

## Tipologia i cicle de vida de les dades. Màster de Ciència de Dades.

### Alumnes:

J. de Curtò i Díaz. [decurto@uoc.edu](mailto:decurto@uoc.edu)

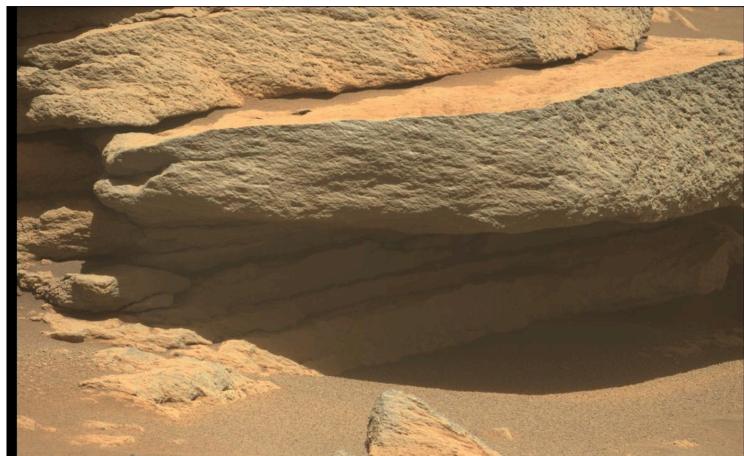
I. de Zarzà i Cubero. [dezarza@uoc.edu](mailto:dezarza@uoc.edu)

## Pràctica 1. Dataset.

### 1. Context.

La informació s'ha recollit en el context de les missions a Mart de la NASA (Spirit [1], Opportunity [2], Curiosity [3] i Perseverance [4,5]). Es tracta d'imatges captades per les càmeres a bord dels respectius robots (rover, lander i en el cas de la missió de 2020, també el dron Ingenuity) al llarg de la seva visita al cos celeste. La informació es troba pública a les corresponents pàgines però la dificultat escau en l'obtenció automatitzada i massiva de les dades. De fet, no es disposa d'una API concreta per fer la descàrrega i només és públic un entorn gràfic que permet la visualització de les imatges de manera seqüencial. La publicació en obert d'aquestes dades pot ser particularment valuosa per fer recerca en visió per computador (aprenentatge automàtic i profund per la navegació autònoma) i també en geofísica i astrofísica per l'anàlisi de les característiques del terreny (per exemple per crear simuladors acurats o entendre bé les característiques morfològiques).

### Mars Perseverance Sol 242: Left Mastcam-Z Camera



#### Download

1648 x 1200  
image/png

[DOWNLOAD](#)

#### Embed

Embed this resource by pasting the following code into your website:

```
<iframe
src="https://mars.nasa.gov/layout/embed/image/500raw/?i=ZLO_0242_0688424944_098EBY_N0072440ZCAM03246_1100LMJ"
width="500" height="500"
scrolling="no" frameborder="0">
```

### 2. Títol.

Títol: CyZ: MARS Space Exploration Dataset.

Subtitol: Images from NASA missions of the celestial body.

Títol: CyZ: Dataset d'exploració espacial de Mart.

Subtítol: Imatges de les missions de la NASA del cos celeste.



### 3. Descripció del data set.

El dataset conté imatges captades per les càmeres del rover i el lander de les respectives missions de la NASA a Mart. Cada imatge té el seu respectiu metadata incorporat; seguint la corresponent nomenclatura establerta per la NASA al títol de la imatge. En el cas de la missió Perseverance, també s'han recopilat imatges captades pel dron Ingenuity. També s'ha fet l'extracció d'altra informació textual (vegeu figura adjunta) com per exemple dades de posicionament donat el sol (dia marcíà) en què es va registrar la imatge.

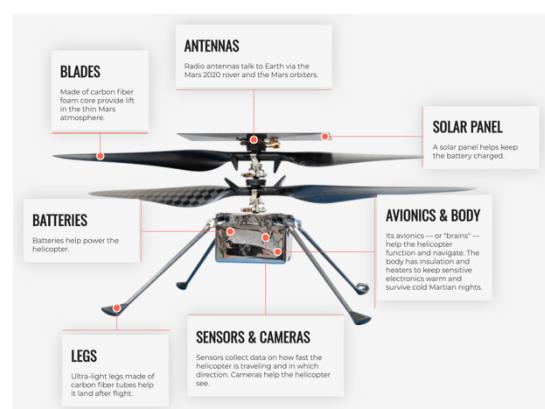
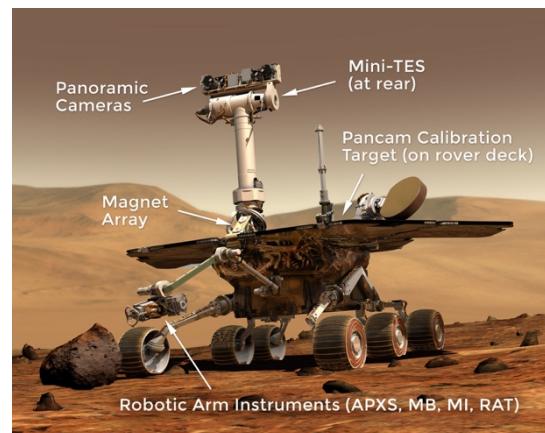
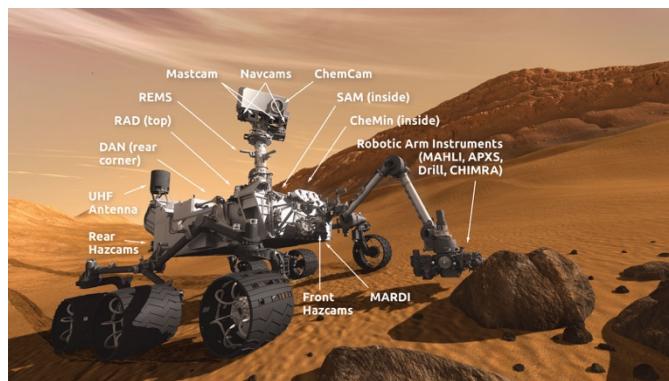


Flight Log

Flight	Sol	Date	Horizontal Distance		Max. Altitude		Max. Groundspeed		Duration	Route of Flight	
			m	ft	m	ft	m/s	mph		From	To
1	58	April 19, 2021	0	0	3	~10	0	0	39.1	Wright Brothers Field	
2	61	April 22, 2021	4	~13	5	~16	0.5	~1	51.9		
3	64	April 25, 2021	100	~328	5	~16	2	~4.5	80.3		
4	69	April 30, 2021	266	~873	5	~16	3.5	~8	116.9		
5	76	May 7, 2021	129	~423	10	~33	2	~4.5	108.2		
6	91	May 22, 2021	215	~705	10	~33	4	~9	139.9		
7	107	June 8, 2021	106	~348	10	~33	4	~9	62.8		
8	120	June 21, 2021	160	~525	10	~33	4	~9	77.4		
9	133	July 5, 2021	625	~2051	10	~33	5	~11	166.4		

#### 4. Representació gràfica.

A continuació s'inclouen diagrames en què es pot observar els diferents instruments de captura de dades dels corresponents rovers i el dron per entendre millor la composició visual del dataset. També es mostra als diagrames i taules el nombre total d'imatges captades per cadascuna de les corresponents càmeres i que s'inclouen al dataset (\*per incloure en la versió final del document\*).



#### 5. Contingut.

Les dades capturades són essencialment imatges captades per les càmeres. Aquestes imatges estan diferenciades pel dia marcià (sol) en què es van enregistrar pels sensors. Cada rover disposa de diferents tipus de càmeres, amb diferents formats: color i escala de grisos; i normalment estèreo (imatge esquerra i dreta). El dataset també contindrà un CSV amb les URL de descàrrega de les imatges i el metadata extret en format text. El codi possibilitarà també la descàrrega massiva i l'emmagatzematge en cas que no fem públiques la totalitat de les dades o es necessiti fer descàrrega de noves imatges penjades posteriorment a la publicació del dataset.

#### 6. Agraïments.

Les dades són obertes i han estat obtingudes per la NASA. En cas d'ús s'ha de fer cita corresponent a la missió de Mart on es van capturar.

## 7. Inspiració.

L'objectiu és fer ús d'aquestes dades en la nostra recerca com a investigadors en visió per computador. Les imatges són valuoses en el sentit que és difícil obtenir massivament captures reals (i en general tota classe de dades dels sensors) fora del nostre planeta [6,7] que permeten entrenar algoritmes d'aprenentatge automàtic i profund per la navegació autònoma en futures missions a Mart. També és interessant per millorar els simuladors que possibiliten les proves de hardware i de les condicions visuals que es donaran en les missions reals.



## 8. Llicència.

Utilitzarem una llicència CC BY-SA 4.0 pel nostre codi i dades generades que permeten prèvia citació de la nostra feina l'ús de les dades en el context més ampli possible. Les imatges són propietat de la NASA però no tenen copyright en si, per tant es poden usar per fer ús educacional i recerca sense problemes a condició que se citin les fonts d'obtenció.

<https://creativecommons.org/licenses/by-sa/2.0/>

## 9. Codi.

<https://github.com/decurtoidiaz/cyz> (encara no disponible)

## 10. Dataset.

A publicar al zenodo <https://zenodo.org/> amb el nom:

CyZ: MARS Space Exploration Dataset.

## Referències:

- [1] Spirit: <https://mars.nasa.gov/mer/gallery/all/spirit.html>
- [2] Opportunity: <https://mars.nasa.gov/mer/gallery/all/opportunity.html>
- [3] Curiosity <https://mars.nasa.gov/msl/multimedia/raw-images/>
- [4] Perseverance: <https://mars.nasa.gov/mars2020/multimedia/raw-images/>

[5] Maki et al. 2020. The Mars 2020 Engineering Cameras and Microphone on the Perseverance Rover: A Next-generation Imaging System for Mars Exploration.

<https://link.springer.com/article/10.1007/s11214-020-00765-9>

[6] Lamarre et al. 2020. The Canadian Planetary Emulation Terrain Energy-Aware Rover Navigation Dataset.

<https://starslab.ca/enav-planetary-dataset/>

[7] ESA Robotics Dataset. 2015. Katwijk Beach Planetary Rover Dataset.

<https://robotics.estec.esa.int/datasets/katwijk-beach-11-2015/>