno de ejecución**, **Herramientas**, and **Ayuda**. Below the menu are buttons for **+ Código**, **+ Texto**, and **Copiar en Drive**.

Hugging Face interface for the dataset **keremberke/satellite-building-segmentation**. The page shows the dataset title, a search bar, and navigation options for **Models** and **Datasets**. The dataset is categorized under **Image Segmentation** with tags **roboflow**, **roboflow2huggingface**, and **Aerial**. It has **6 likes**. The **Dataset card** is selected, and the **Dataset Viewer** is open, showing a **Subset** of **full (9.67k rows)** and a **Split** of **train (6.76k rows)**. There are also links for **Auto-converted to Parquet**, **API**, and **Go to dataset viewer**.

¿No nos gusta  
la segmentación?  
Hacemos  
*fine tuning*:  
iniciando el  
entrenamiento...

Imagen original



Máscara real



Máscara inferida pre



Máscara inferida post





***Fine tuning:***  
el descenso del  
gradiente  
responde

Imagen original



Máscara real



Máscara inferida pre



Máscara inferida post



**Conclusión:**  
podemos adaptar  
el modelo y las  
herramientas a  
nuestras  
necesidades





# The AI community building the future.

The platform where the machine learning community collaborates on models, datasets, and applications.

🏠 Geo-SAM

latest

Search docs

## SAM: A Image Segmentation Foundation Model

Segment Anything Model by Facebook's Meta AI Research





Editor: [Hugo Ladoux](#)

Reviewers:

- [@ayoubft](#)
- [@aazuspan](#)
- [@SimonMolinsky](#)

Submitted: 18 May 2023  
Published: 12 September 2023

## sameo: A Python package for segmenting geospatial data with the Segment Anything Model (SAM)

Qiusheng Wu<sup>1</sup> and Lucas Prado Osco<sup>2</sup>

<sup>1</sup> Department of Geography & Sustainability, University of Tennessee, Knoxville, TN 37996, United States <sup>2</sup> Faculty of Engineering and Architecture and Urbanism, University of Western São Paulo, Rod. Raposo Tavares, km 572 - Limeiro, Pres. Prudente 19067-175, SP, Brazil

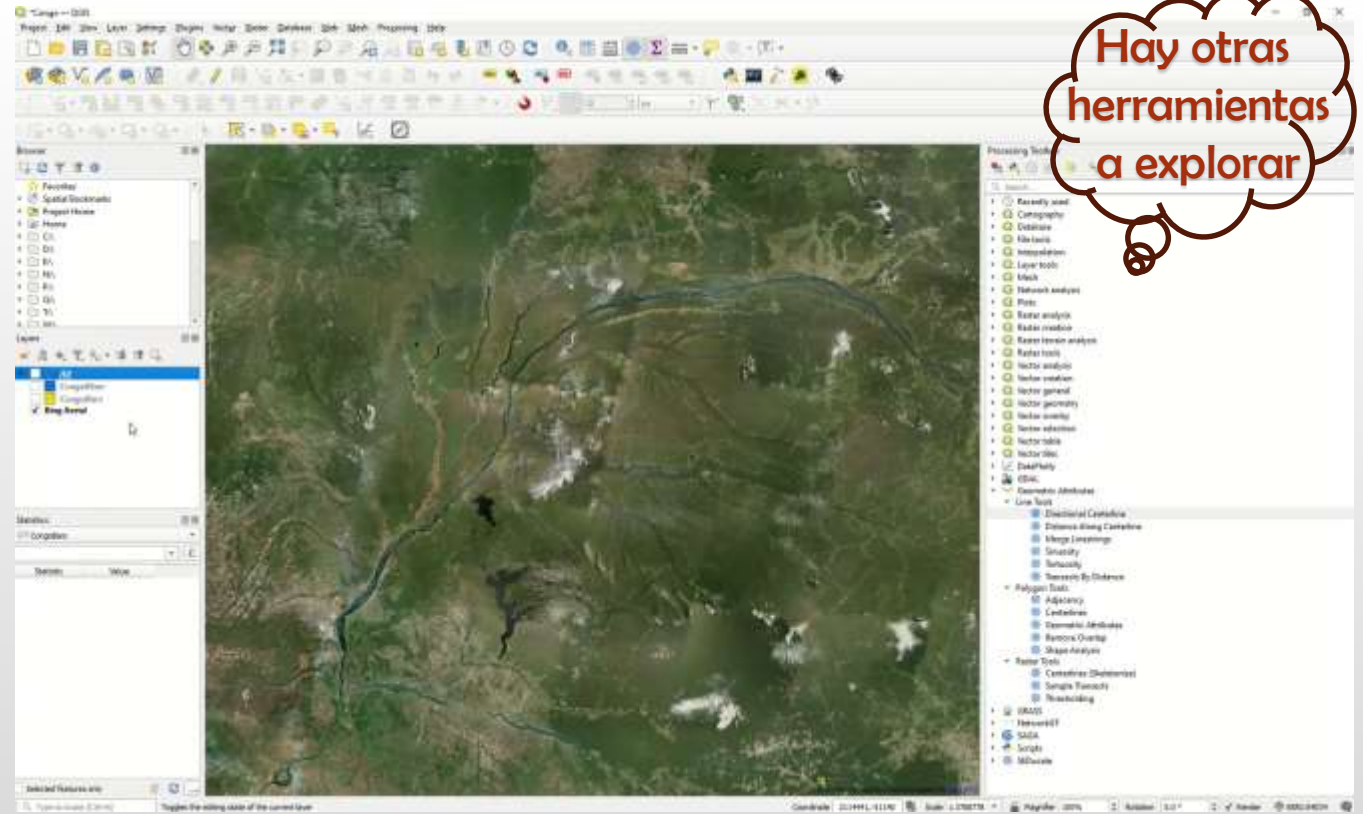
DOI: [10.21105/joss.05663](https://doi.org/10.21105/joss.05663)

Welcome to sameo

[Open](#)
[Studio Lab](#)
[Open](#)
[Planetary Computer](#)
[Open in Colab](#)
[pypi v0.10.2](#)
[conda-forge v0.10.2](#)
[pulls 319](#)

[downloads 55k](#)
[downloads 7k total](#)
[JOSS 10.21105/joss.05663](#)

A Python package for segmenting geospatial data with the Segment Anything Model (SAM)



### Using SAM with Desktop GIS

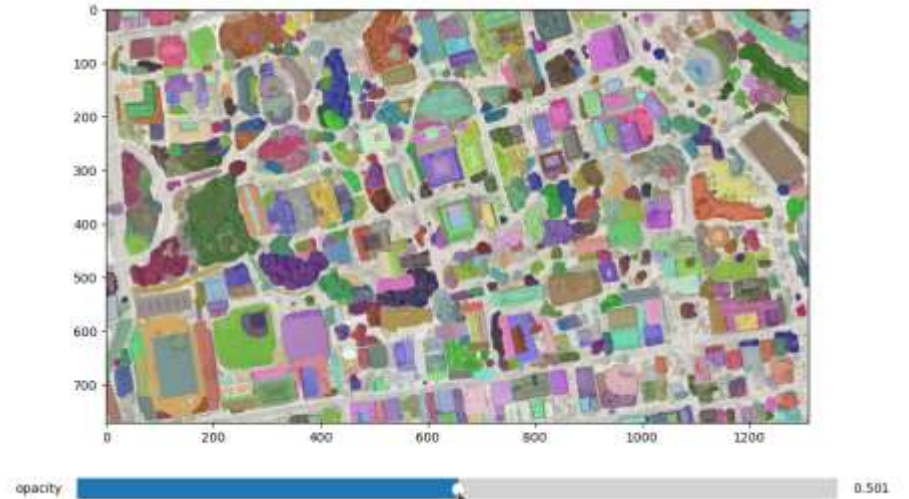
- **QGIS:** Check out the [Geometric Attributes plugin for QGIS](#). Credit goes to Bjorn Nyberg.
- **ArcGIS:** Check out the [Segment Anything Model \(SAM\) Toolbox for ArcGIS](#) and the [Resources for Unlocking the Power of Deep Learning Applications Using ArcGIS](#). Credit goes to Esri.

## segment-geospatial

[Open](#)
[Studio Lab](#)
[Open](#)
[Planetary Computer](#)
[Open in Colab](#)
[pypi v0.10.2](#)
[conda-forge v0.10.2](#)
[pulls 316](#)

[downloads 55k](#)
[downloads 7k total](#)
[JOSS 10.21105/joss.05663](#)

A Python package for segmenting geospatial data with the Segment Anything Model (SAM)



Hay otras herramientas a explorar





*Muchas gracias*



Jose Bengoa

[jose.bengoa@jcyf.es](mailto:jose.bengoa@jcyf.es)



# TREE-GPT: MODULAR LARGE LANGUAGE MODEL EXPERT SYSTEM FOR FOREST REMOTE SENSING IMAGE UNDERSTANDING AND INTERACTIVE ANALYSIS

S. Q. Du<sup>a,b,c</sup>, S.J. Tang<sup>a,b,c</sup>\*, W. X. Wang<sup>a,b,c</sup>, X. M. Li<sup>a,b,c</sup>, R. Z. Guo<sup>a,b,c</sup>

<sup>a</sup> Research Institute for Smart Cities, School of Architecture and Urban Planning, Shenzhen University, Shenzhen, P.R. China

<sup>b</sup> State Key Laboratory of Subtropical Building and Urban Science, Shenzhen, P.R. China

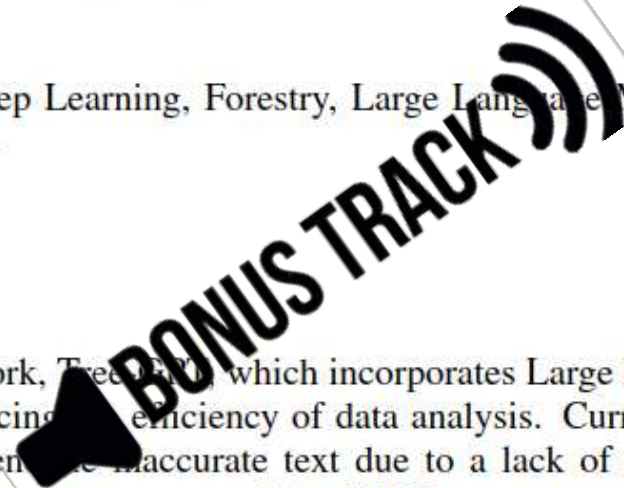
<sup>c</sup> Guangdong–Hong Kong–Macau Joint Laboratory for Smart Cities, P.R. China

**KEY WORDS:** Remote Sensing, Deep Learning, Forestry, Large Language Model, Individual Tree Segmentation, Tree Factor Estimation, Segment Anything Model.

## ABSTRACT:

This paper introduces a novel framework, Tree-GPT, which incorporates Large Language Models (LLMs) into the forestry remote sensing data workflow, thereby enhancing the efficiency of data analysis. Currently, LLMs are unable to extract or comprehend information from images and may generate inaccurate text due to a lack of domain knowledge, limiting their use in forestry data analysis. To address this issue, we propose a modular LLM expert system, Tree-GPT, that integrates image understanding modules, domain knowledge bases, and toolchains. This empowers LLMs with the ability to comprehend images, acquire accurate knowledge, generate code, and perform data analysis in a local environment. Specifically, the image understanding module extracts structured information from forest remote sensing images by utilizing automatic or interactive generation of prompts to guide the Segment Anything Model (SAM) in generating and selecting optimal tree segmentation results. The system then calculates tree structural parameters based on these results and stores them in a database. Upon receiving a specific natural language instruction, the

7 Oct 2023





Esta presentación es una versión retocada de la expuesta en las Jornadas sobre Información de Biodiversidad y Administraciones Ambientales de 2023 en el Jardín Botánico de Madrid. Se ha incluido en este apartado de notas una transcripción revisada y adaptada para facilitar la visualización.



5 abril 2023 →  Meta

## Segment Anything | Meta AI

COMPUTER VISION

### Segment Anything

April 05, 2023

 Cornell University





 arXiv:2304.02643

Computer Science > Computer Vision and Pattern Recognition  
(Submitted on 5 Apr 2023)

**Segment Anything**

 Product  Solutions  Open Source  Pricing

[facebookresearch/segment-anything](https://facebookresearch.github.io/segment-anything/)

 Code  Issues  Pull requests  Actions



El 5 de abril de este año, Meta publicó un modelo que se llama Segment Anything. Un modelo pensado para poder segmentar cualquier imagen.  
Pero ¿Qué es segmentar? Aquí tenemos un ejemplo.





## Jornadas sobre Información de Biodiversidad y Administraciones Ambientales 2022

Jornadas sobre Información de Biodiversidad y Administraciones Ambientales 2022 23/11/2022

Y, si lo traemos a nuestro terreno, aquí tenemos otro ejemplo,



Decíamos 'ayer' 23/11/2022

OpenAI GPT-3 DALL-E 2

Texto predictivo

Playground

Load a preset

Joe: No es costumbre poner tu propio nombre a una planta que describes tú mismo. De-  
 igüentes ampliar la reseña con el lugar de localización, la descripción del biotopo y alguna  
 cómo supiste que estabas ante una nueva especie?

Salvador: La especie *Sideritis Rivas-Godoyi* se localiza en la Península Ibérica, concretamente en la  
 comunidad autónoma de Andalucía. El biotopo en el que se encuentra es un matorral esclerófilo  
 mediterráneo.

La anécdota que cuenta el descubrimiento de esta especie es que un día, mi hijo y yo estábamos  
 caminando por un sendero en Andalucía, de repente, nos tropezamos con una planta que no  
 habíamos visto antes. Nos dimos cuenta de que se trataba de una especie nueva y desde entonces, la  
 hemos estado estudiando.

Joe: Gracias Salvador, ¿puedes escribir la reseña completa (unas 10-15 líneas) con todos los datos que  
 me has dado y terminar con algún latín o comentario en latín que le de nivel al texto?

Salvador:

La especie *Sideritis Rivas-Godoyi* se encuentra en la Península Ibérica, concretamente en la  
 comunidad autónoma de Andalucía. El biotopo en el que se encuentra es un matorral esclerófilo  
 mediterráneo.

La especie se distingue por sus flores amarillas y sus hojas oblongas. El descubrimiento de esta  
 especie se debió a una casualidad, ya que mi hijo y yo nos encontramos con ella mientras  
 caminábamos por un sendero en Andalucía.

Model: text-davinci-002

Temperature: 0.75

Maximum length: 164

Stop sequences: (none)

Top P: 1

Frequency penalty: 0

Presence penalty: 0

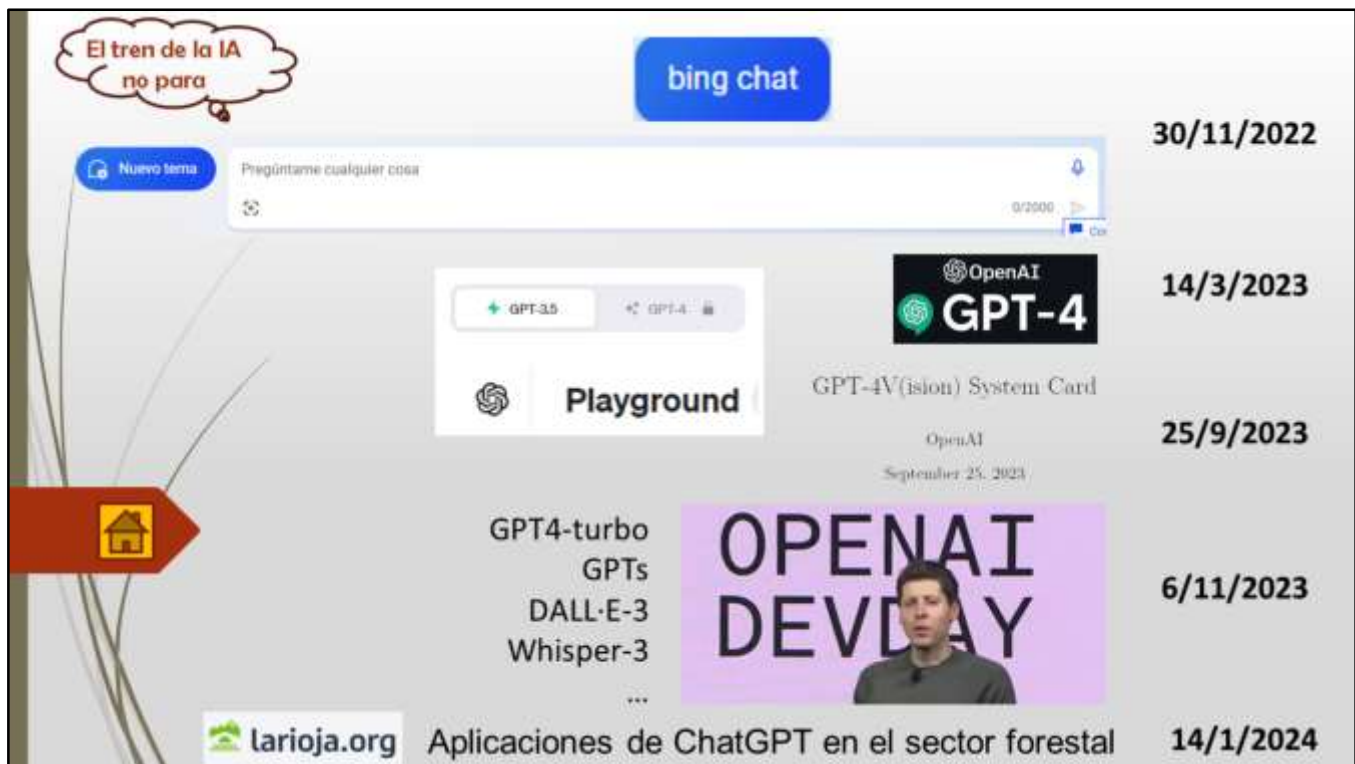
Best of: 1

Stable Diffusion

Texto + Imagen

Jornadas sobre Información de Biodiversidad y Administraciones Ambientales 2022 23/11/2022

Tampoco parece nada del otro mundo, de hecho, hace un año estábamos hablando en este mismo foro de generar imágenes a partir de texto, estábamos hablando de GPT3, de modelos multimodales, tuneamos el modelo Stable Diffusion con imágenes propias, y hablamos del protagonismo que estaba adquiriendo OpenAI, Deep Mind, Stability AI y otras.



Y se dio la coincidencia de que una semana después se publicó chatGPT, y con ello se abrió la caja de pandora de los grandes modelos del lenguaje que llevaban dos años ebullendo en el mundo de los desarrolladores pero no habían saltado al conjunto de la sociedad. Desde entonces ha llovido bastante en este campo y sin duda seguirá lloviendo abundantemente.

Hace tan solo una semana OpenAI ha lanzado algunas novedades en su evento Developer Day y con ello abre nuevos horizontes.

La pregunta es ¿Por qué detenernos en un modelo, SAM, que hace algo superado hace ya unos cuantos años?





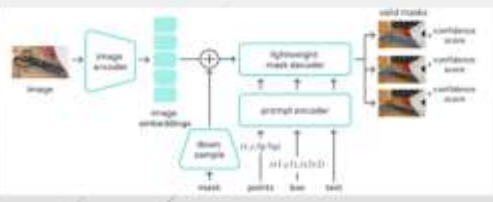
He traído este modelo aquí por dos razones:

-> La primera, porque me permite ilustrar, desde mi modesta experiencia, qué pasa desde que se publica un modelo por parte de una de las grandes tecnológicas, hasta que un usuario final lo puede utilizar para algo útil.

-> La segunda razón es que no hay que minusvalorar algunos modelos que, en un primer vistazo, no parece que vayan a ser útiles, porque si nos pueden servir productos derivados de ellos o está en nuestra mano adaptarlos a nuestras necesidades.

## SAM: A Image Segmentation Foundation Model

Segment Anything Model by Facebook's Meta AI Research



Cornell University

**arXiv** - cs - arXiv:2304.02643

Computer Science > Computer Vision and Pattern Recognition

(Submitted on 5 Apr 2023)

**Segment Anything**

---


Product · Solutions · Open Source · Pricing

[facebookresearch / segment-anything](https://facebookresearch.com/segment-anything) Public

Code · Issues · API · Pull requests · Actions

¿Qué pasa cuando se publica un modelo?  
5/4/2023 → 15/11/2023

# Segment Anything | Meta AI



SA-1B Dataset Explorer

Access the full dataset

10 GB

Antes de hablar de lo que pasa con este modelo desde que se publica hasta que se crean herramientas para el usuario final es conveniente destacar cuatro particularidades de este modelo:

- > La primera es que, según Meta, se trata de un modelo fundacional, es decir, es la madre de todos los modelos de segmentación de imágenes. Hay que decir que hasta cierto punto.
- > La segunda, que se publica en abierto, pero abierto de verdad, ya que hay varias formas de entender lo que significa un modelo abierto. Digo que este se publica en abierto porque se publica la arquitectura del modelo, el modelo entrenado con los correspondientes parámetros y código para reproducirlo y trabajar con él o con las variantes que queramos.
- > La tercera, que se publica el dataset con el que ha sido entrenado, que es un conjunto de 11 millones de imágenes y mil millones de máscaras, y generar ese dataset y publicarlo no es algo accesorio sino bastante meritorio.
- > La cuarta, que alrededor de este modelo hay bastante actividad por parte de la comunidad de desarrolladores y eso facilita mucho su uso.



## Meta publica SAM -> ¿Que pasa después?



La siguiente pregunta es ¿qué pasa después de que se publica un modelo? Cuando se publica uno como éste en abierto, todo el mundo se pone a jugar con él:

-> Desde el mundo académico esto da lugar a la publicación de artículos científicos.

-> También es muy activa la comunidad de desarrolladores que hacen pública gran cantidad de código y algunos datasets. Se trata mayoritariamente de material libre, disponible en plataformas abiertas como Huggingface, un sitio fantástico para la inteligencia artificial, con modelos, datos, tutoriales, demos, etc.

-> Aquí también juegan un papel fundamental los repositorios como GitHub, que todo el mundo conoce.

-> Finalmente surgen las herramientas destinadas al usuario final. Herramientas que pueden venir de la comunidad de desarrolladores o podemos elaborarlas o promoverlas cualquiera de nosotros.

Yo invito a que seamos conscientes de esta secuencia de acontecimientos y que intervengamos en el momento más oportuno.

Para ayudar a ser conscientes de este proceso, comento algunas cuestiones sobre cada fase de la secuencia.



Respecto a los papers, la gente tiene muchas ganas o necesidad de escribir y publicar. En pantalla nuestro 20 artículos científicos relativos a este modelo, y no he buscado lo publicado este último año, sino que todos estos papers, que hablan de SAM, han sido remitidos para publicar en el último mes. Hay que ser conscientes que los artículos científicos a veces son difíciles de digerir, entre otras cosas, porque el soporte papel/pdf tiene sus limitaciones.



Hugging Face Search models, datasets, users...

Models Datasets Spaces Docs Solutions Pricing

**The AI community building the future.**  
The platform where the machine learning community collaborates on models, datasets, and applications.

Datasets 77,782  
Models 4,272

Computer Vision

- Depth Estimation
- Image Classification
- Object Detection
- Image Segmentation
- Image-to-Image
- Unconditional Image Generation
- Video Classification
- Zero-Shot Image Classification

## Model Card for Segment Anything Model (SAM)

Overview

- SamConfig
- SamVisionConfig
- SamMaskDecoderConfig
- SamPromptEncoderConfig
- SamPrompter
- SamImageProcessor
- SamModel
- T2iSamModel

Resources:

- [Demo notebook](#) for using the model.
- [Demo notebook](#) for using the automatic mask generation pipeline.
- [Demo notebook](#) for inference with MedSAM, a fine-tuned version of SAM on the medical domain.
- [Demo notebook](#) for fine-tuning the model on custom data.

Te damos la bienvenida a Colab

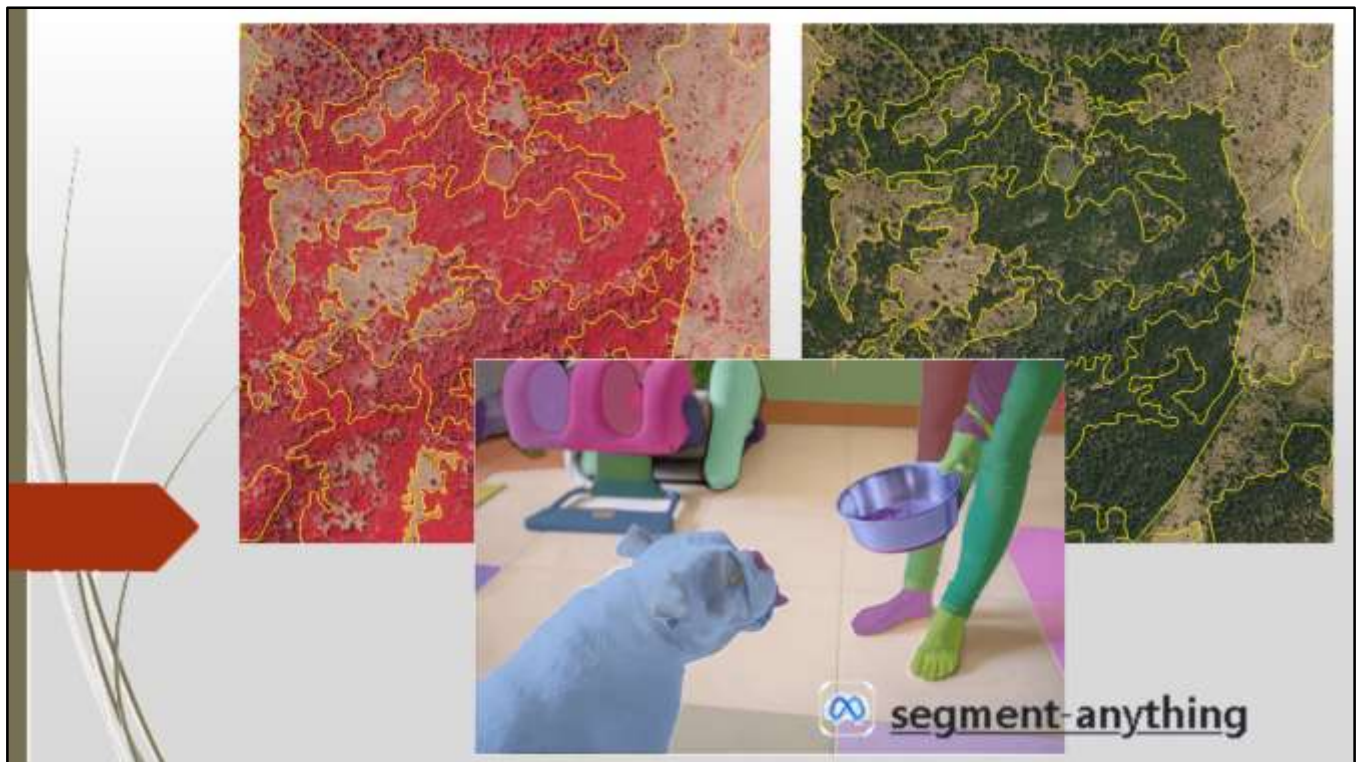
Por suerte, además del mundo académico, está la comunidad de desarrolladores, que aporta herramientas y facilita la digestión de todo lo que hay alrededor de los modelos de IA. En Huggingface encontramos fichas resumen, demos, ejemplos de uso con su código, datasets, etc.



Todo esto acaba en plataformas como GitHub, donde podemos encontrar el código que hemos comentado y mucho más, discusiones relacionadas con ese código, instrucciones, librerías, etc.

Digamos, que la cosa ya está pre-cocinada, a nuestra disposición, y este es un buen momento para que desde nuestras instituciones promovamos iniciativas para sacar partido a estos recursos.

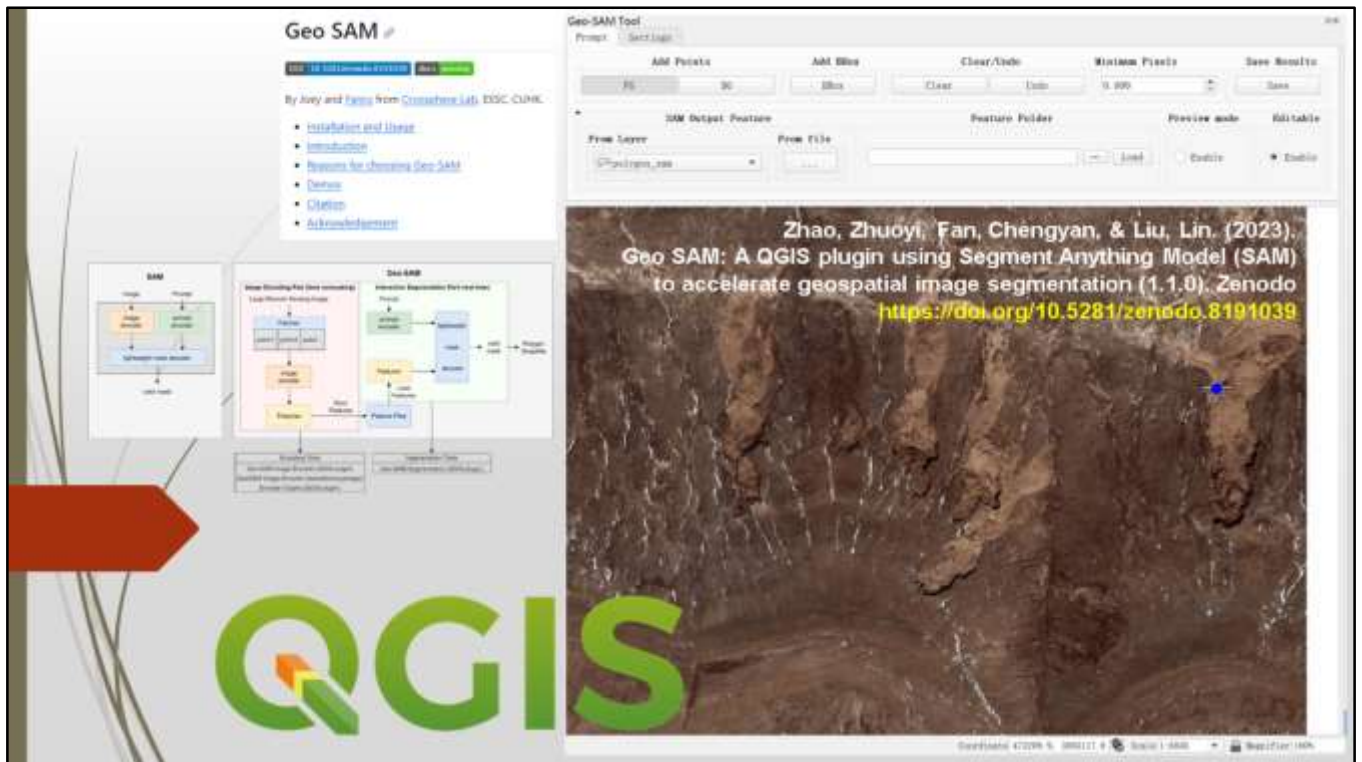




En todo caso, el usuario final quiere herramientas de uso inmediato, eficiente y sencillo. Veamos cómo podemos relacionar el modelo SAM con nuestra actividad científica y profesional:

Meta dice que su modelo lo segmenta todo, pero... puede distinguir objetos o tipos de cubierta en una ortofoto?

A eso nosotros lo llamamos fotointerpretar o, si lo hacemos en una imagen de satélite también lo llamamos segmentar.



Con esa idea dos técnicos de una universidad de Hong Kong prepararon un complemento de Qgis basado en SAM y lo hicieron público en el repositorio de plugins. Aunque es seguro que ellos tienen necesidades distintas a las nuestras, ya que pertenecen a un área de trabajo muy específica (Cryosphere Lab de la Universidad China de Hong Kong), su complemento nos sirve para analizar hasta qué punto SAM también nos puede ser útil a nosotros en tareas de cartografía de tipos de cubierta, hábitats o ecosistemas a partir de ortofotos y/o imágenes de satélite. El complemento se llama GeoSAM y lo que hace es usar el modelo de Meta y adaptarlo para hacerlo más ágil pre-procesando las imágenes para que la inferencia se haga en milisegundos, con lo que convierten a SAM en una herramienta interactiva dentro de Qgis.





Después de probar el complemento podemos llegar a la conclusión de que GeoSAM no nos sirve porque no buscamos una herramienta interactiva o porque la segmentación que hace no nos convence.

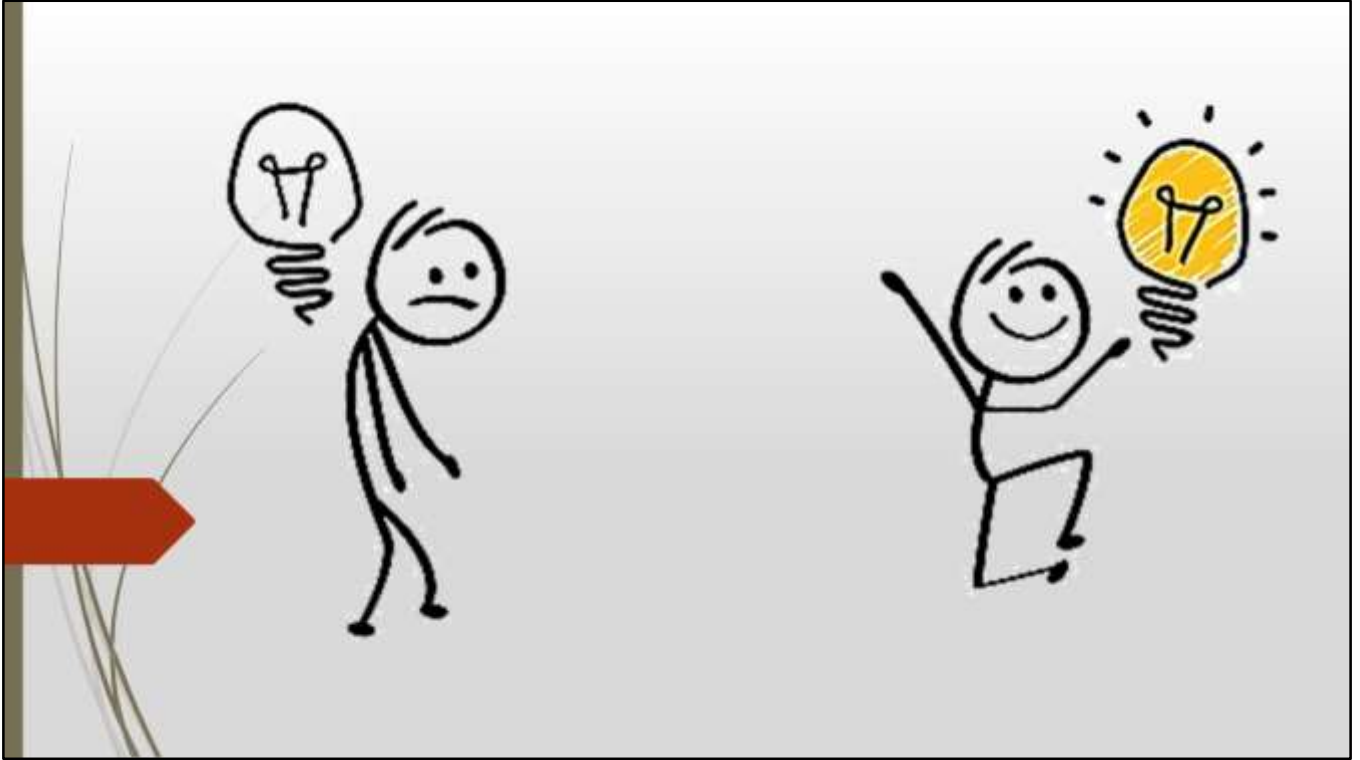
El propio modelo SAM invita a esta óptica interactiva porque está diseñado para segmentar en respuesta a un *prompt* (un punto o un rectángulo en la imagen) y se apoya en un entrenamiento con imágenes que no son fotografías aéreas, por lo que a priori no es el modelo más orientado hacia la fotointerpretación.



¿Significa eso que SAM y GeoSAM no nos sirven? Mi respuesta es el mensaje principal de esta presentación:

**Que un modelo de IA o una herramienta que lo pone en servicio no nos sirvan en primera instancia no implica que no los podamos adaptar o reciclar para nuestras necesidades concretas.**



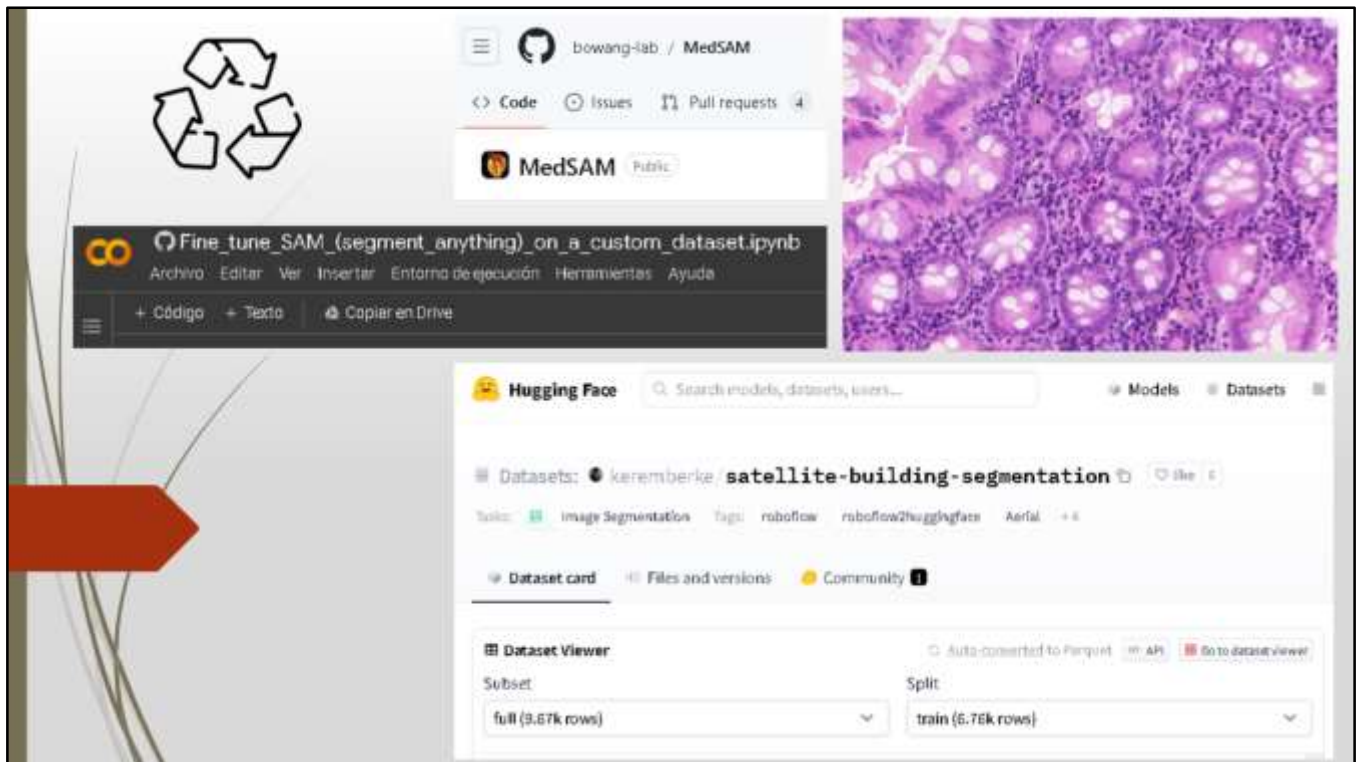


Invito a que desde nuestra condición de administraciones gestoras o centros tecnológicos o de investigación contribuyamos a adaptar o crear herramientas que satisfagan nuestras necesidades o las de otros usuarios en nuestro campo.

Elaborar un complemento de Qgis (o la herramienta que estimemos oportuna) está muy al alcance de cualquiera de nuestras instituciones, pero

¿Nos puede ayudar SAM en tareas de fotointerpretación?

La respuesta la encontramos haciendo una prueba de fine-tuning del modelo.



Para hacernos una idea rápida de si podemos mejorar la capacidad de fotointerpretación de SAM con fine-tuning podemos reciclar recursos disponibles en internet:

-> Código ya escrito con otros objetivos:

\* P. ej. NielsRogge/Transformers-

Tutorials/blob/master/SAM/Fine\_tune\_SAM\_(segment\_anything)\_on\_a\_custom\_dataset.ipynb

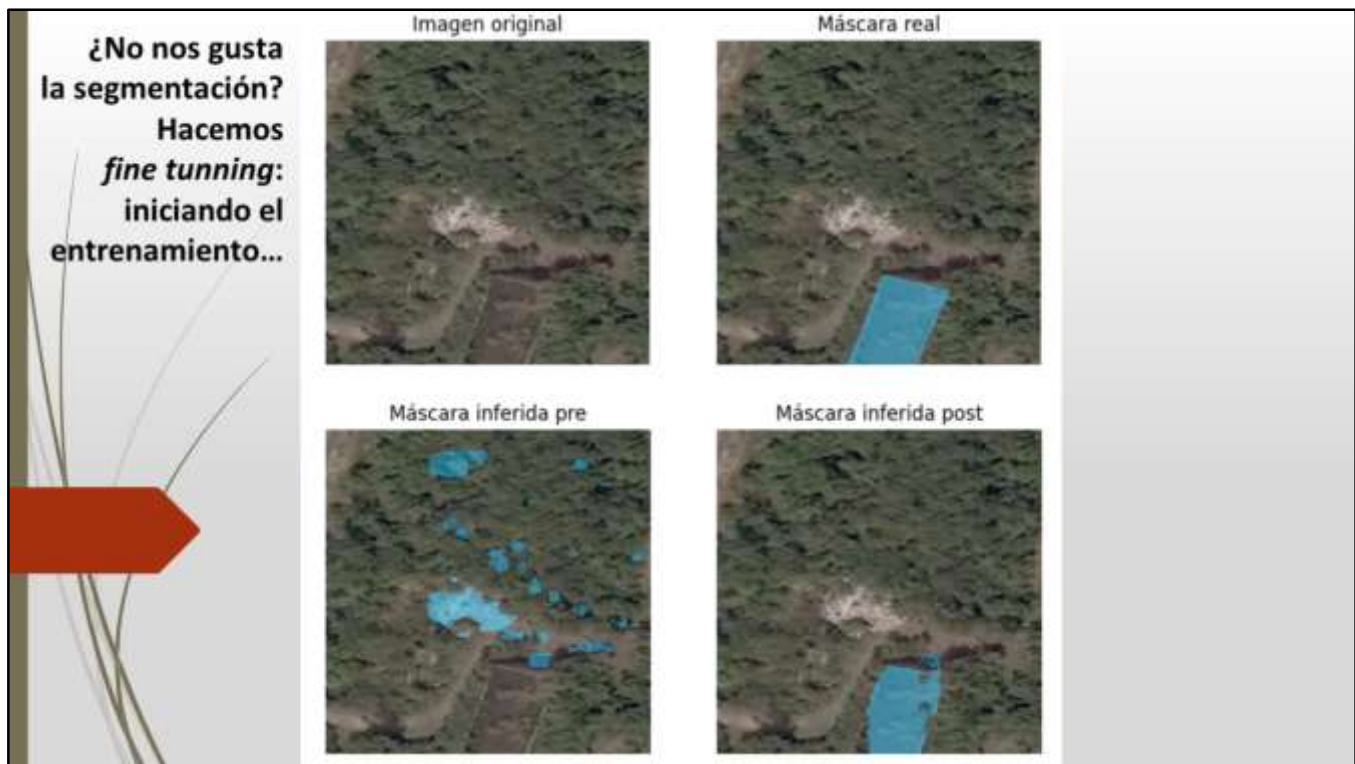
basado en <https://github.com/bowang-lab/MedSAM>)

-> Datasets de Huggingface:

\* P. ej. keremberke/satellite-building-segmentation, que tiene imágenes aéreas y máscaras de edificios

Nota: elaborar un dataset de este tipo en España es especialmente sencillo porque contamos con toda la información que necesitamos

pero, para una prueba de fine-tuning, nos vale este dataset ya preparado.



Esta prueba está hecha esta misma semana y el entrenamiento ejecutado en la herramienta gratuita colab de google.  
Tras unos pocos ciclos de entrenamiento con una pequeña muestra del dataset de entrenamiento, se puede comprobar que el tipo de segmentación se adapta mejor a nuestras necesidades.



***Fine tuning:***  
el descenso del  
gradiente  
responde


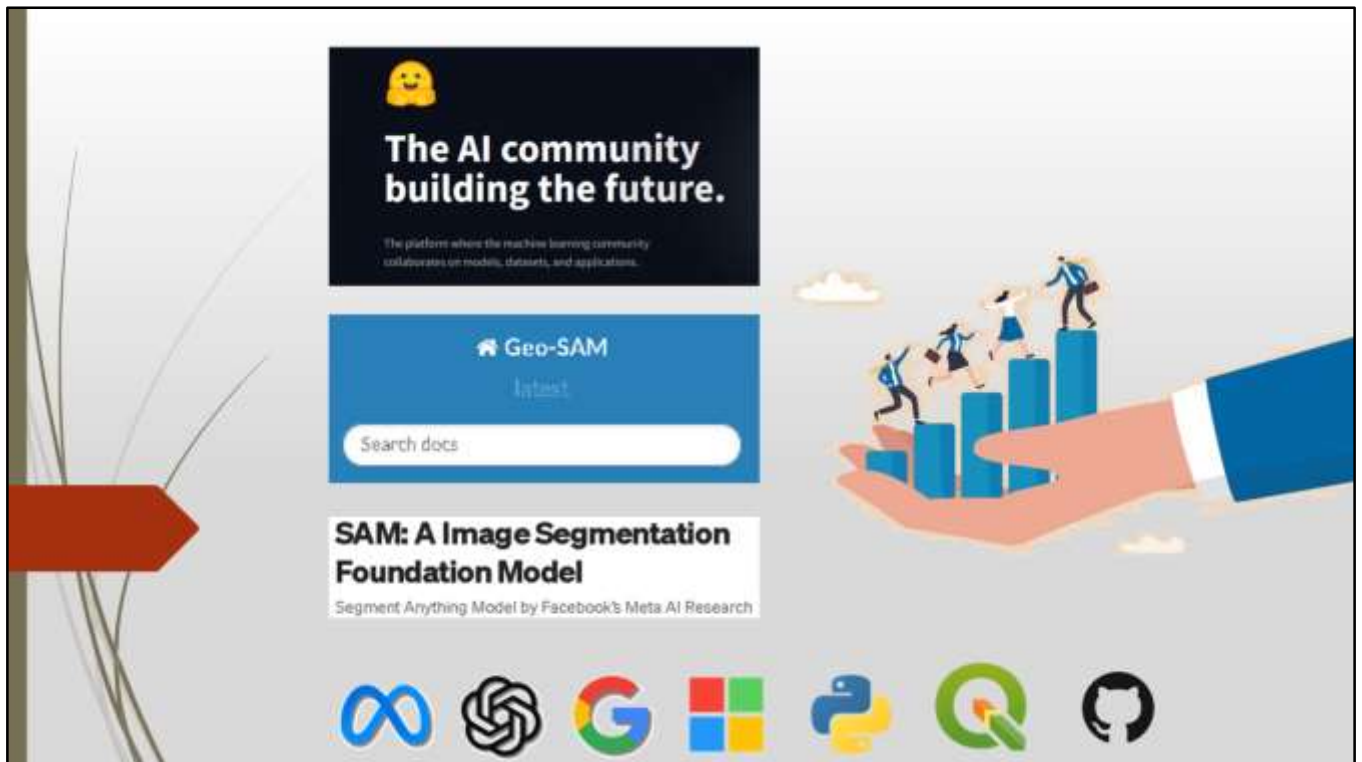


Imagen original	Máscara real
Máscara inferida pre	Máscara inferida post

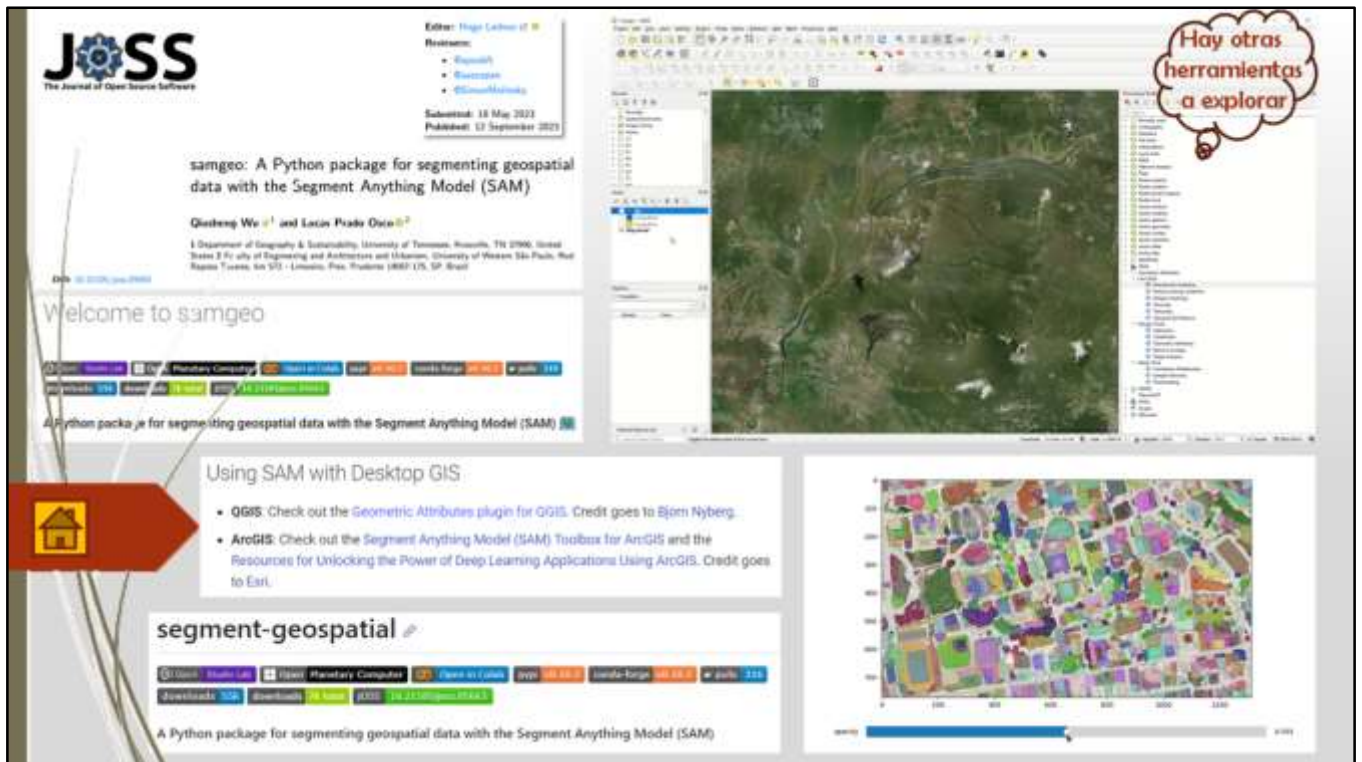
**Conclusión:**  
podemos adaptar  
el modelo y las  
herramientas a  
nuestras  
necesidades

Tras un segundo entrenamiento, apenas quince minutos, se puede comprobar que hay claras mejorías en la forma en que el modelo segmenta los edificios.  
Conclusión: aquí hay recorrido, hay margen de mejora.



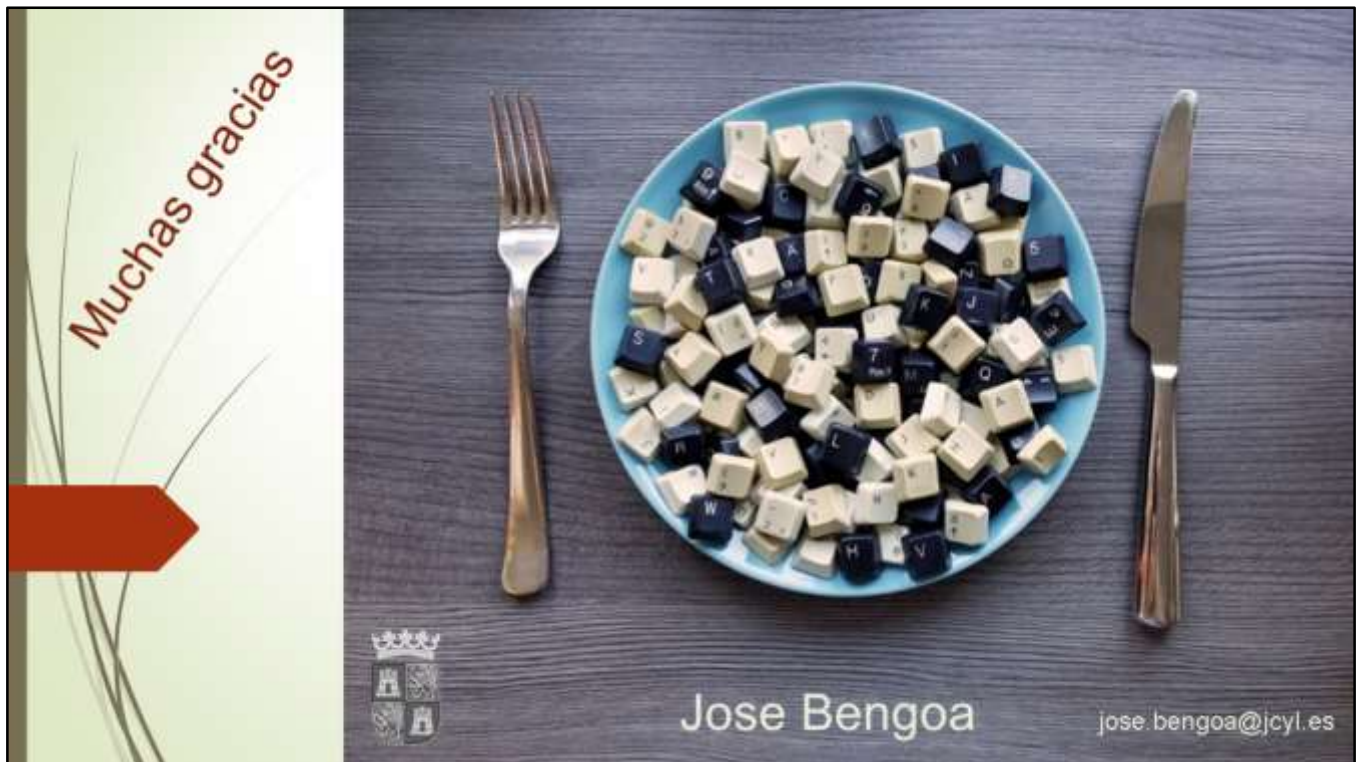
GeoSAM nos ha ayudado a abrir la perspectiva con que miramos al modelo SAM porque nos ayuda a definir mejor el tipo de herramienta que buscamos y, lo que es aún más importante, tenemos un escalón ganado en ese proceso ya que en este campo todo se puede reciclar y unos desarrollo se apoyan siempre en otros previos.

Así se construye en el mundo del código abierto y GeoSAM es un peldaño.



Ayer mismo, mientras ejecutaba el fine-tuning, me encontré con esta otra herramienta (samgeo) que apunta en esta misma dirección y que tiene su correspondiente plugin de Qgis. Esta herramienta fue publicada apenas un mes después de la salida de SAM, reflejo del enorme dinamismo que podemos encontrar en este campo. Queda pendiente explorar samgeo, analizar su funcionalidad y ver si han hecho un fine-tuning previo.





En resumen, el mensaje de esta presentación es doble:

Por un lado recordar que están pasando trenes, que son modelos de IA delante de nosotros y el que no encontremos una utilidad directa de muchos de ellos o no se adapten exactamente a nuestras necesidades no significa necesariamente que debemos mirar a otro lado, sino que es posible trabajar con ellos para traerlos a nuestro terreno.

Por otro lado, y con esto reitero el mensaje del año pasado, que podemos adaptar los grandes modelos de IA o los entrenados por algunas de las Big Tech a nuestras necesidades.

Cuando hablé con Paco sobre mi posible participación en estas jornadas, le comenté que tenía esta idea pero los productos no estaban maduros, y que era mejor descartar este contenido para las jornadas, porque sería una presentación un tanto pobre.

Pero después entendí el interés del asunto y de ahí la orientación de esta charla; en este campo es probable que nos encontremos con recursos que no están del todo cocinados pero eso no es impedimento para que nos impliquemos y contribuyamos a llevar estas técnicas de la cocina a la mesa de nuestros usuarios finales.

Muchas gracias.