



Towards an **exoplanet census** in the **cloud**



Néstor Espinoza

In collaboration with: Rachel Cooper (STScI) &
Matthew Bourque (STScI)



github.com/nespinoza



twitter.com/nespinozap

1

(De?)

A motivational story

i.e., the “exciting” life of an exoplaneteer



(The first transiting exoplanet!)

**“ I'm interested in HD209458 *b*.
Most up-to-date properties?**

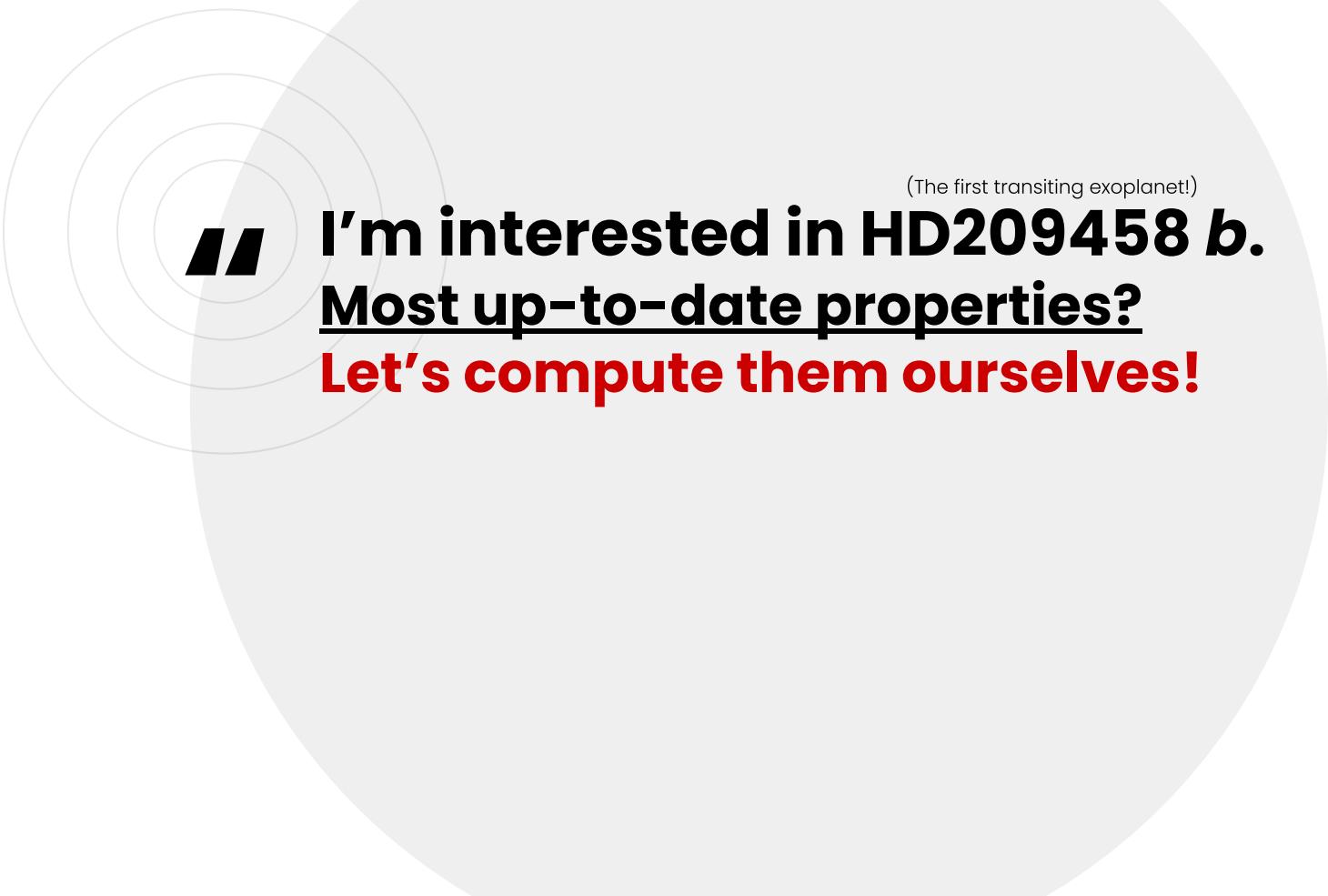
(The first transiting exoplanet!)

I'm interested in HD209458 b. Most up-to-date properties?



NASA EXOPLANET ARCHIVE
NASA EXOPLANET SCIENCE INSTITUTE

Planet	Planet Parameters						Reference		
	(Jupiter Mass)	(Earth Mass)	(Jupiter Mass)	(Earth Mass)	(Jupiter Radii)	(Earth Radii)			
b	null	0.685 ^{+0.014} _{-0.011}	217 ⁺³ ₋₃	1.359 ^{+0.018} _{-0.017}	15.20 ^{+0.18} _{-0.17}	0.333 ^{+0.014} _{-0.013}	null	Bonomo et al. 2017	
b	null	0.734 ^{+0.04}	232 ⁺¹³	1.38 ^{+0.02}	15.6 ^{+0.2}	0.362 ^{+0.014}	null	Stassun et al. 2017	
b	null	null	null	null	null	1448	null	Barstow et al. 2017	
b	null	0.743 ^{+0.06}	235 ⁺¹⁹	1.41 ^{+0.06}	15.8 ^{+0.7}	0.32 ^{+0.05}	null	del Burgo & Allende Prieto 2016	
b	null	0.69	219	1.36	15.2	0.450	null	Sing et al. 2016	
b	null	null	null	null	null	1484 ⁺¹⁸	null	Evans et al. 2015	
b	null	0.69	219.29	1.4	15.7	0.34	1450	Saito et al. 2015	
b	null	0.643 ^{+0.09}	203 ⁺⁴²	1.451 ^{+0.074}	16.264 ^{+0.074}	0.261 ^{+0.044}	null	Burgasser et al. 2015	
b	null	null	null	null	null	1450	null	Albrecht et al. 2015	
b	null	0.714 ^{+0.017}	227 ⁺⁵	1.380 ^{+0.017}	15.476 ^{+0.19}	0.362 ^{+0.006}	1459 ⁺¹²	null	Southworth 2010
b	null	0.685 ^{+0.015} _{-0.012}	217.70 ^{+47.07}	1.359 ^{+0.018} _{-0.017}	15.233 ^{+0.177}	0.338 ^{+0.016}	1449 ⁺¹²	Torres et al. 2008	
b	null	null	null	1.320 ^{+0.024}	14.796 ^{+0.269}	0.318 ^{+0.018}	null	Knutson et al. 2007	
b	0.699 ^{+0.007}	222.154 ^{+2.25}	null	null	null	null	null	Naeff et al. 2004	
b	null	0.685 ^{+0.020}	218 ⁺⁶	1.42 ^{+0.13}	15.9 ^{+1.1}	0.31 ^{+0.04}	null	Cody & Sasselov 2002	
b	null	null	null	1.347 ^{+0.060}	15.10 ^{+0.67}	0.31 ^{+0.018}	null	Brown et al. 2001	
b	null	0.624 ^{+0.05}	197.05 ^{+15.89}	1.42 ^{+0.10}	15.92 ^{+1.12}	0.27 ^{+0.04}	null	Henry et al. 2000	
b	null	null	0.63	200.22	1.27 ^{+0.02}	14.24 ^{+0.22}	0.38	Charbonneau et al. 2000	

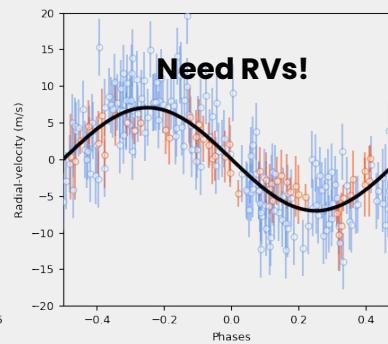
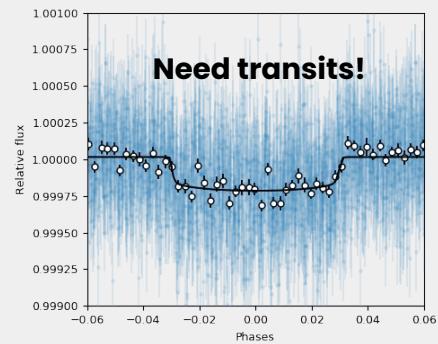


(The first transiting exoplanet!)

**“ I'm interested in HD209458 *b*.
Most up-to-date properties?
Let's compute them ourselves!**

(The first transiting exoplanet!)

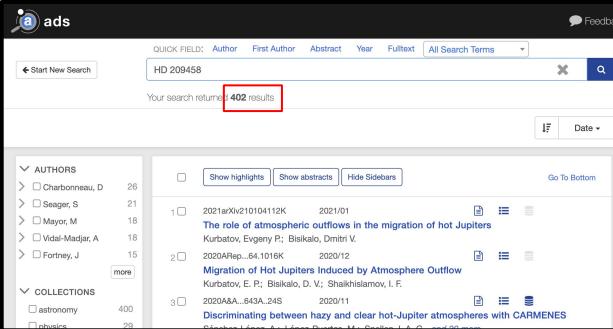
“ I’m interested in **HD209458 b.**
Most up-to-date properties?
Let’s compute them ourselves!



(The first transiting exoplanet!)

“ I’m interested in HD209458 *b*.
Most up-to-date properties?
Let’s compute them ourselves!

(ADS = not good to get datasets)



The screenshot shows the ADS search interface with the query "HD 209458". The search results page displays 402 results. The left sidebar shows filters for "AUTHORS" (including Charbonneau, D; Seager, S; Mayor, M; Vidal-Madjar, A; Fortney, J) and "COLLECTIONS" (including astronomy, 400; physics, 29). The main results table lists three papers:

Publication	Year	Authors
2021arXiv210104112K	2021/01	Kurkhatov, Evgeny P.; Biskalo, Dmitri V.
2020ARep...64.1016K	2020/12	Kurkhatov, E. P.; Biskalo, D. V.; Shaikhislamov, I. F.
2020AAA...643A..24S	2020/11	Seager, S.; Rauscher, M.; Coughlin, J. L.; et al.

(The first transiting exoplanet!)

I'm interested in HD209458 b. Most up-to-date properties? Let's compute them ourselves!

(ADS - not good to get datasets)



(NASA Exoplanet Archive = awesome, but overwhelming!)

A screenshot of the NASA Exoplanet Archive search interface. The search bar at the top contains 'HD209458'. The results page shows a table of datasets for HD209458 b. One specific dataset is highlighted with a red box: 'HD 209458 b (Borucki et al. 2013)'. The table includes columns for 'Type', 'Start Time', 'End Time', 'Number of Data Points', 'Wavelength', 'Instrument/Source', 'Links', 'File', and 'Reference'. The 'File' column for the highlighted dataset shows 'OCO - MODIS'. A large red 'X' is drawn across the entire screenshot.

(The first transiting exoplanet!)

“ I’m interested in HD209458 b. Most up-to-date properties? Let’s compute them ourselves!

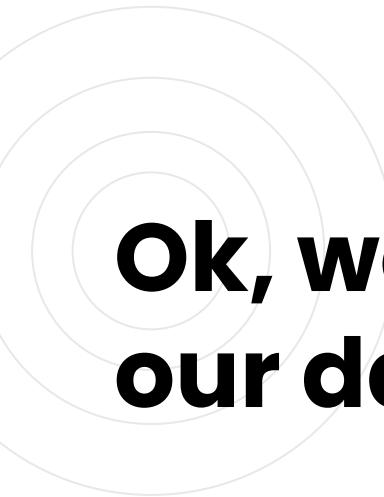
(ADS = not good to get datasets)



The screenshot shows the NASA Exoplanet Archive search interface for the query "HD209458". The search results page displays a list of datasets, with a red box highlighting the "402 results" link at the top of the list.

(NASA Exoplanet Archive = awesome, but overwhelming!)

This is a >100 parameter fit!



**Ok, we have
our data!**

EXOFAST



Ok, we have our data!



Joint analysis of exoplanetary transits & RVs

J
Juliet

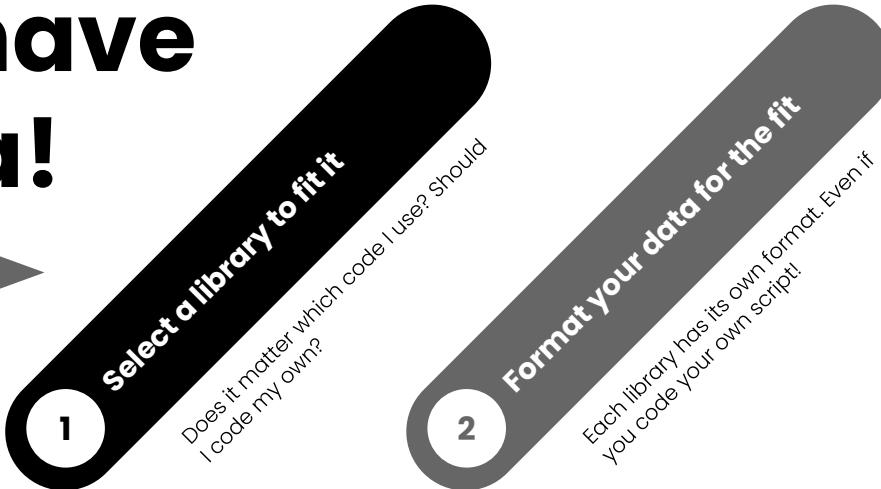
pyaneti

MCMCI



Exo-Striker

Ok, we have our data!

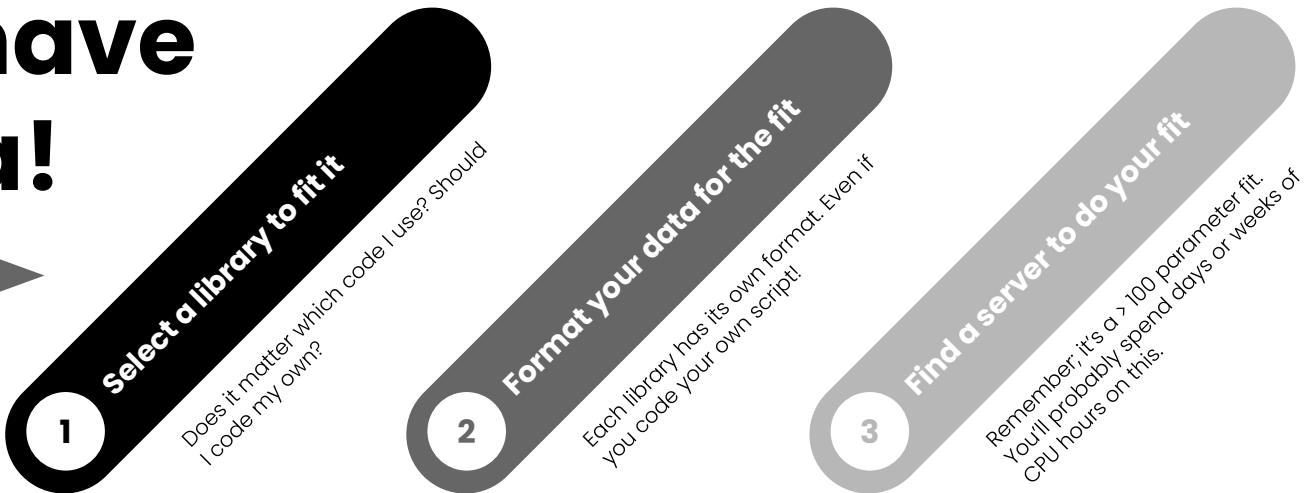


Ok, we have our data!



- 1 Select a library to fit it**
Does it matter which code I use? Should I code my own?
- 2 Format your data for the fit**
Each library has its own format. Even if you code your own script!
- 3 Find a server to do your fit**
Remember, it's $\alpha > 100$ parameter fit. You'll probably spend days or weeks of CPU hours on this.

Ok, we have our data!



This entire process can take **weeks of work** (per system!)

(And good luck if there's a TESS/PLATO/CHEOPS/Gaia data release!)

Ok, we have
our data!

This is unsustainable.

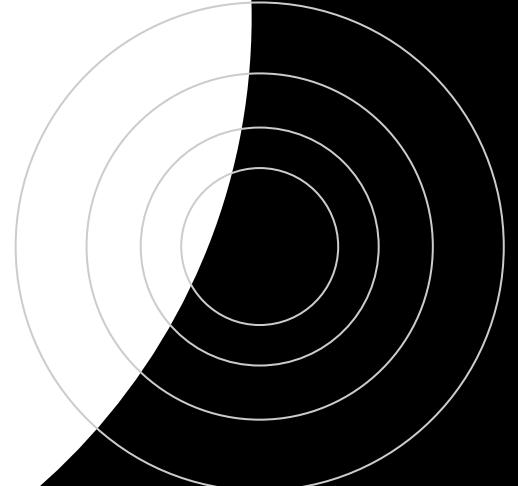
This entire process can take **weeks of work** (per system!)
(And good luck if there's a TESS/PLATO/CHEOPS/Gaia data release!)



2

The dream

An exoplanet census in the cloud





The dream

(explained with Star Wars' Bespin planet)
(a gas giant!)



Service constantly
**generating up-to-date
planetary properties**

The dream



Service constantly
generating up-to-date
planetary properties

The dream

Users **can use their own data**
to generate planetary
properties (e.g., for discovery papers)



Service constantly
**generating up-to-date
planetary properties**

The dream

Hosted & running
in the cloud

Users **can use their own data
to generate planetary
properties** (e.g., for discovery papers)



Service constantly
generating up-to-date
planetary properties

Hosted & running
in the cloud

The dream

Users **can use their own data**
to generate planetary
properties (e.g., for discovery papers)



Service constantly
**generating up-to-date
planetary properties**

Pay as
you go

Hosted & running
in the cloud

The dream

Users **can use their own data**
to generate planetary
properties (e.g., for discovery papers)



Service constantly
**generating up-to-date
planetary properties**

Pay as
you go

Hosted & running
in the cloud

Flexible

Users **can use their own data**
to generate planetary
properties (e.g., for discovery papers)



Service constantly
generating up-to-date
planetary properties

Pay as
you go

Reliable

Hosted & running
in the cloud

Flexible

Users **can use their own data**
to generate planetary
properties (e.g., for discovery papers)



Service constantly
generating up-to-date
planetary properties

The dream

Pay as
you go

Reliable

Hosted & running
in the cloud

"Mobile"

Flexible

Users can use their own data
to generate planetary
properties (e.g., for discovery papers)



Service constantly
generating up-to-date
planetary properties

Pay as
you go

Reliable

Hosted & running
in the cloud

aws

"Mobile"

Flexible

Users can use their own data
to generate planetary
properties (e.g., for discovery papers)



3

What we've been up to

a.k.a., our "technological exploration" phase





3

What we've been up to

a.k.a., our "technological exploration" phase

(The “core” team)
Who are “we”



Matthew Bourque (STScI)

Software Engineer



Rachel Cooper (STScI)

Science Support Analyst



Néstor Espinoza (STScI)

Assistant Astronomer

(Transit, RVs, Nested Samplers --- juliet.readthedocs.io)

What are we doing?

(currently)

Joint analysis of exoplanetary transits & RVs

Juliet



(Transit, RVs, Nested Samplers --- juliet.readthedocs.io)

What are we doing?

Espinoza et al., 2019 (MNRAS, 490, 2262)

The screenshot shows the Juliet documentation website. The left sidebar includes links for Installation, Getting started, Models, priors and outputs, API, Tutorials, and various guides. The main content area displays a code snippet for plotting phase-folded lightcurves, followed by two side-by-side plots showing Relative Flux vs Time (BJD - 2457000) and Phases. A descriptive text below the plots explains the deterministic and GP components of the model.

```
# Now plot phase-folded lightcurve but with the GP part removed:  
ax2.errorbar(phases, dataset.data_lc['TESS'] - gp_model, \br/>             yerr = dataset.errors_lc['TESS'], fmt = '.', alpha = 0.3)  
  
# Plot transit-only (divide by mflux) model:  
idx = np.argsort(phases)  
ax2.plot(phases[idx], transit_model[idx], color='black', zorder=10)  
ax2.yaxis.set_major_formatter(pt.NullFormatter())  
ax2.set_xlabel('Phases')  
ax2.set_xlim([-0.03, 0.03])  
ax2.set_ylim([0.96, 1.04])
```

Looks pretty good! As can be seen, the `results.lc.model['TESS']['deterministic']` dictionary holds the deterministic part of the model. This includes the transit model which is distorted by the dilution factor (set to 1 in our case) and the mean out-of-transit flux, which we fit together with the other parameters in our joint fit — this deterministic model is the one that is plotted in the right panel in the above presented figure. The `results.lc.model['TESS']['GP']` dictionary, on the other hand, holds the GP part of the model — because this is an additive process in this case, we can just subtract it from the data in order to get the "systematic-corrected" data that we plot in the right panel in the figure above.

Global and instrument-by instrument GP models

In the previous lightcurve analysis we dealt with GP models which are individually defined for each



oint analysis of exoplanetary transits & RVs
Juliet

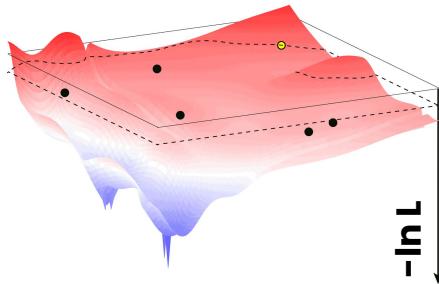


(Transit, RVs, Nested Samplers --- juliet.readthedocs.io)

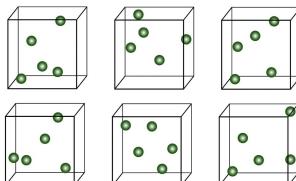
What are we doing?

Why Nested Samplers?

Very good at thoroughly exploring the parameter space. Also expensive.



Unitary cubes (prior)



GIF credit: Noam, Rob, Gabor, Livia ([2015](#))

(currently)



oint analysis of exoplanetary transits & RVs
Juliet



(Transit, RVs, Nested Samplers --- juliet.readthedocs.io)

What are we doing?

Algorithmic efficiency

How to optimally speed-up Nested Samplers with multi-threading?

(currently)



oint analysis of exoplanetary transits & RVs
Juliet



(Transit, RVs, Nested Samplers --- juliet.readthedocs.io)

What are we doing?

Algorithmic efficiency

How to optimally speed-up Nested Samplers with multi-threading?

(currently)

Fundamental for cost estimates!
(Time = money --- literally)



Joint analysis of exoplanetary transits & RVs
Juliet

(Transit, RVs, Nested Samplers --- juliet.readthedocs.io)

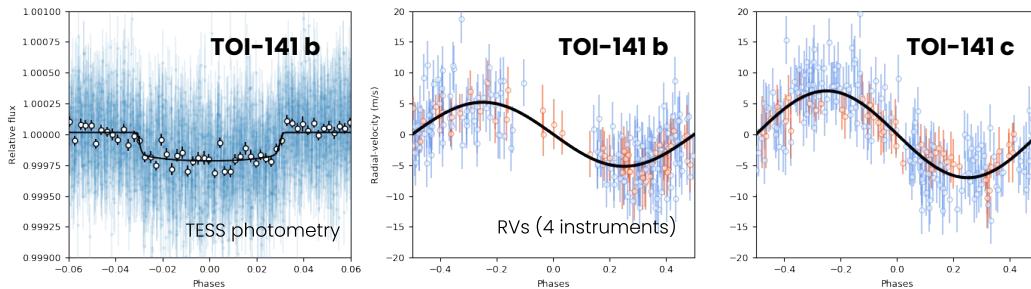
What are we doing?

Algorithmic efficiency

How to optimally speed-up Nested Samplers with multi-threading?

(currently)

Test case: TOI-141 system (Espinoza et al., 2020, MNRAS, 491, 2, 2982)



(+ Gaussian Process on the RVs --- not shown here)

oint analysis of exoplanetary transits & RVs
Juliet

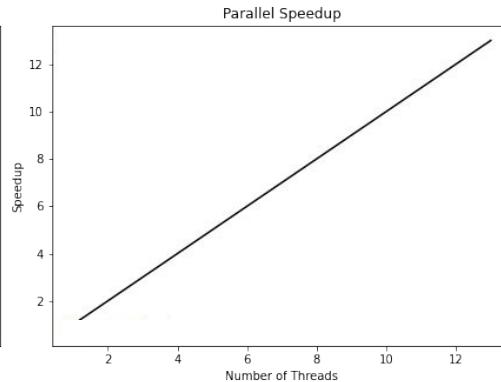
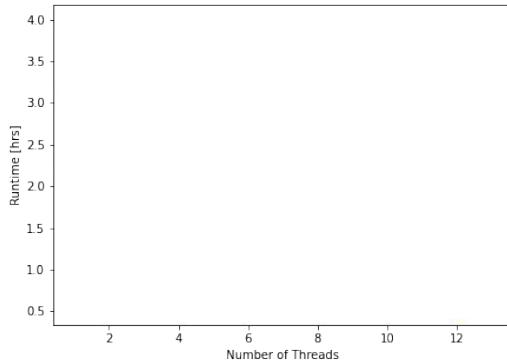


What are we doing?

Algorithmic efficiency

How to optimally speed-up Nested Samplers with multi-threading?

(currently)



oint analysis of exoplanetary transits & RVs

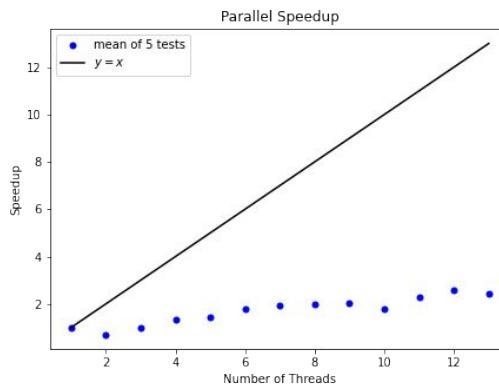
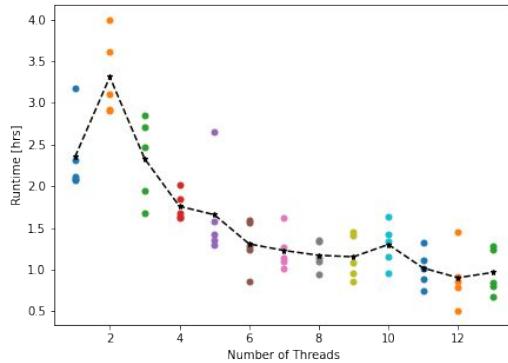
Juliet



What are we doing?

Algorithmic efficiency

How to optimally speed-up Nested Samplers with multi-threading?

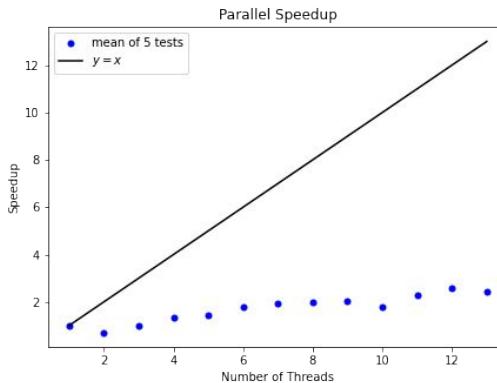
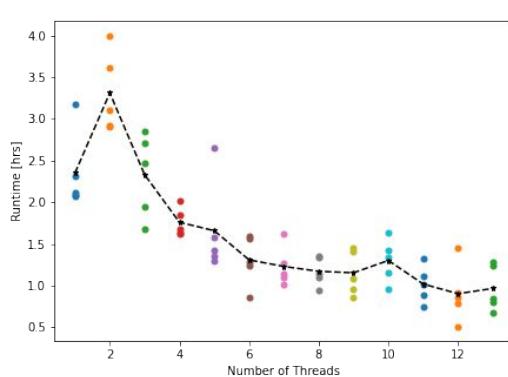


What are we doing?

Algorithmic efficiency

How to optimally speed-up Nested Samplers with multi-threading?

Nested Samplers evaluate \mathbf{N}_{live} times the likelihood

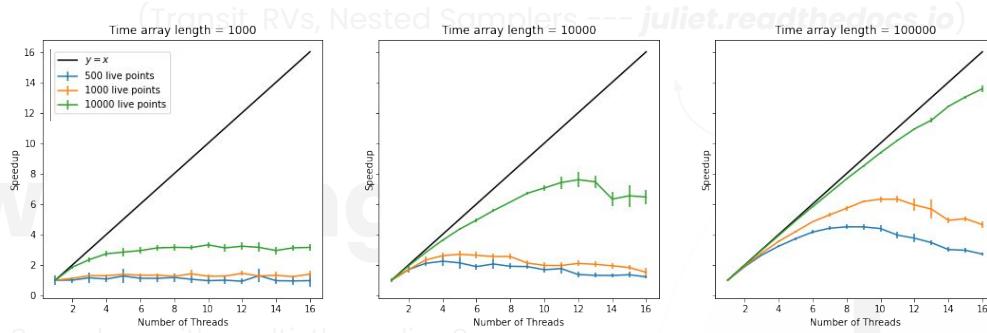


Who are we?

batman
**(i.e., transit
model evaluation)**

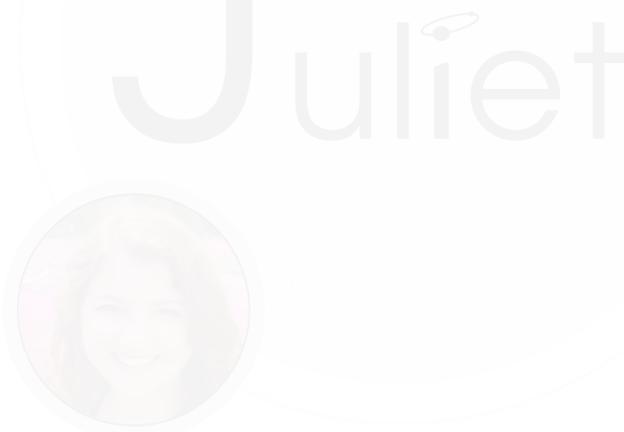
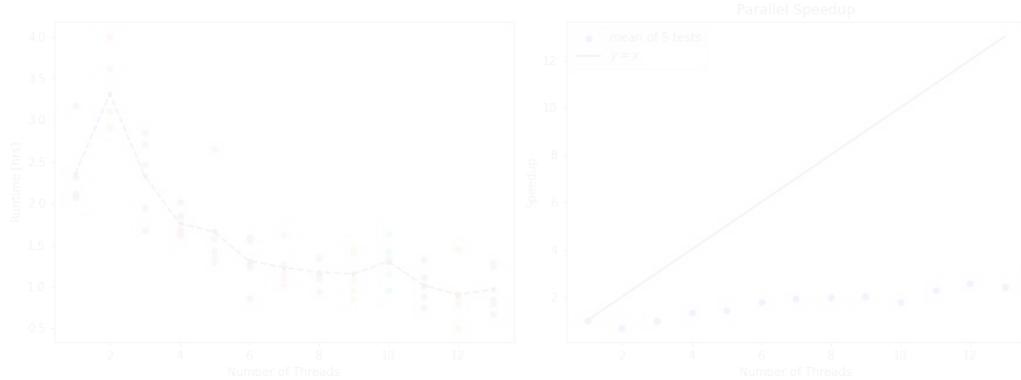
Algorithmic efficiency

How to optimally speed-up Nested Samplers with multi-threading?



of exoplanetary transits & RVs

Nested Samplers evaluate N_{live} times the likelihood



Who are we?

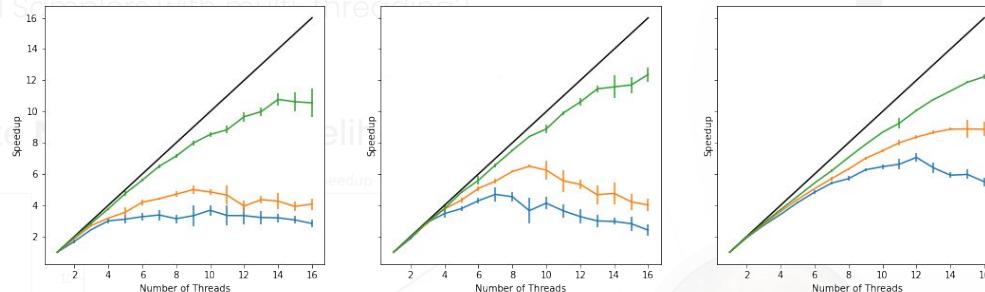
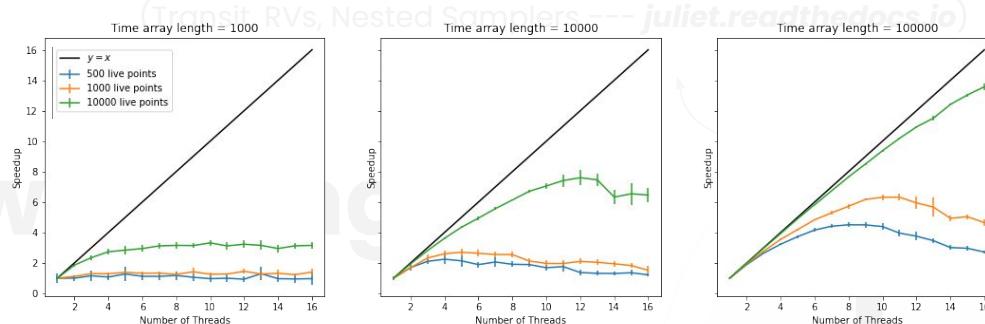
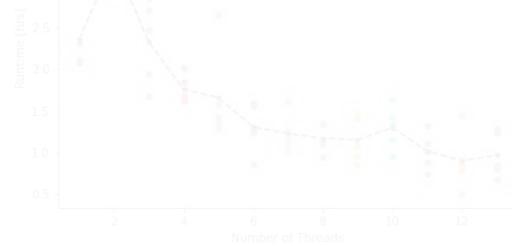
batman
(i.e., transit
model evaluation)

Algorithmic efficiency

How to optimally speed-up Nested Sampling in $\mathcal{O}(N \log N)$ threads

celerite

Nested sampling
(i.e., fast GP
likelihood evaluation)



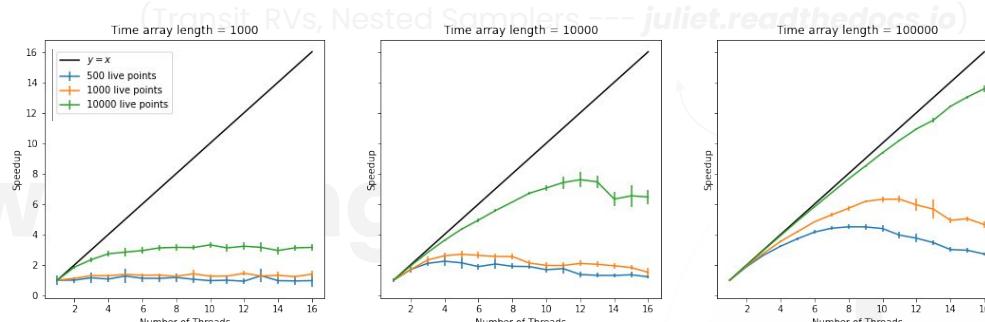
Who are we?

Algorithmic efficiency

How to optimally speed-up Nested Sampling with threads?

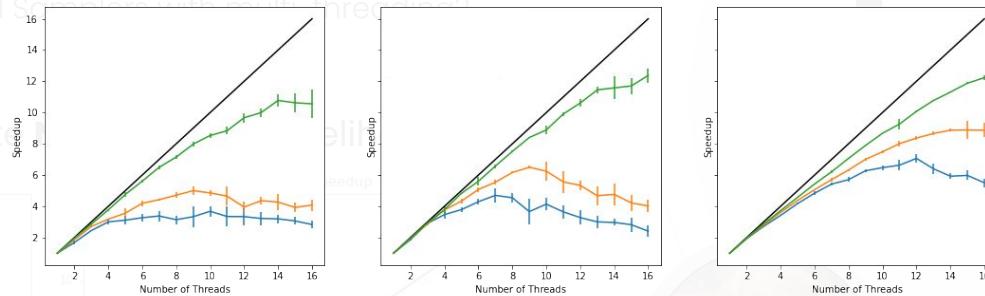
batman

(i.e., transit
model evaluation)



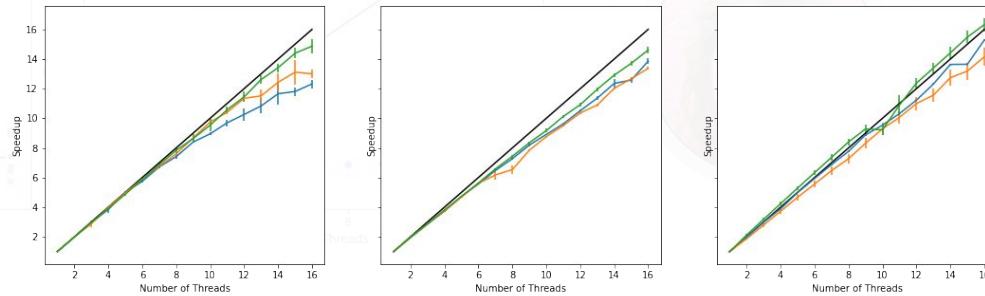
celerite

Nested Sampling
(i.e., fast GP
likelihood evaluation)



george

(i.e., slow GP
likelihood evaluation)



Memory intensive

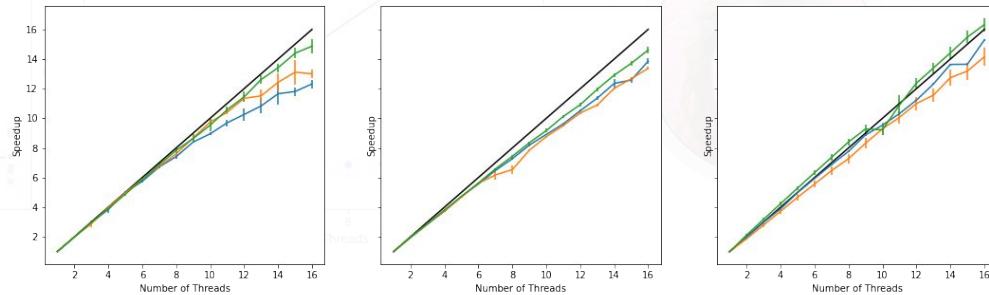
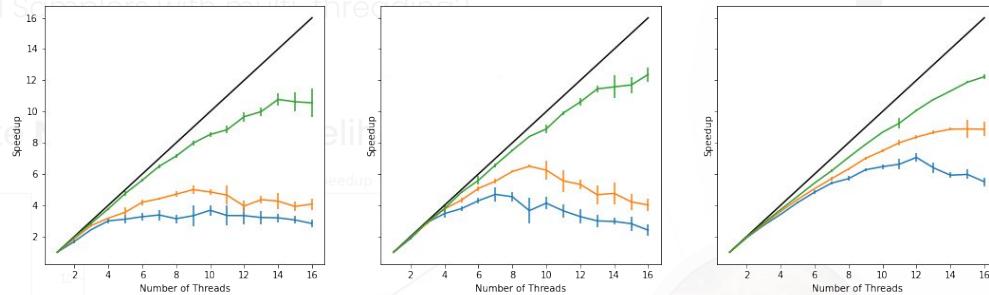
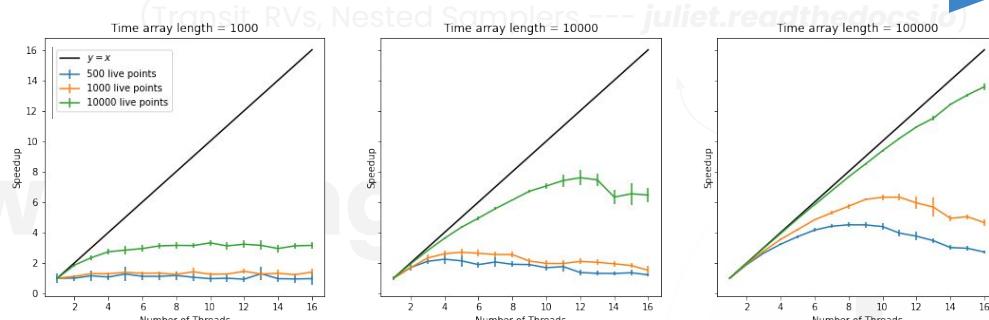
batman
(i.e., transit
model evaluation)

Algorithmic efficiency

How to optimally speed-up Nested Sampling with threads

celerite
(i.e., fast GP
likelihood evaluation)

george
(i.e., slow GP
likelihood evaluation)



Computationally intensive

(Transit, RVs, Nested Samplers --- juliet.readthedocs.io)

What are we doing?

How to use & deploy cloud services on the web?

(currently)



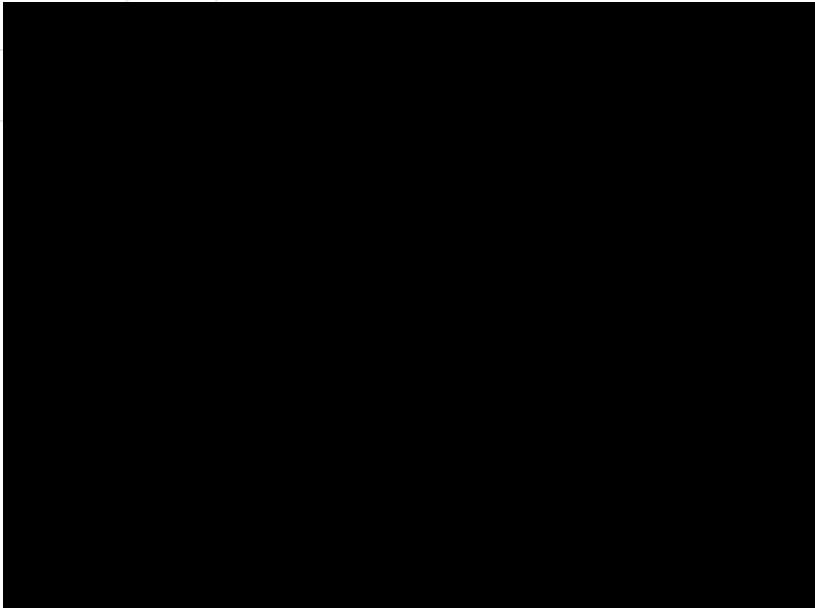
Joint analysis of exoplanetary transits & RVs
Juliet



(Transit, RVs, Nested Samplers --- juliet.readthedocs.io)

What are we doing?

How to use & deploy cloud services on the web?



(currently)



oint analysis of exoplanetary transits & RVs
Juliet



exoctk.stsci.edu

(Transit, RVs, Nested Samplers --- juliet.readthedocs.io)

What are we doing?

How to use & deploy cloud services on the web?

Objective: link **any** python code with the cloud!

(currently)



oint analysis of exoplanetary transits & RVs
Juliet



(Transit, RVs, Nested Samplers --- juliet.readthedocs.io)

What are we doing?

How to use & deploy cloud services on the web?

Objective: link **any** python code with the cloud!

The screenshot shows a GitHub repository page for 'exo_bespin/exo_bespin'. At the top, there's a pull request from 'bourque' titled 'Merge pull request #26 from exo-bespin/support-ubuntu'. Below the pull request, there's a list of commits showing changes to various files like '.github/workflows', 'exo_bespin', 'notebooks', '.gitignore', '.pep8speaks.yml', 'CODE_OF_CONDUCT.md', 'LICENSE', 'README.rst', 'environment.yml', and 'setup.py'. The commits are dated from November 9, 2020, to just over 3 months ago. At the bottom of the page, there's an 'Installation' section with instructions for users.

(currently)



oint analysis of exoplanetary transits & RVs
Juliet



(Transit, RVs, Nested Samplers --- juliet.readthedocs.io)

What are we doing?

How to use & deploy cloud services on the web?

Objective: link **any** python code with the cloud!

First milestone: deployed clickable button that launches a cloud computing instance!

(currently)



oint analysis of exoplanetary transits & RVs

Juliet



Summary

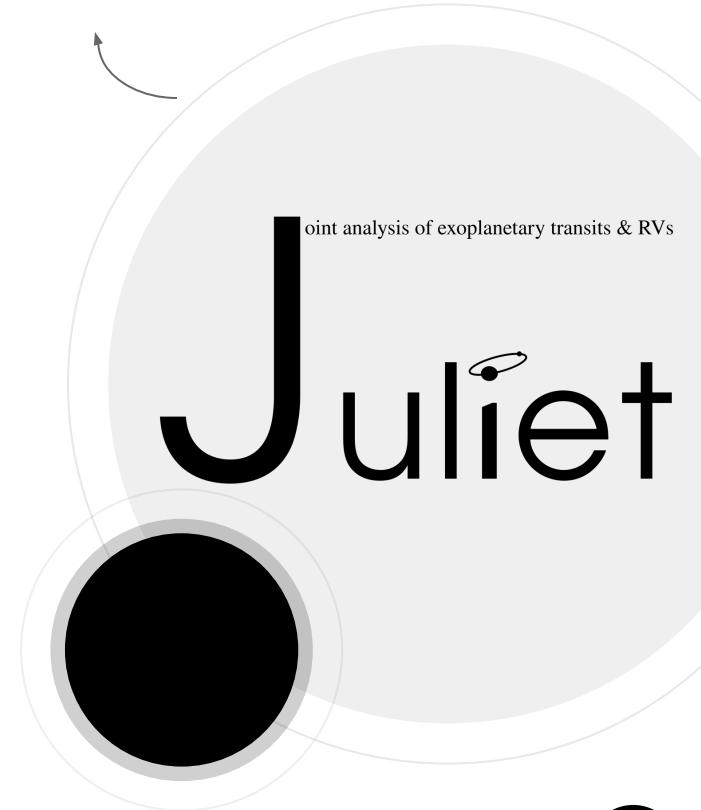
(Transit, RVs, Nested Samplers --- **juliet.readthedocs.io**)

Joint analysis of exoplanetary transits & RVs
Juliet

(Transit, RVs, Nested Samplers --- **juliet.readthedocs.io**)

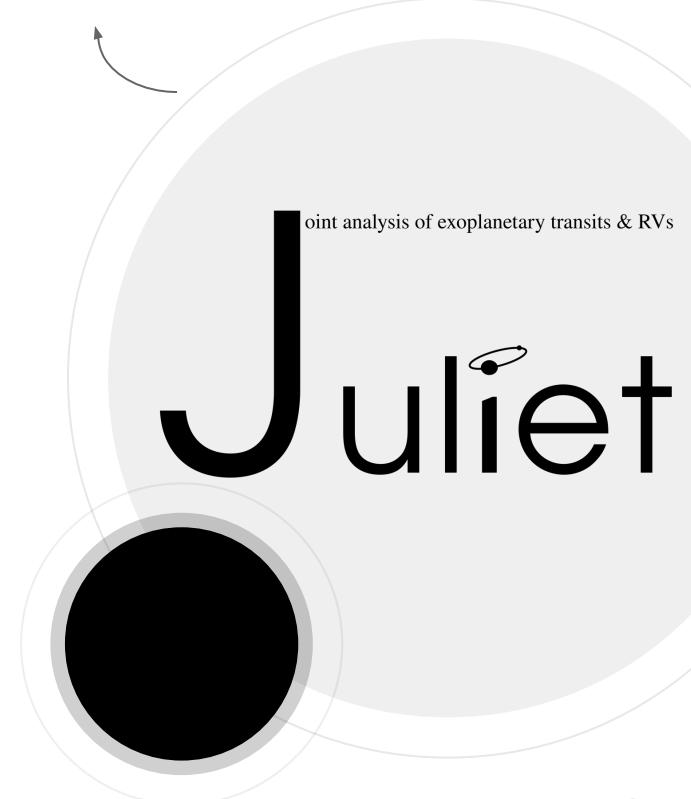
Summary

1. Current methodology to update exoplanet properties **is unsustainable.**
 - The Dream: an exoplanet census in the cloud!
 - To fully exploit cloud-computing benefits begs **optimizing algorithmic parameters (e.g., N_{live}).**
 - Minimal working example to **launch juliet runs in the cloud from the web** coming soon!



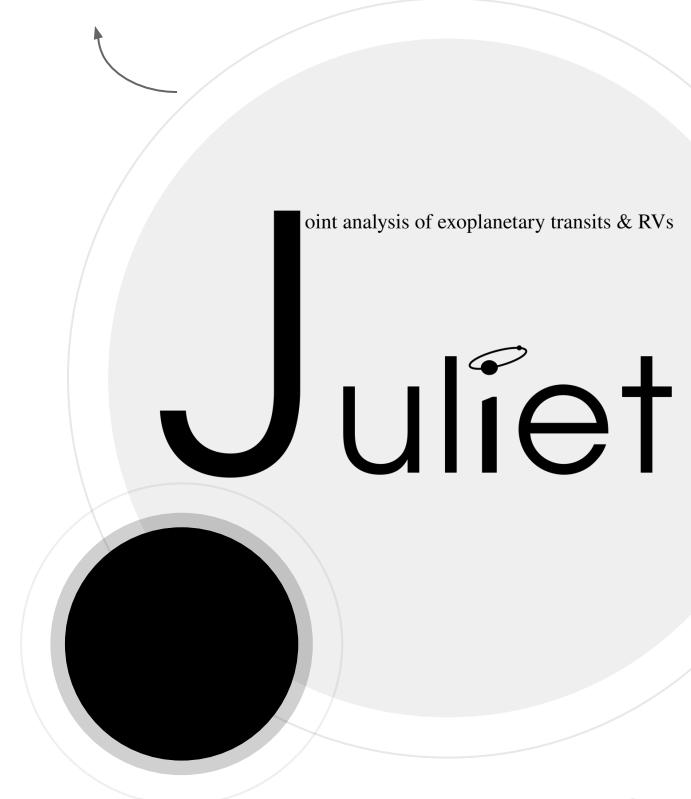
Summary

1. Current methodology to update exoplanet properties **is unsustainable**.
2. **The Dream:** an exoplanet census in the cloud!
 - To fully exploit cloud-computing benefits begs **optimizing algorithmic parameters (e.g., N_{live})**.
 - Minimal working example to **launch juliet runs in the cloud from the web** coming soon!



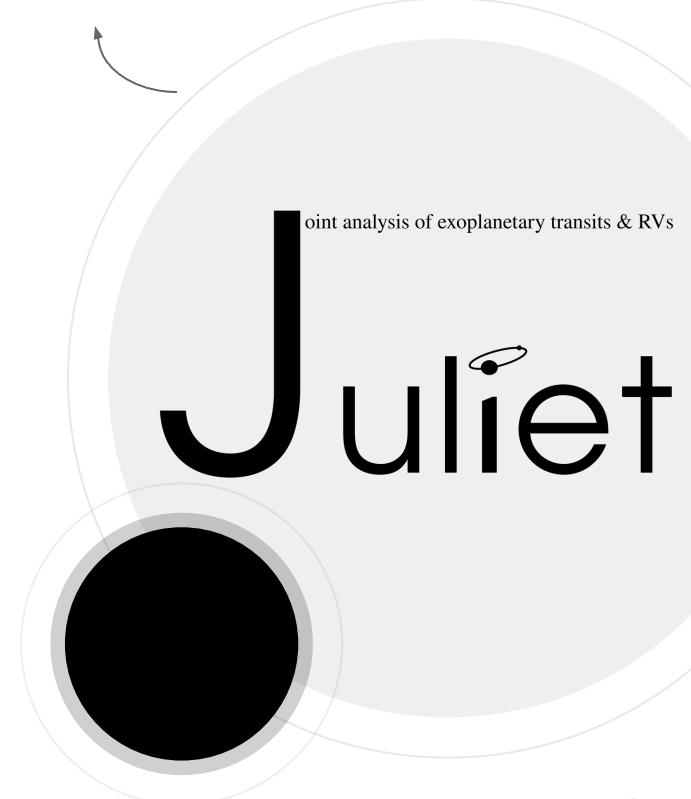
Summary

1. Current methodology to update exoplanet properties **is unsustainable**.
2. **The Dream:** an exoplanet census in the cloud!
3. To fully exploit cloud-computing benefits begs **optimizing algorithmic parameters (e.g., N_{live})**.
4. Minimal working example to **launch juliet runs in the cloud from the web** coming soon!



Summary

1. Current methodology to update exoplanet properties **is unsustainable**.
2. **The Dream:** an exoplanet census in the cloud!
3. To fully exploit cloud-computing benefits begs **optimizing algorithmic parameters (e.g., N_{live})**.
4. Minimal working example to **launch juliet runs in the cloud from the web** coming soon!



What are we doing?

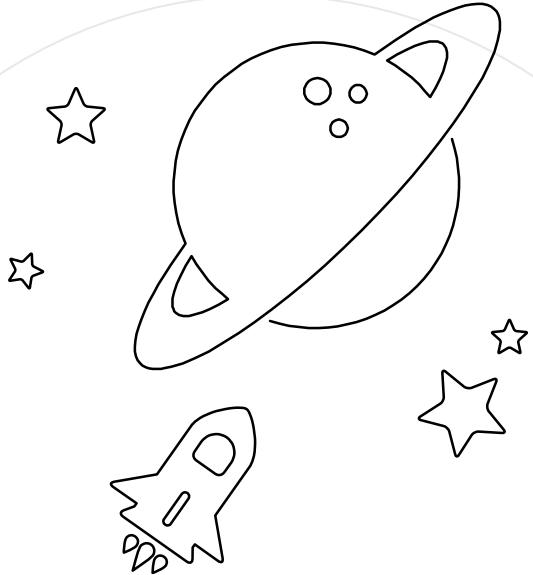
White

Is the color of milk and fresh snow, the color produced by the combination of all the colors of the visible spectrum.

Black

Is the color of ebony and of outer space. It has been the symbolic color of elegance, solemnity and authority.





Big concept

Bring the attention of your audience over a key concept using icons or illustrations

A picture is worth a thousand words

A complex idea can be conveyed with just a single still image, namely making it possible to absorb large amounts of data quickly.



And tables to compare data

	A	B	C
Yellow	10	20	7
Blue	30	15	10
Orange	5	24	16

Maps



89,526,12

4

Whoa! That's a big number, aren't you proud?



89,526,124\$

That's a lot of money

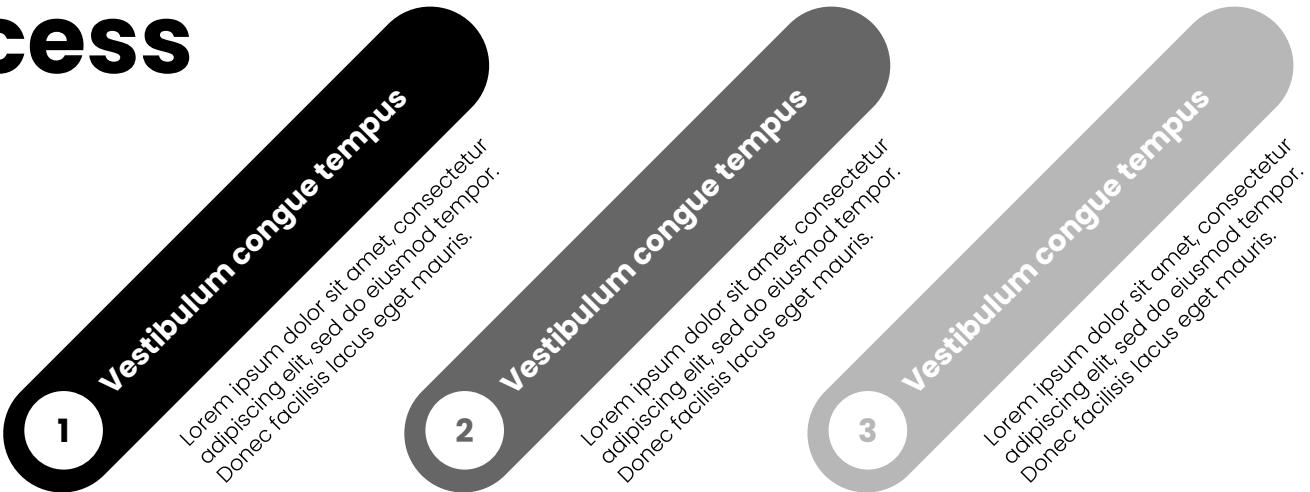
185,244 users

And a lot of users

100%

Total success!

Our process is easy



Let's review some concepts

Yellow

Is the color of gold, butter and ripe lemons.

Blue

Is the colour of the clear sky and the deep sea.

Red

Is the color of blood, danger and courage.

Yellow

Is the color of gold, butter and ripe lemons.

Blue

Is the colour of the clear sky and the deep sea.

Red

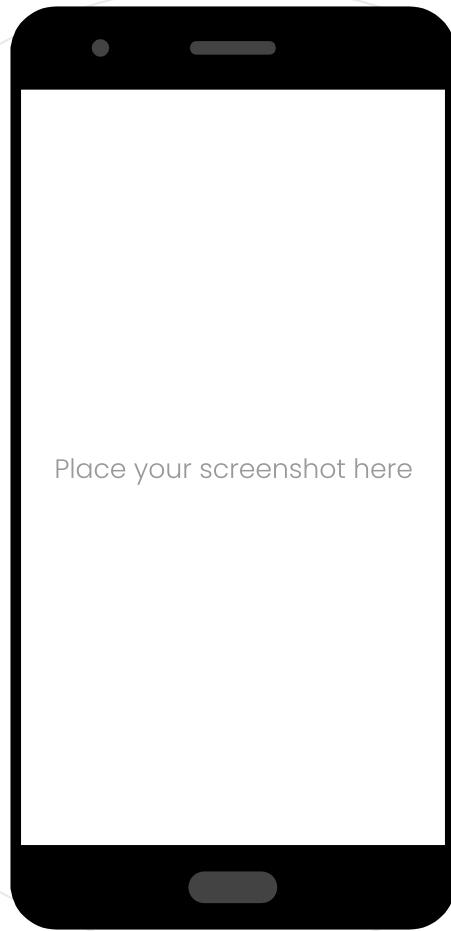
Is the color of blood, danger and courage.





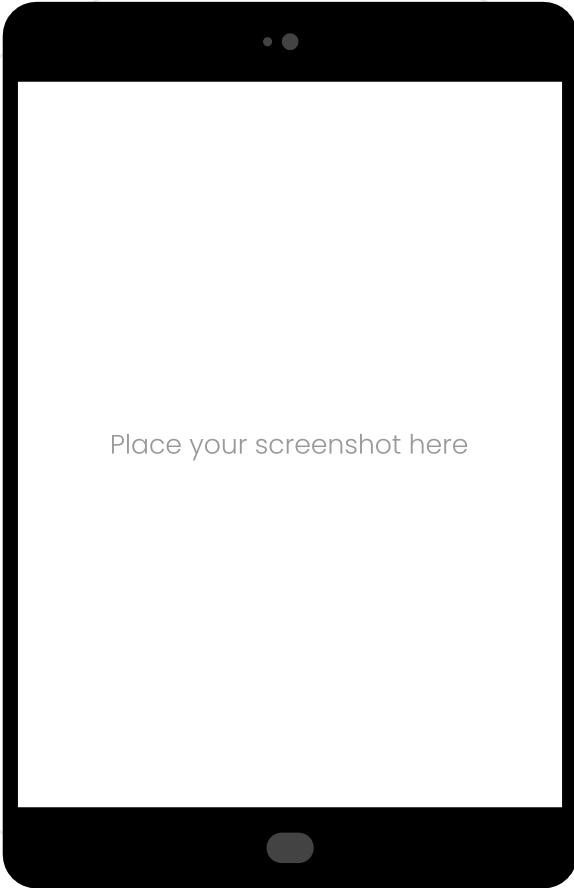
You can insert graphs from [Google Sheets](#)

Mobile project



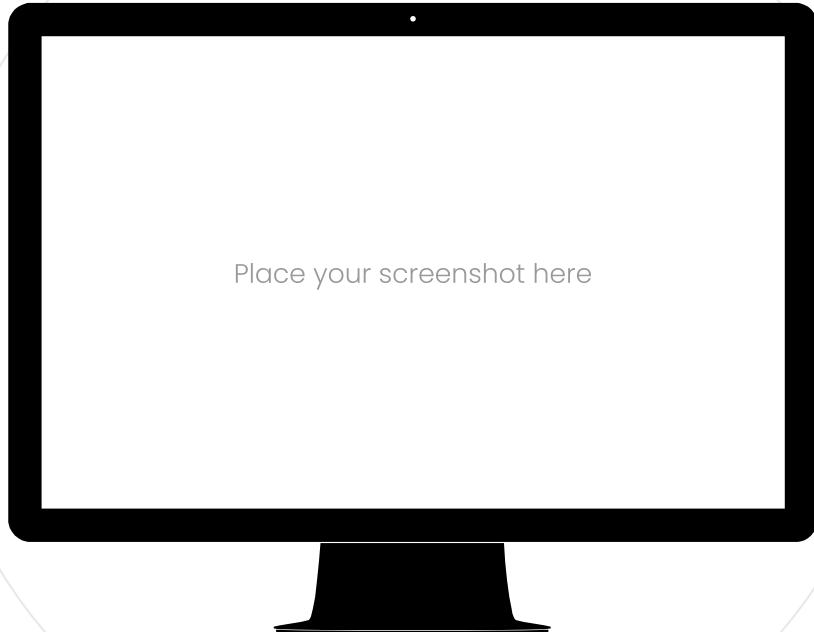
Show and explain your web, app or software projects using these gadget templates.

Tablet project



Show and explain your web, app or software projects using these gadget templates.

Desktop project



Place your screenshot here

Show and explain your web, app or software projects using these gadget templates.



Thanks!

Any questions?

You can find me at

@username

user@mail.me

Credits

Special thanks to all the people who made and released these awesome resources for free:

- Presentation template by [SlidesCarnival](#)
- Photographs by [Unsplash](#)

Presentation design

This presentation uses the following typographies:

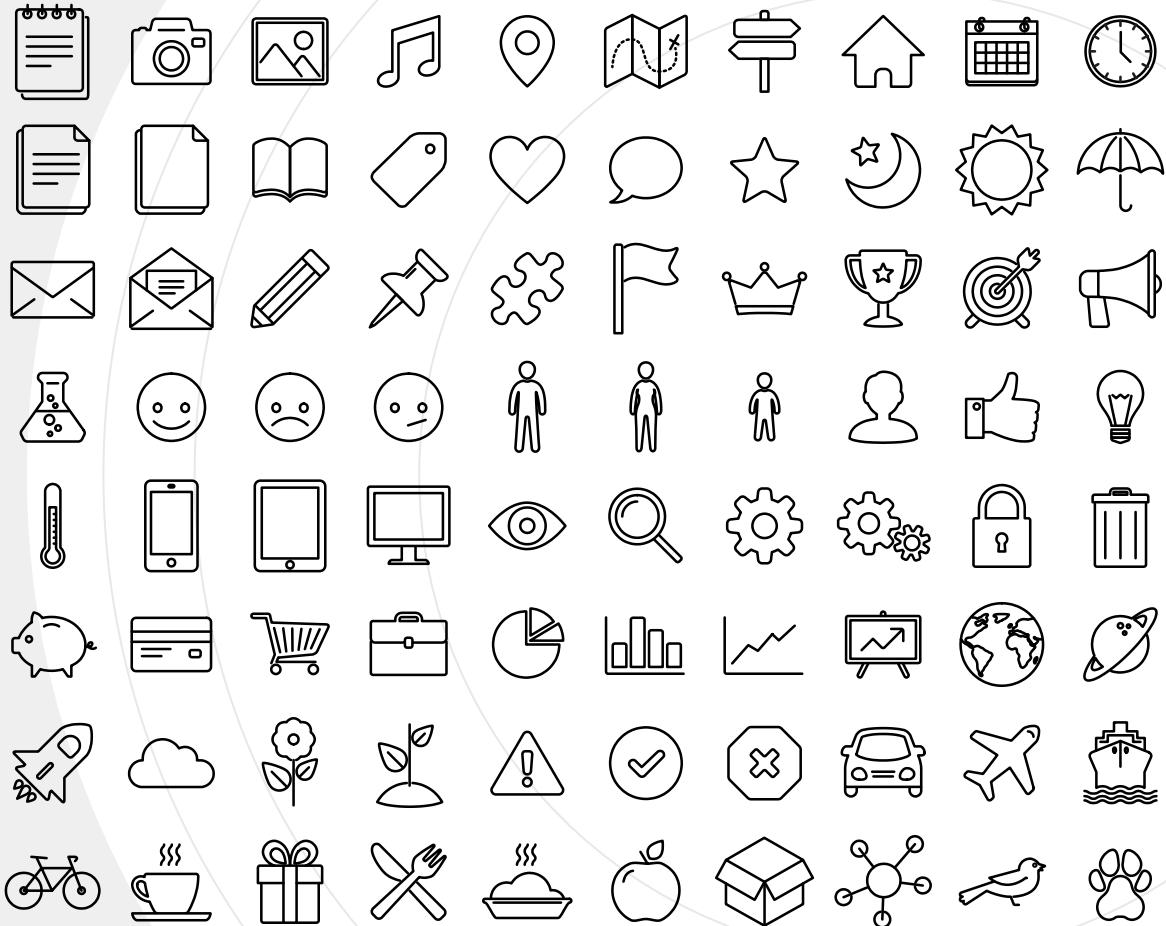
- Titles: Poppins Bold
- Body copy: Poppins Light

You can download the fonts at:

<https://www.fontsquirrel.com/fonts/poppins>

You don't need to keep this slide in your presentation. It's only here to serve you as a design guide if you need to create new slides or download the fonts to edit the presentation in PowerPoint®





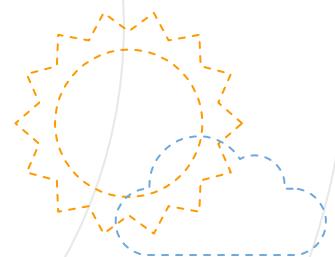
SlidesCarnival icons are editable shapes.

This means that you can:

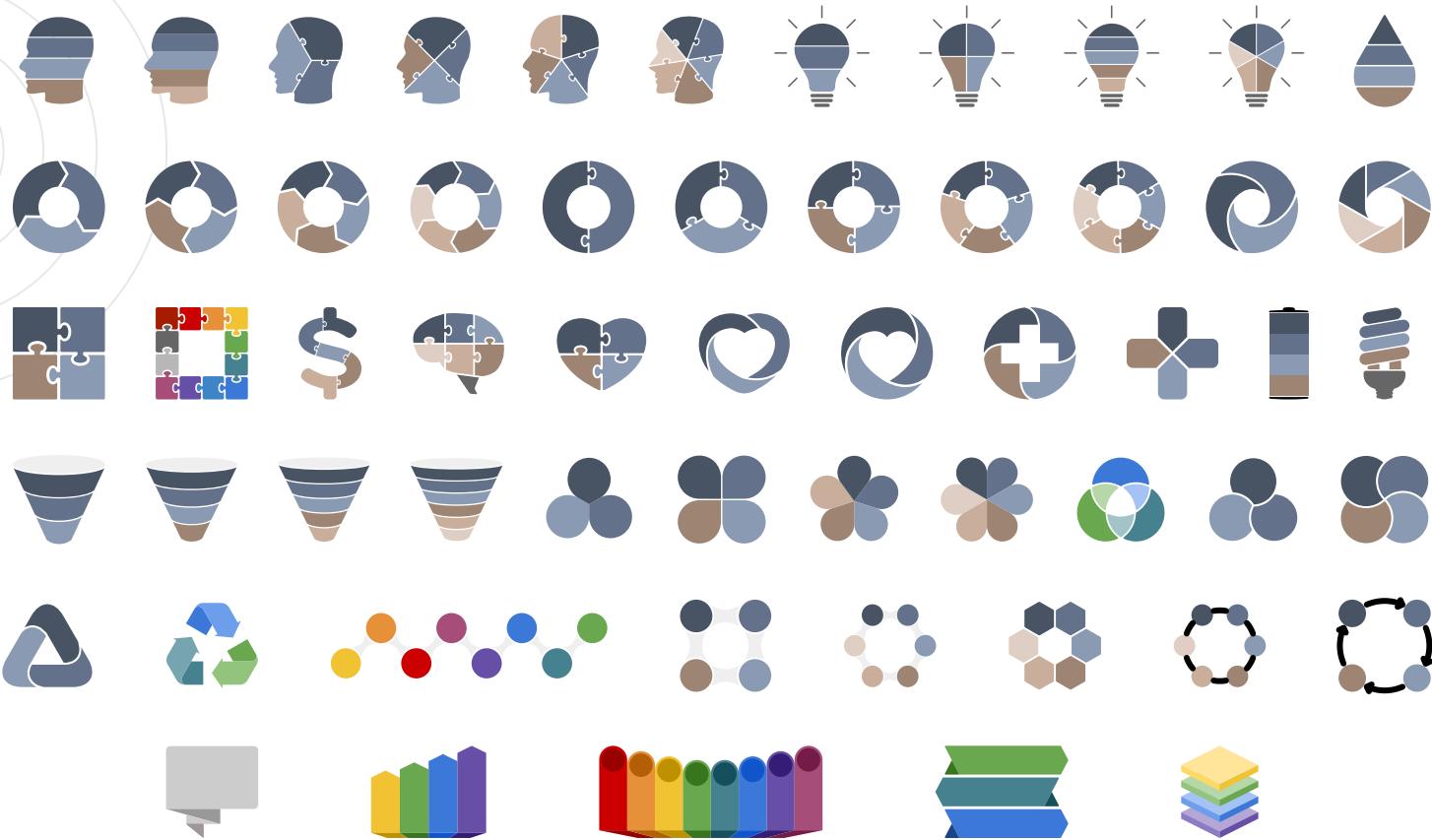
- Resize them without losing quality.
- Change line color, width and style.

Isn't that nice? :)

Examples:



Diagrams and infographics



Now you can use any emoji as an icon!
And of course it resizes without losing quality and you can change the color.

How? Follow Google instructions

<https://twitter.com/googledocs/status/730087240156643328>



and many more...



Free templates for all your presentation needs



For PowerPoint and
Google Slides



100% free for personal
or commercial use



Ready to use,
professional and
customizable



Blow your audience
away with attractive
visuals