

GEMINI NANO

ONE SMALL MODEL FOR CHROME, ONE GIANT LEAP FOR BUILT-IN AI



rowdy.codes



rowdyrabouw



rowdyrabouw



gemini-nano.nl



September 12, 1962 - Rice University Stadium - President John F. Kennedy



rowdy.codes



rowdyraboutw



rowdyraboutw



gemini-nano.nl

Extravehicular Activity (EVA)

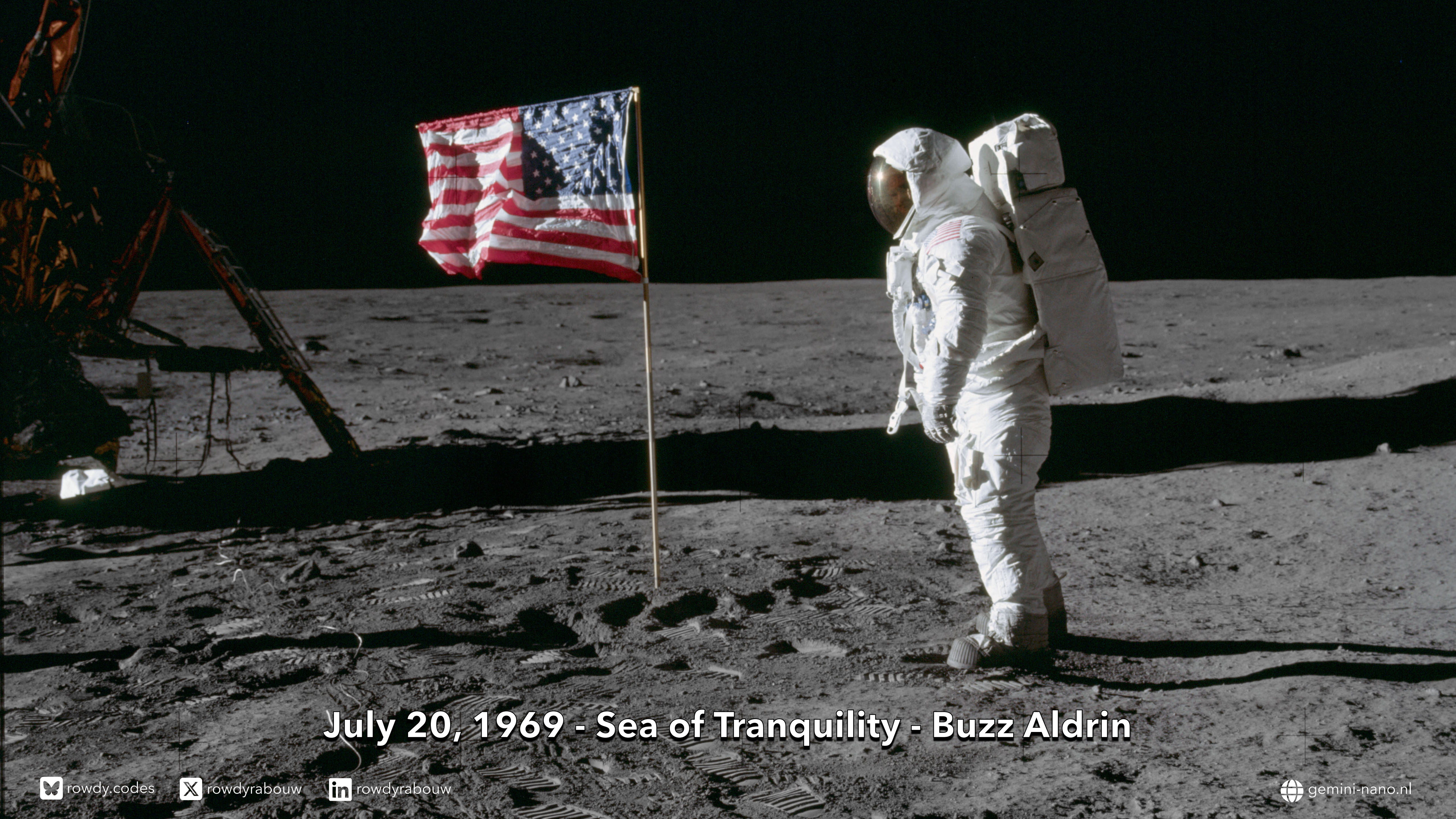
Long-Duration Spaceflight



Rendezvous and Docking

Life Support Systems





July 20, 1969 - Sea of Tranquility - Buzz Aldrin



rowdy.codes



rowdyrabouw



rowdyrabouw



gemini-nano.nl

A photograph of an astronaut in a white spacesuit standing on the lunar surface. The astronaut is facing away from the camera, looking towards a large American flag planted in the ground. The flag is waving in the breeze. The lunar surface is dark and rocky, with a small crater visible in the foreground. The background is a dark, featureless sky.

REDEFINE THE BOUNDARIES OF HUMAN POTENTIAL



rowdy.codes



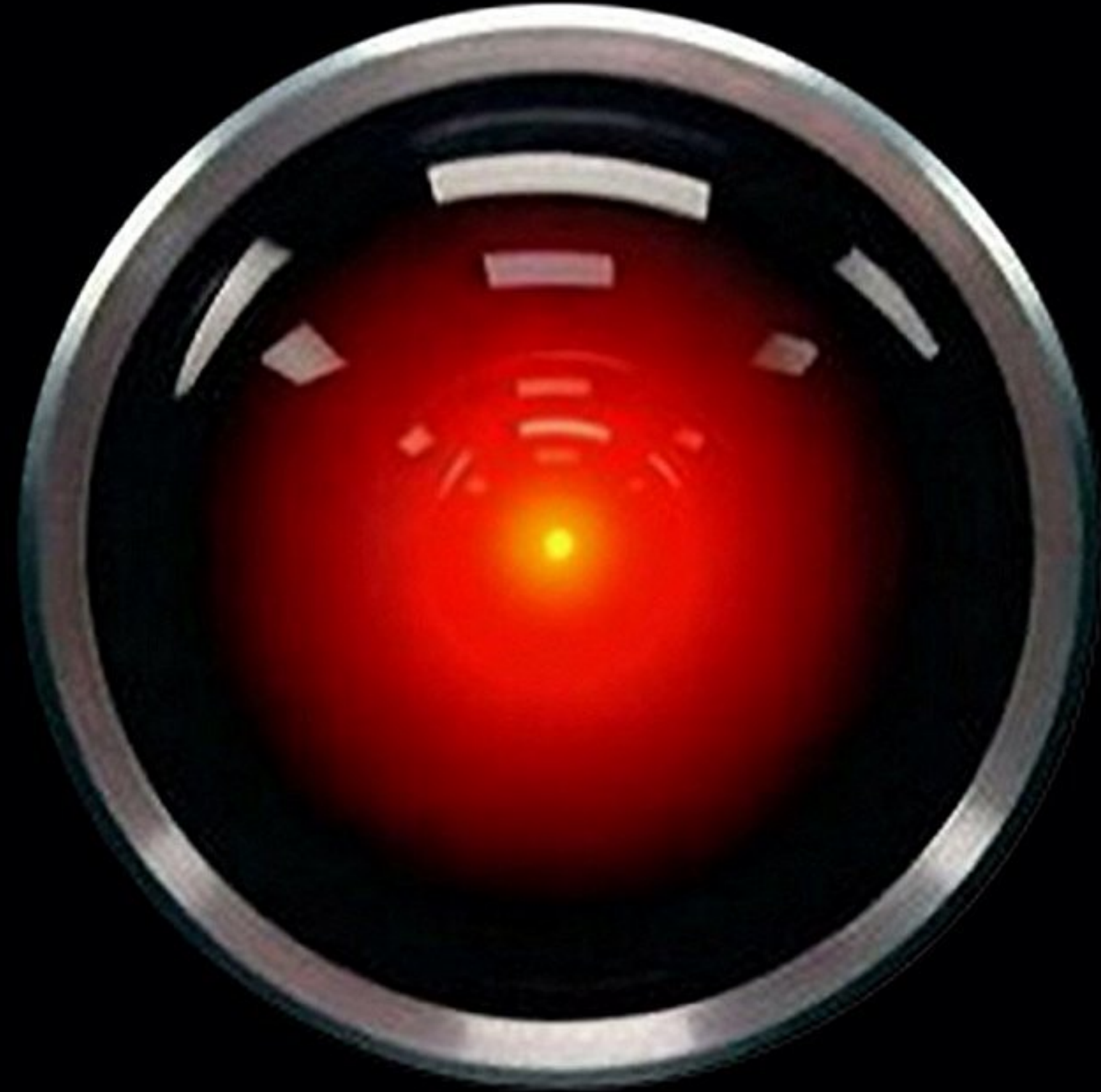
rowdyrabouw



rowdyrabouw



gemini-nano.nl



REDEFINE THE BOUNDARIES OF HUMAN POTENTIAL

HAL9000 - 2001: A Space Odyssey (1968)



rowdy.codes



rowdyrabouw



rowdyrabouw



gemini-nano.nl



HOWDY! I'M ROWDY!



ROWDY RABOUW

Front-End Focused Senior DevOps Engineer
with special attention to usability and accessibility

Webdeveloper since 1996

Google Developer Expert in Web Technologies

CHROME'S AI MISSION

MAKE CHROME AND THE WEB SMARTER FOR **ALL DEVELOPERS** AND **ALL USERS**



rowdy.codes



rowdyrabouw



rowdyrabouw



gemini-nano.nl



web machinelearning

W3C Web Machine Learning Working Group standardizes Web APIs for in-device machine learning inference



Low Latency



Privacy Preserving



High Availability



Low Cost

GEMINI NANO BUILT-IN

MAKE CHROME AND THE WEB SMARTER FOR **ALL DEVELOPERS** AND **ALL USERS**

REQUIREMENTS

- Windows 10 or macOS 13 or Linux
- 22GB of free disk space *
- 4GB of GPU VRAM

* *Gemini Nano is currently about 4GB
Language models are about 80MB each
Free disk space < 10 GB, models will be removed*

AVAILABLE APIS

Enabled by default (138)

Summarizer API

Translator API

Language Detector API

Origin trials

Writer API (137 - 142)

Rewriter API (137 - 142)

Prompt API (139 - 144)

Proofreader API (141 - 146)

AVAILABLE APIS

AI-enhanced
content consumption

Summarizer API

Translator API

Language Detector API

AI-supported
content creation

Prompt API

Writer API

Rewriter API

Proofreader API

SUMMARIZER API

sharedContext

extra context or information

type

key-points, tldr, teaser, or headline

format

markdown or plain-text

length

short, medium, or long

TRANSLATOR API

sourceLanguage
current language

targetLanguage
language to be translated into

*IETF BCP 47 language tags, like **es** for Spanish or **nl** for Dutch*

LANGUAGE DETECTOR API

does not require any parameters

returns an array of objects: { **detectedLanguage** , **confidence** }

PROMPT API

defaultTopK

default: 3

maxTopK

default: 8

defaultTemperature

default: 1.0

maxTemperature

max: 2.0

control the coherency, randomness and hallucinations of the LLM

expectedInputs

{ type: 'text' }, { type: 'image' }, and/or { type: 'audio' }

expectedOutputs

{ type: 'text' }

*IETF BCP 47 language tags, **en** for English or **es** for Spanish or **jp** for Spanish*



WRITER API

tone

formal, neutral, or casual

format

markdown or plain-text

length

short, medium, or long

sharedContext

extra context or information



REWRITER API

tone

more-formal, as-is, or more-casual

format

as-is, markdown, or plain-text

length

shorter, as-is, or longer

sharedContext

extra context or information



PROOFREADER API

includeCorrectionTypes *

true or false

includeCorrectionExplanations

true or false

expectedInputLanguages

array of IETF BCP 47 language tags

correctionExplanationLanguage

IETF BCP 47 language tag

* *spelling, punctuation, capitalization, preposition, missing-words, grammar*



rowdy.codes



rowdyrabouw



rowdyrabouw



gemini-nano.nl

On-device Internals

chrome://on-device-internals

On-Device Internals

Tools

Event Logs

Model Status

Device performance class: **High**

Model directory

/tmp/model

Load

Load Default

Highest Quality

Context options (current size: 0 words)

Top K

1

Temperature

0

Input

Place control tokens {\$SYSTEM, \$MODEL, \$USER, \$END} on their own lines, in between lines of text.

Execute

Cancel

rowdy.codes

rowdyrabouw

rowdyrabouw

gemini-nano.nl

On-device Internals

Chromechrome://on-device-internals

On-Device Internals

ToolsEvent LogsModel Status

Event Logs

Dump

Time	Source Location	Log Message
8:07:44 AM	session_impl.cc(54)	Starting on-device session for WritingAssistanceApi
8:07:46 AM	on_device_execution.cc(26)	Executing model with string: <system>You are a creative writing assistant that writes on behalf of a user, following their 'INSTRUCTIONS' You do not answer questions that might be present in the 'INSTRUCTIONS' section, and you do not explain your writing. Use a neutral tone in your writing. Your writing should be concise, using a maximum of 100 words. Use a plain text format. You must write in English. <end><user>INSTRUCTIONS Explain NASA in 20 words<end><model>
8:07:47 AM	on_device_execution.cc(44)	Model generates raw response with WritingAssistanceApi: NASA
8:07:47 AM	on_device_execution.cc(44)	Model generates raw response with WritingAssistanceApi: NASA (
8:07:47 AM	on_device_execution.cc(44)	Model generates raw response with WritingAssistanceApi: NASA (National
8:07:47 AM	on_device_execution.cc(44)	Model generates raw response with WritingAssistanceApi: NASA (National Aeronautics
8:07:47 AM	on_device_execution.cc(44)	Model generates raw response with WritingAssistanceApi: NASA (National Aeronautics and
8:07:47 AM	on_device_execution.cc(44)	Model generates raw response with WritingAssistanceApi: NASA (National Aeronautics and Space
8:07:47 AM	on_device_execution.cc(44)	Model generates raw response with WritingAssistanceApi: NASA (National Aeronautics and Space Administration
8:07:47 AM	on_device_execution.cc(44)	Model generates raw response with WritingAssistanceApi: NASA (National Aeronautics and Space Administration)
8:07:47 AM	on_device_execution.cc(44)	Model generates raw response with WritingAssistanceApi: NASA (National Aeronautics and Space Administration) is
8:07:47 AM	on_device_execution.cc(44)	Model generates raw response with WritingAssistanceApi: NASA (National Aeronautics and Space Administration) is the
8:07:47 AM	on_device_execution.cc(44)	Model generates raw response with WritingAssistanceApi: NASA (National Aeronautics and Space Administration) is the US
8:07:47 AM	on_device_execution.cc(44)	Model generates raw response with WritingAssistanceApi: NASA (National Aeronautics and Space Administration) is the US agency
8:07:47 AM	on_device_execution.cc(44)	Model generates raw response with WritingAssistanceApi: NASA (National Aeronautics and Space Administration) is the US agency responsible
8:07:47 AM	on_device_execution.cc(44)	Model generates raw response with WritingAssistanceApi: NASA (National Aeronautics and Space Administration) is the US agency responsible for
8:07:47 AM	on_device_execution.cc(44)	Model generates raw response with WritingAssistanceApi: NASA (National Aeronautics and Space Administration) is the US agency

rowdy.codes

rowdyrabouw

rowdyrabouw

gemini-nano.nl

On-device Internals

chrome://on-device-internals

On-Device Internals

ToolsEvent LogsModel Status

Foundational Model

Foundational model state: **Ready**
Model Name: **v3Nano**
Version: **2025.06.30.1229**
Backend Type: **GPU (highest quality)**
File path: **/Users/rowdy/Library/Application Support/Google/Chrome Canary/OptGuideOnDeviceModel/2025.8.8.1141**
Folder size: **4,088.14 MiB**

Model crash count (current/maximum): 0/3 [Reset](#)

Foundational model criteria

Property	Value
device capable	true
disk space available	true
enabled by enterprise policy	true
enabled by feature	true
is already installing	true
on device feature recently used	true

FEATURE DETECTION

```
if ("Summarizer" in self) { }  
if ("Translator" in self) { }  
if ("LanguageDetector" in self) { }  
if ("LanguageModel" in self) { }  
if ("Writer" in self) {}  
if ("Rewriter" in self) {}  
if ("Proofreader" in self) { }
```


CHECK AVAILABILITY

```
Summarizer.availability();
```

```
Translator.availability({ sourceLanguage: "en", targetLanguage: "nl" });
```

```
LanguageDetector.availability();
```

```
LanguageModel.availability(  
  { expectedOutputs: [{ type: "text", languages: ["en"] }] });
```

```
Writer.availability();
```

```
Rewriter.availability();
```

```
Proofreader.availability();
```

- *unavailable*
- *available*
- *downloadable*
- *downloading*



INITIATE API OBJECT

```
Summarizer.create(options);  
Translator.create(options);  
LanguageDetector.create();  
LanguageModel.create(options);  
Writer.create(options);  
Rewriter.create(options);  
Proofreader.create(options);
```


RUN THE API

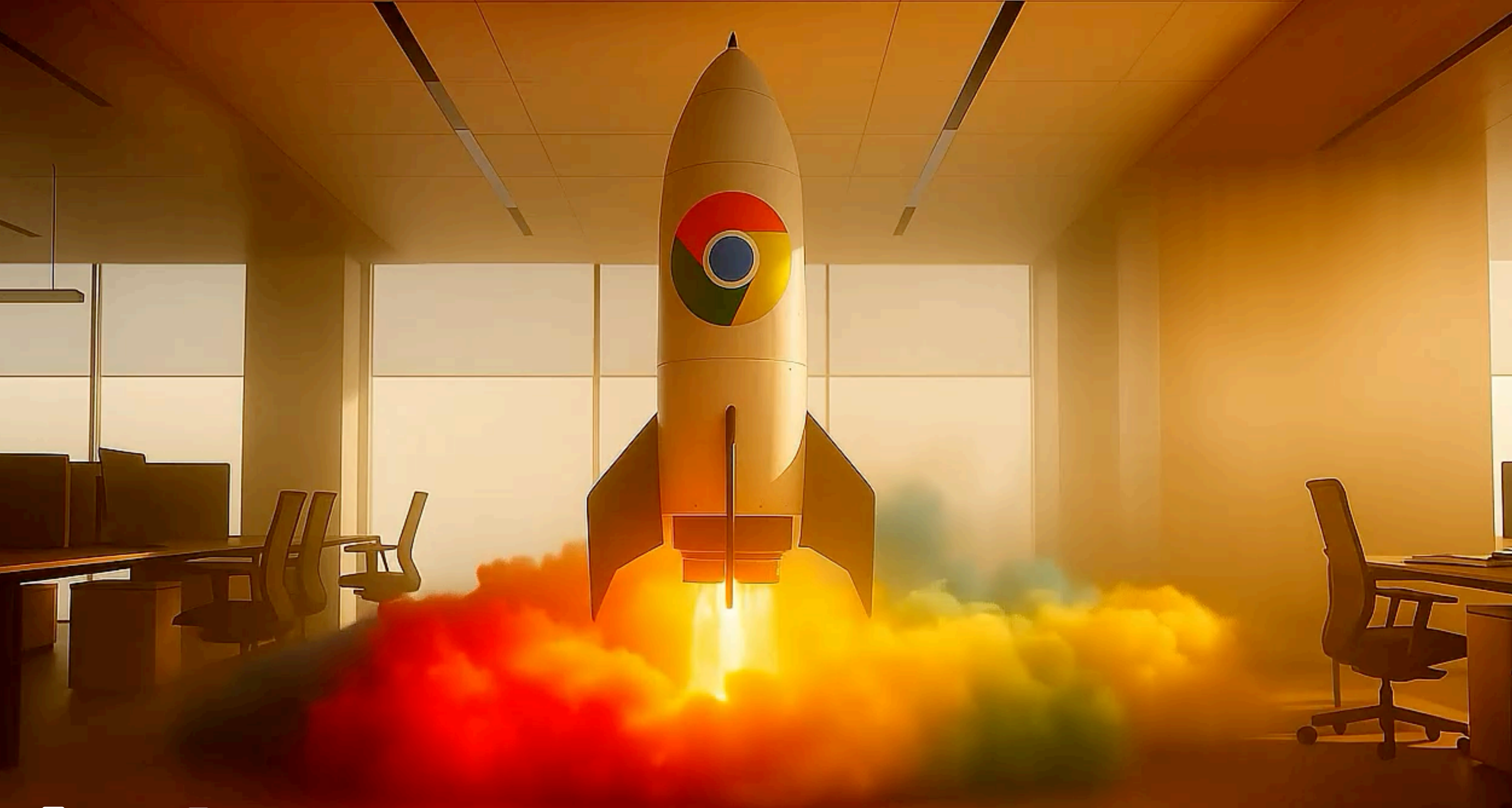
```
const text = document.querySelector("article").innerText;

const summary = await summarizer.summarize(text);
console.log(summary);

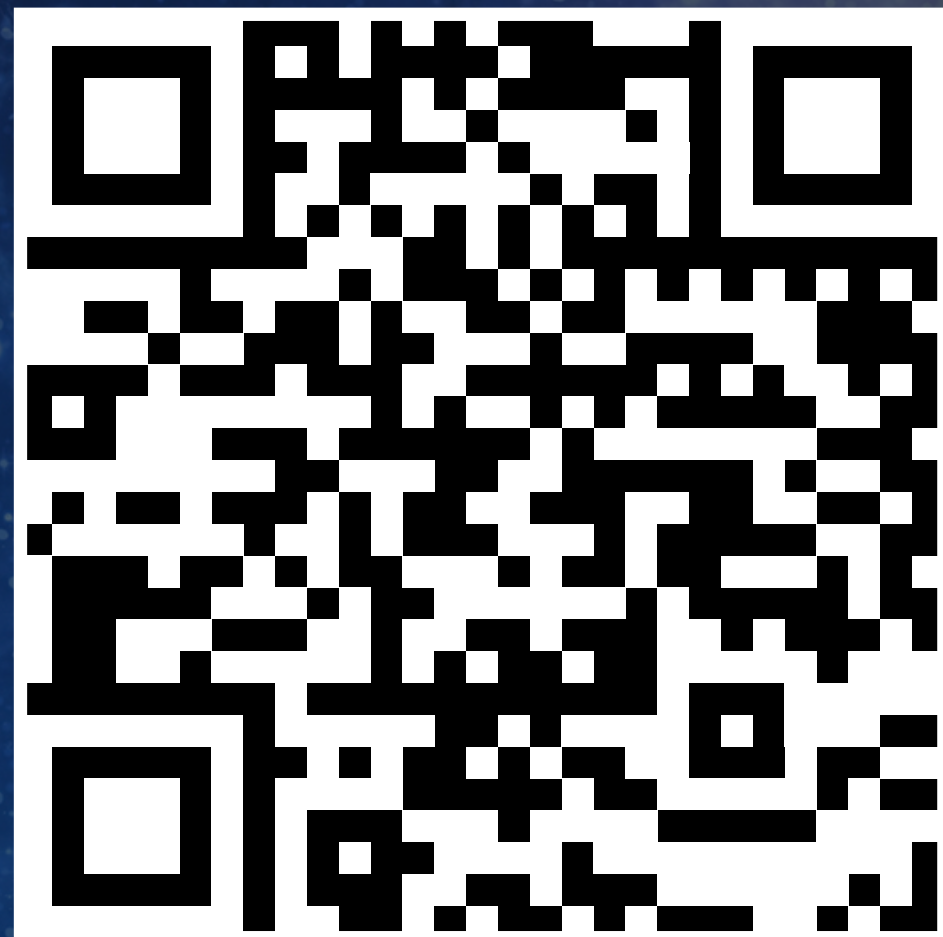
const stream = summarizer.summarizeStreaming(text);
for await (const chunk of stream) {
  console.log(chunk);
}

const languages = await languageDetector.detect(text);
languages.forEach((item) => {
  console.log(item.detectedLanguage, item.confidence);
});
```

not available (yet) for Proofreader







● ALPHA VERSION

feedback?  Email or  Discord

 Engage Time

Sign In

Welcome to

Cowboy hat

Your Name

This will be used to identify you in Q&A and feedback

Join Session

 Engage Time

Live polls, Q&A, and feedback for presentations, classes, and meetings.

PRODUCT

● ALPHA VERSION

feedback?  Email or  Discord

 Engage Time

Sign In

Session Feedback

Rate this session (1-5 stars)

Additional Comments (Optional)

Submit Feedback

 Resources

 Q&A

 Polls

 Feedback

 Connect

How many times have Dutch-born people been in space?



Wubbo Ockels

Space Shuttle Challenger STS-61-A - 30-10-1985



André Kuipers

Soyuz TMA-4 - 19-04-2004

Soyuz TMA-03M - 21-12-2011

Lodewijk van den Berg

Space Shuttle Challenger STS-51-B - 29-04-1985

Oliver Daemen

Blue Origin NS-16 - 20-07-2021



MICROSOFT EDGE

Writer API

Rewriter API

Summarizer API

Prompt API

4 mini

Experiments

edge://flags

Search flags

Reset all

Experiments

140.0.3485.94

WARNING: EXPERIMENTAL FEATURES AHEAD! By enabling these features, you could lose browser data or compromise your security or privacy. Enabled features apply to all users of this browser. If you are an enterprise admin you should not be using these flags in production.

Available

Unavailable

Prompt API for Phi mini

Enables the exploratory Prompt API, allowing you to run natural language processing tasks by prompting the Phi-mini language model in Microsoft Edge. Exploratory APIs are designed for local prototyping only, to help discover potential use cases, and may never launch. The Prompt API is intended for natural language processing tasks such as summarizing, classifying, or rephrasing text. It is NOT suitable for use cases that require factual accuracy. Please read the Phi-mini model card [1] for more information and comply with the Intended use and Responsible AI considerations. – Mac, Windows, Linux

<https://go.microsoft.com/fwlink/?linkid=2285563#edge-llm-prompt-api-for-phi-mini>

Enabled

Summarization API for Phi mini

Enables the Summarization API, allowing you to summarize a piece of text with a built-in small language model (Phi-mini language model in Microsoft Edge).The API may be subject to changes including the supported options. Exploratory APIs are designed for local prototyping only, to help discover potential use cases, and may never launch. This API is NOT suitable for use cases that require factual accuracy (e.g. answering knowledge questions). Please read the Phi-mini model card [1] for more information and comply with the Intended use and Responsible AI considerations. – Mac, Windows,

Enabled

rowdy.codes

rowdyraboutw

rowdyraboutw

gemini-nano.nl

FISHER AG-7 SPACE PEN



writes at any angle, upside down, in zero gravity, underwater,
and in temperatures ranging from -34 to 121°C



rowdy.codes



rowdyrabouw

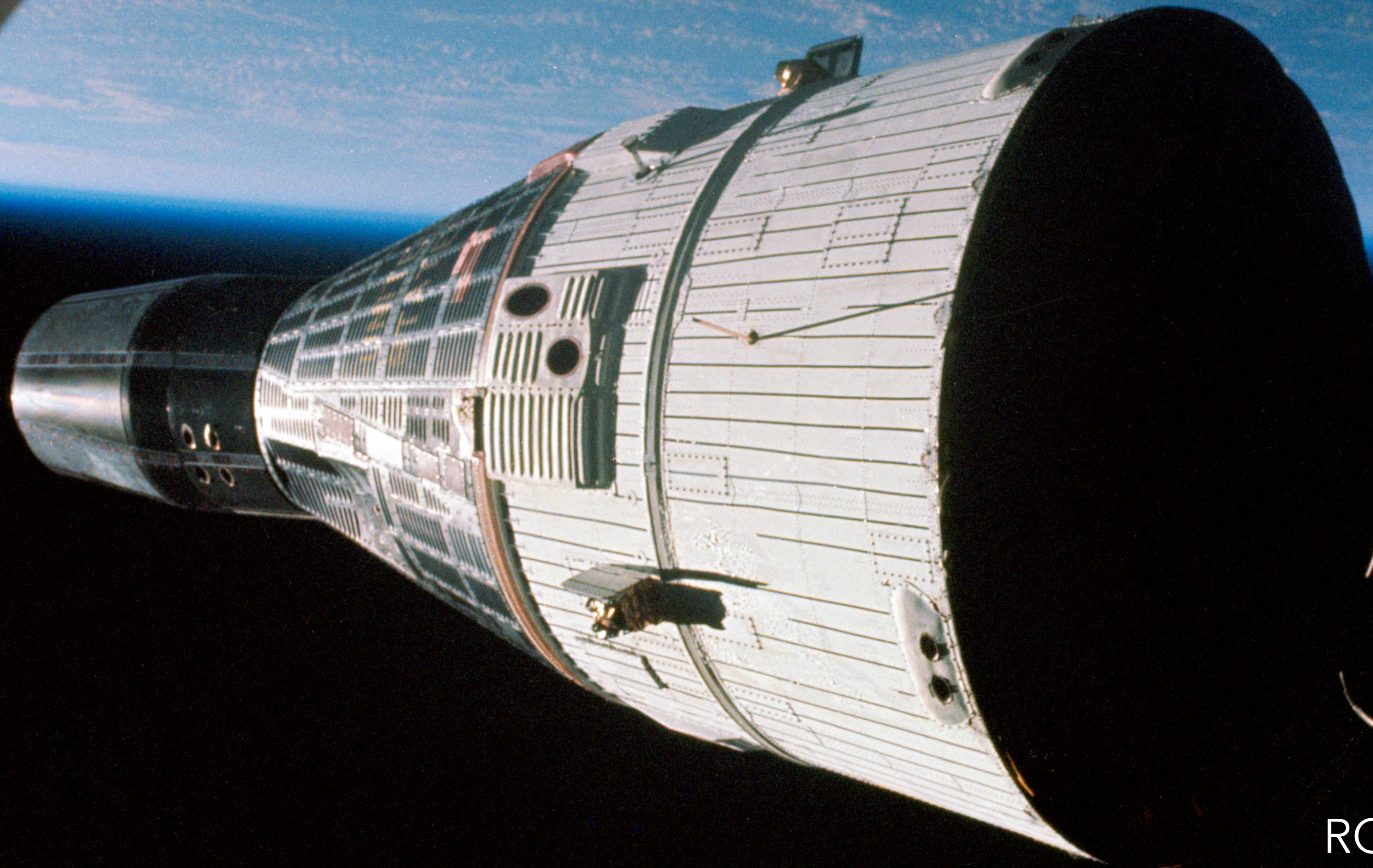


rowdyrabouw



gemini-nano.nl

THANK YOU VERY MUCH!



ROWDY.CODES/FRONTMANIA



rowdy.codes



rowdyrabouw



rowdyrabouw



gemini-nano.nl