IAC-24-E1.IP.36.x81945

Advancing worldwide interest in spaceflight through a nonprofit open data initiative

A. Muller[a]*, J. Rascagneres[b], C. Jones[c], D. Weijers[d], I. Li[e], L. Stauder[f], R. Bouchet[g]

[a] *The Space Devs, Toulouse, France*, arnaud.muller@thespacedevs.com
[b] *The Space Devs, Hartford, United Kingdom*, jacques.rascagneres@thespacedevs.com
[c] *The Space Devs, Sugar Hill, GA, United States*, caleb.jones@thespacedevs.com
[d] *The Space Devs, Hooglanderveen, The Netherlands*, derk.weijers@thespacedevs.com
[e] *The Space Devs, Hong Kong, China*, ivan.li@thespacedevs.com
[f] *The Space Devs, Darmstadt, Germany*, lucas.stauder@thespacedevs.com
[g] *The Space Devs, Levallois-Perret, France*, robin.bouchet@thespacedevs.com
* Corresponding Author

Abstract

The free and open accessibility of spaceflight information serves as a cornerstone for involving the public in space activities. Organized by a non-profit called *The Space Devs*, this endeavour is dedicated to the collection of public spaceflight information in a database updated in real-time by a team of volunteers located around the world. It is then made available to the public through an open standardized data retrieval portal. The information includes, but is not restricted to, all orbital launches since Sputnik 1, space agencies and companies, launch vehicles and pads, astronauts, spacecraft, space stations, as well as spaceflight events and news. This project distinguishes itself by its expansive user base, ranging from hobbyist app and website developers to content creators, teachers, academic groups, museums, analytics firms, and space agencies. The available data and features have progressed over multiple years in response to user feedback and industry trends, while their increasing popularity can be measured through the number of projects leveraging this service, their combined public reach, and their standing in the industry. Current assessments find that this initiative has garnered a global reach well into the multiple millions worldwide. This work elucidates key learnings for stakeholders in the field, underscoring the effectiveness of combining crowd-sourced data with precedence-based decision-making. It also highlights the importance of advocating for conservative consumption of public resources and finds that incorporating integrations with established entities from inception can significantly enhance interest in new offerings. Additionally, it emphasizes the benefits of instant and informal communication channels between users and team members for faster implementation of corrections and updates. Finally, it illustrates how finely tuned automations can help handle an increasing workload without compromising on quality of service. Overall, it showcases how a project such as *The Space Devs* stemming from a team of passionate volunteers collaborating across borders can meaningfully impact the landscape of space outreach and education.
Keywords: nonprofit, open data, spaceflight, API, OSINT

Acronyms/Abbreviations

The Space Devs (TSD), Launch Library 2 (LL2), Spaceflight News API (SNAPI)

1. Introduction

The Space Devs is a fully online nonprofit association of volunteers from around the world created in 2020. It strives to advance worldwide interest in spaceflight by improving accessibility of information.

To achieve this goal, it maintains two datasets related to space activities and makes them easily accessible to developers of public-reaching services through standardized data retrieval portals called APIs.

The first and largest API, called Launch Library 2, is since its release in 2020 the official successor of the original Launch Library API that became public in 2014. It contains extensive and up-to-date information on all space launches and many other kinds of spaceflight activities since Sputnik 1 in 1957.

The Spaceflight News API was created in 2018 and joined TSD at its inception to be maintained alongside LL2. It automatically references space-related news articles in English language from curated sites, as well as blog posts and mission reports from official sources.

These datasets are comparable with commercially available ones designed for market research and analytics but differ by their free access through public APIs. They also serve a different purpose than comparable free options, which usually are web pages designed for human consumption (e.g. wikis, blogs). Indeed, this project does not target the public directly but focuses on providing extensive and reliable data to developers and analysts, hobbyists and professionals alike. These users then leverage this data in public-facing content and services, such as websites, apps, infographics, museum exhibitions, research, community forums, podcasts, educational material, etc.

As of September 2024, the TSD team comprises 13 members from 10 countries across 4 continents and 7 time zones. It handles all the software development work for both APIs, as well as the data maintenance work, also called librarian work. Its members, all passionate about spaceflight, come from various backgrounds often but not always related to software and aerospace engineering.

2. Objectives

To achieve its mission of advancing worldwide interest in spaceflight, this project must pursue and balance two distinct objectives. The first is to provide open access to a large breadth of data for free in a format that makes it ideal for consumption by large-scale public-facing services. The second one is to ensure its own long-term financial sustainability.

This second objective is a necessary condition for establishing the durable presence in the ecosystem that is required to be trusted as a dependable source of data. As such, the nonprofit legal status of this project ensures that it can collect the resources needed to achieve its primary mission and only uses them towards that goal.

3. Methodology & scale

This section provides an overview of the information workflow in both TSD APIs, starting with the simplest.

3.1 Spaceflight News API information workflow

The Spaceflight News API, as a small API referencing simple objects from various sources, is a good example of how automation can play an important role. In fact, the entire data ingestion process is automated, i.e. all news articles, blog posts, reports, and other objects in its database are automatically created by a custom-made service detecting new releases on source websites. The database does not actually store all the content of the source items, but only basic information such as the source URL, publisher, title, date, summary, and another URL to the accompanying image.

To allow for wide coverage of space activities through services using both SNAPI and LL2, objects in SNAPI can be linked to LL2 launch and event objects through their unique identifier. This can be used to create e.g. "related news" sections in launch and event pages, making it possible to trace back the history of a space activity from its first mention to the present day. The process of linking SNAPI objects to LL2 objects is currently done fully manually through an administration website, as it requires accurate interpretation of not only the title of the source item, but also its full content. Partial automation of this process could be explored in the future but is not a priority since the current manual process requires little actual time.

On the data consumption side, TSD users perform requests for data by querying the API through standard HTTP requests. The API server then processes the query and collects the data from the backend, before formatting it and returning it in the HTTP response to the user. These API calls are repeated by users with varying frequency depending on their requirement for up-to-date information.
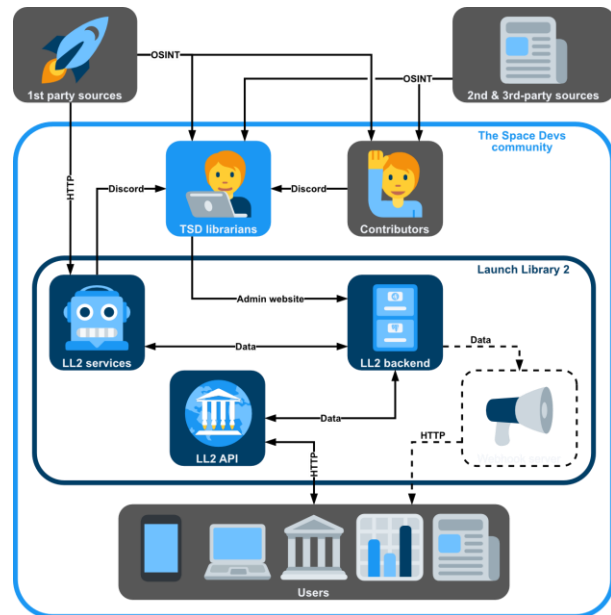

Fig. 1. Workflow of information around LL2

3.2 Launch Library 2 information workflow

Figure 1 schematizes the workflow of information between the TSD community, the set of software products that makes up LL2, and external sources.

These external sources can either be 1st-party, i.e. official sources of information such as launch service providers, space agencies, satellite operators, etc., 2nd-party, i.e. trusted relaters of information such as news outlets and journalists, or 3rd-party, i.e. community-run forums and other individuals. Sources of all three types are followed, and information gathered from them gets analysed either directly by the TSD librarian team, or by contributors who then relay it to the team.

The TSD librarian team performs changes in the dataset through a highly customized administration website. It features numerous visual aids and validation processes that help make data changes as fast and easy as possible, while reducing the risk of human error. These data changes can be made based on information gathered from external sources or contributors, but also from information relayed by automated LL2 services.

These custom-built services can e.g. track websites for information, which is then pre-processed for easy interpretation and relayed to the librarian team. They can also autonomously perform data changes based on 1st-party sources that have proven themselves reliable. As of September 2024, there are 16 permanents trackers

in operation, as well as temporary ones that are automatically enabled during launches. Another important service maintains statistics fields based on other objects in the dataset, e.g. the time spent in space by an astronaut using launch and landing times.

The data consumption pipeline works in the exact same way as for SNAPI. The only difference being that on LL2, the number of API calls is limited on the main API server to 15 hourly per user. This helps prevent abuse and retain control of infrastructure costs incurred by the ever-increasing user traffic, while being sufficient to run a live service that does not constantly require sub-5-minutes data timeliness. For projects consistently requiring a higher access frequency, API keys are made available through subscriptions that help cover server costs. While this free tier rate limit can be acceptable for operational services, it does not allow easy testing of user code during development. That need is covered by a development API server that has a stale and limited dataset but no access rate limits.

In the future, a webhook-based pipeline will allow information to be sent directly to users on data changes only. This will remove the need for repetitive API calls and help reduce the incurred infrastructure costs.

3.3 API documentation & tutorials

For both APIs, interactive documentation webpages are made available to help current and prospective users get familiar with the many data endpoints. These automatically generated pages include exhaustive information on the available sorting, filtering, and search options, as well as specifications for their usage. For example, these options make it possible to find launches from a specific launch vehicle, location, or launch service provider within a given timeframe. They also allow sorting active, all-time, or currently in space astronauts by their total time spent in space or on extravehicular activities.

Additionally, these documentation pages link to a GitHub repository that provides various FAQs and step-by-step tutorials helping beginners get started with TSD APIs. The introduction of these resources, and their direct linking from the web documentation and the Discord server, has proven very effective. Indeed, a steady traffic of more than 50 individuals per week has consistently been achieved on the repository, while the repetitiveness of questions asked on the Discord server and by email has since decreased dramatically.

While making documentation available for humans is essential to help new users understand what data is made available through the APIs and how they can customize their requests, adhering strictly to OpenAPI standards is just as important. This ensures that automatically generated specifications cover all fields of all possible responses from the API, making their formatting and typing fully predictable. Common tools

can then leverage these specifications to effortlessly generate code in any language that is purpose-built to ingest API responses and make the parsed data readily available. This allows developers to focus on what makes their service unique and interesting for end users, instead of spending unnecessary time on data ingestion.

Table 1. Overview of the main objects counts in the SNAPI and LL2 datasets as of September 7, 2024.

Object type (LL2, unless specified)	Occurrences
Blog posts (SNAPI)	1,304
Mission reports (SNAPI)	1,415
News articles (SNAPI)	22,570
Agencies	333
Astronauts	782
Dockings	541
Expeditions	156
Landings	1361
Launch pads	227
Launch sites	63
Launches	7284
Mission patches	963
Reusable first stages	150
Rocket configurations	493
Space programs	39
Space stations	15
Spacecraft	576
Spacecraft flights	849
Spaceflight events	912
Spacewalks	476

Note: only primary objects are listed, many secondary objects requiring little maintenance also exist.

3.4 Scale & temporality of the dataset

To provide a sense of scale of the datasets for both APIs, an overview of the main objects counts is available in Table 1. It lists the number of occurrences of the main objects that contain at least multiple fields that are manually maintained. Secondary objects that only serve as links between main ones are not included, as they do not require nearly as much manual work.

As can be seen in the first part of the table, the vast majority of SNAPI objects are news articles. These are automatically imported, and relevant LL2 launches and events are manually linked to them within a few hours of their import. Generally, between 10 and 30 to 50 articles are imported on each weekday.

The second part of the table illustrates the significant diversity of the LL2 dataset. While launches are at the

core of the data, many other objects revolve around them and are also kept updated in real-time by the librarian team. These can be objects that are directly linked to the rocket launch itself, e.g. the launch site and pad, rocket configuration, reusable stages, or launch service provider. They can also be objects more generally linked to the payload and its mission, be it a satellite, crewed spacecraft, or resupply cargo ship.

On the temporal side, launches are typically added anywhere between days ahead, based on marine navigation warnings or airspace closures, and sometimes up to many years ahead, based on official announcements such as launch contracts. They are then kept up to date by the librarian team from the time of their addition to the dataset until completion of the mission, after which only minor data fixes can still occur. This also applies to all other objects, which are added as needed following new information, and kept updated as long as necessary.

4. Results

As a provider of data that is not destined for the end user, or even humans, but for intermediate digital services, applications, analytics, or experiences, the results and the success of this endeavour are not only measured with the number and the variety of user services, but also and more importantly with the number of people reached worldwide through them.

It is important to note that, without resorting to any sort of advertising or active prospection of potential future users, the main way that TSD gets discovered is through its users, i.e. digital word-of-mouth. In practice, this is usually thanks to users crediting the source of the data in their project. While being a slow discovery method, this has not prevented TSD services from growing steadily in popularity over the years, both in terms of global reach and credibility.

Additionally, TSD requires no registration or authentication to use its services, and as such cannot possess an exhaustive list of its users. The following estimates are based exclusively on the list of known projects and entities using TSD data.

4.1 Global reach

By combining public information on the reach of known TSD users with internal analytics data some accepted to share, it is possible to estimate a total reach worldwide for TSD data of more than 5 million people. A more detailed breakdown is provided in Table 2. It is also known to be a conservative estimate, as the reach data of some high-profile users is neither public nor shared with TSD, and therefore not included.

4.2 Variety of users

The number and variety of its users can be used as a success metric transcribing the popularity of TSD

services in different spheres. The various user types and their respective counts are presented in Table 3.

As can be expected for a fully digital service built around an online community, many TSD users are digital services themselves, such as apps, websites, Discord bots, or a subreddit. Based on the available reach information visible in Table 2, these types are also believed to contribute the most to the global reach achieved by TSD services.

There are however also a few notable users of other types, such as public entities of the space industry (e.g. the US Space Force or the Australian Space Agency), private companies of the space industry (e.g. Kayhan Space), a specialized museum (the Australian Space Discovery Centre), an academic lab from University College London, as well as multiple projects using TSD data as for analytics on space history and the space industry. The reach achieved through these users is in most cases unknown, but the confidence they show towards TSD by using its data represents an important success indicator on its own.

Table 2. Non-exhaustive list of known TSD users with public or specially provided global reach data.

User	Type	Reach
r/SpaceX [7]	Subreddit	> 3 M
Space Launch Now [8]	Apps/Website	1.172 M
OKTO [9]	Discord bot	397 k
SpaceX Launch Bot [10]	Discord bot	230 k
LiveLaunch [11]	Discord bot	138 k
Australian Space Discovery Centre [12]	Museum	> 100 k
Go4Liftoff [13]	App/Website	92 k
Horizon Bot [14]	Discord bot	90 k
RocketLaunch.org [15]	Website	58 k
Space Launch Schedule [16]	Apps/Website	> 50 k
SpaceX Launch Tracker [17]	App	48.5 k
T-Minus [18]	App	19.4 k
Flyoverbot [19]	Discord bot	7 k
ForumAstronautico [20]	Forum	3.8 k
Liftoff [21]	App	> 1 k
Rakietomania [22]	App	> 1 k
SpaceVoid [23]	Apps	> 1 k
Rock It! [24]	App/Website	778
SpaceX Progress [25]	Twitter bot	750
Launch-Stats.com [26]	Analytics	708
Rocket Launch Space Race [27]	App	> 500
Home Assistant [28]	Integration	315
ApoBot [29]	Discord bot	276
TheSpaceXFans [30]	Discord bot	200
Total		> 5.4 M

Note: other known users with no global reach data are included in the counts of Table 3.

Table 3. Distribution of known TSD user types

User service type	Count
Website	18
iOS app	14
Discord bot	10
Android app	9
Analytics	4
Software development kit	2
Home automation integration	2
Others: academic lab, desktop app, museum, podcast, slack bot, subreddit, telegram bot, telegram channel, twitter bot	1 each

Note: as no registration is required to use its services, there is no exhaustive registry of all TSD users.

5. Takeaways

Five main lessons can be drawn from years of TSD operations following an increasingly fast-paced space industry and evolving to serve a growing user base ranging from small enthusiast and hobbyist projects to larger educational and commercial ventures.

5.1 Crowd-sourcing data & making decisions

Core to the data-sourcing process is the combination of crowd-sourced data, e.g. information gathered from social media, forums, and official communication channels, with precedence-based decision-making by an experienced team that always remains aware of the current context. This ensures both timeliness in reacting to new information from any source, as well as accuracy and long-term consistency in how this information is interpreted when making changes in the dataset.

5.2 Financial sustainability of public resources

To ensure long-term financial sustainability of the project, actively encouraging the conservative use of public resources has proved essential. For LL2, this translates into limiting the access rate on the main API to prevent abuse, while setting up a secondary API with a limited dataset but no access limits that can be used by both current and prospective users during their development efforts. This method has proved effective in moderating the incurred infrastructure costs. Additionally, making higher access rate limits on the main API available through subscriptions has made it possible for both generous supporters and revenue-generating projects that require faster access rates to cover the incurred infrastructure costs.

5.3 Integrating with the ecosystem

To make it possible for developers and analysts to more comprehensively and accurately cover space activities, integration between TSD data and information from other established sources has been made possible. For example, the corresponding international launch designator (also known as COSPAR ID) is provided with every launch object in LL2, allowing easy cross-matching with many other sources such as the NSSDCA Master Catalog [1], Space-Track [2], CelesTrak [3], Jonathan McDowell's GCAT [4], or Gunter's Space Page [5]. A direct link to the trajectory data computed and provided by Flight Club [6] is also included. These integrations with the ecosystem have greatly contributed to the confidence that is presently awarded to the data provided by TSD.

5.4 Instant & informal communications

Fostering the development of a community around the services provided by TSD has always been an important aspect of this project. This can be seen through the evolution of the data and features provided, many of the changes having been implemented following user feedback and requests. It can also be regarded as having played a key role in the growing interest showed for TSD over the years. To create this community around the use of TSD services, instant and informal communication between users and team members has been made possible using a Discord server. Most of the conversation, user reports, and feedback are centralised in this server. However, more conventional communication channels such as social media or emails remain in use for public announcements and exchanges with the team. Over time, this has contributed to the emergence of a collaborative community, where established users regularly help prospective ones get started with their project, and proved an efficient way to acknowledge and implement timely data and code corrections as they are reported.

5.5 Workload management through automation

Keeping track of an increasingly large and fast-paced space industry with a dataset that also evolves with new features over time implies an increasing workload for the TSD team. To counterbalance these trends, the manual librarian workflow is under constant scrutiny to identify potential areas of improvement. In recent years, this has led to many changes in the UI of the administration panel used for manual data changes, helping reduce unnecessary or repetitive steps. It has also led to the identification of patterns where a small number of reliable sources regularly triggered similar data changes (e.g. with regular SpaceX launches). These patterns have been at the core of developments of finely tuned automations that can help reduce the overall manual workload, for example by pre-processing the source information, while the human team retains full control of data changes. These improvements have consistently proved beneficial in handling the increasing workload with a small team of volunteers without compromising on data accuracy or timeliness.

6. Conclusions

The primary goal set for this project is to advance worldwide interest in spaceflight by improving the accessibility of information. To achieve this mission, standardized data retrieval portals make extensive real-time datasets on the space industry available to creators of public-facing services. These range from hobbyist app and website developers to content creators, teachers, academic groups, museums, analytics firms, space agencies, and more.

Public reach data demonstrates the significant presence achieved by TSD to date, extending well into the multiple millions globally. This figure is likely to keep increasing in the coming years with the growing popularity of spaceflight, while the data offered will also keep improving in quality and ease of use.

The gradual development and constant growth of this project as well as of its surrounding community have also allowed the team to gather valuable insights and experience. This work elucidates several key takeaways that can help foster new ideas and endeavours contributing to public interest in spaceflight.

Overall, it can be stated that this nonprofit stemming from a global team of passionate volunteers has achieved a measurably positive impact on the landscape of space outreach and education.

Acknowledgements

First and foremost, this work is dedicated to all members of the TSD team past and present, whose passion and commitment have been the driving force behind this endeavour since day one.

Secondly, this project deeply thanks all its financial supporters, without whom it would not be possible.

Much appreciation also goes to contributors in the TSD Discord server and other individuals in the online space community who help track and document the ever-changing world of spaceflight.

Finally, this project owes deep recognition to its many users, whose interest, engagement, and feedback over the years have been instrumental in making it successful.

Appendix (Examples of TSD user services)

This section presents pictures of several public-reaching services leveraging data provided by TSD.
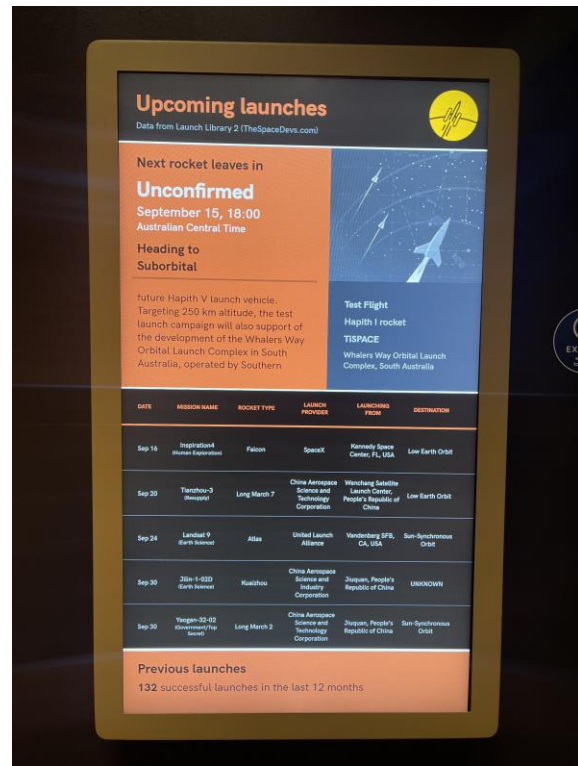
Fig. A1. Real-time list of upcoming space launches at the Australian Space Agency's Australian Space Discovery Centre [12]
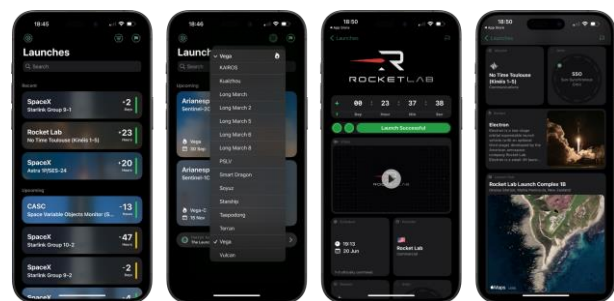
Fig. A2. Departing Earth iOS app [31]
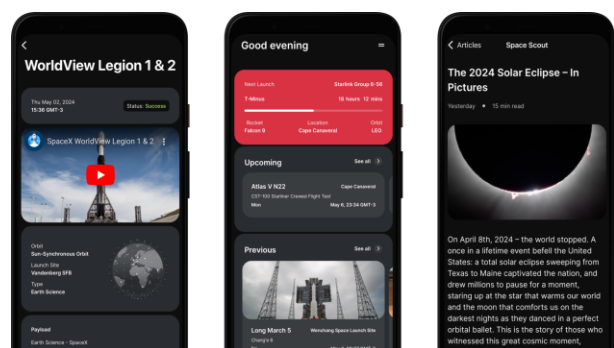
Fig. A3. Liftoff Android app [21]

Fig. A4. Space launch countdown box by L. Stauder
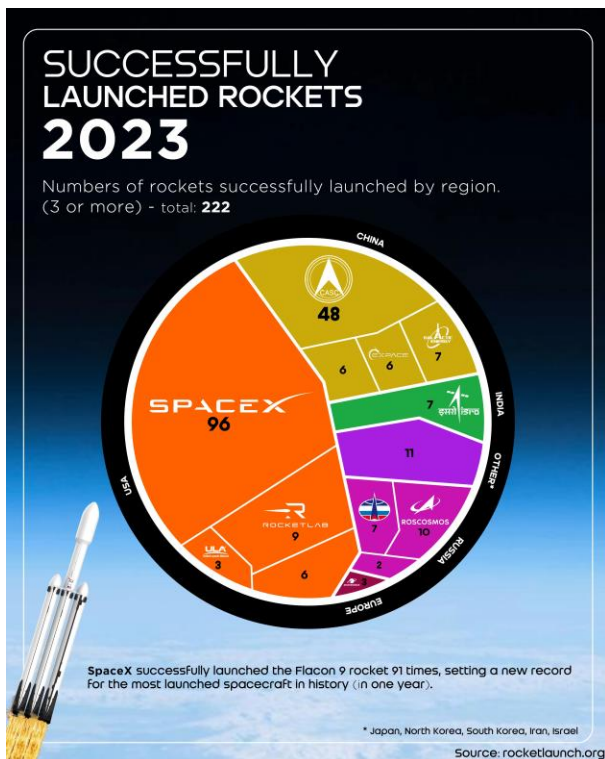
Fig. A5. Infographic by RocketLaunch.org [15]
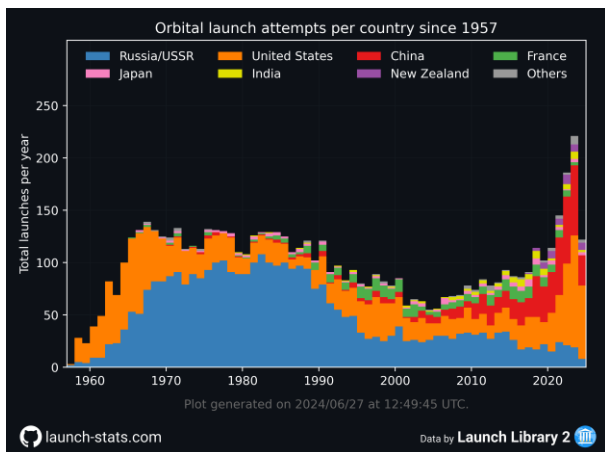
Fig. A6. Launch statistics on Launch-Stats.com [26]

Fig. A7. Upcoming launches list on the Everyday Astronaut website [32]
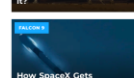

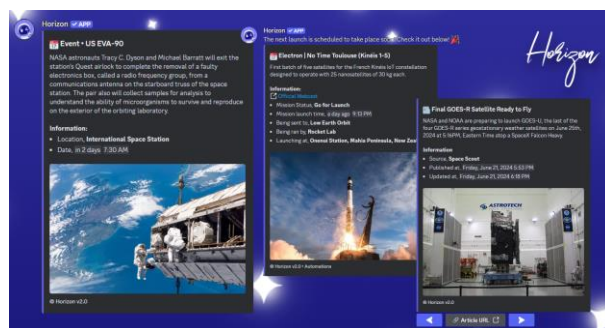Fig. A8. Upcoming launches list on the Primal Nebula website [33]
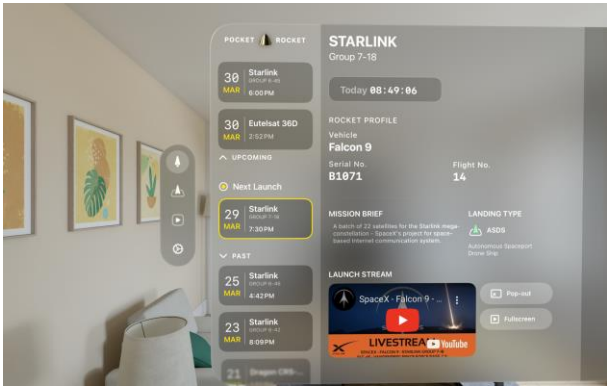

Fig. A9. Horizon Discord bot [14]

Fig. A10. Pocket Rocket VR app [34]

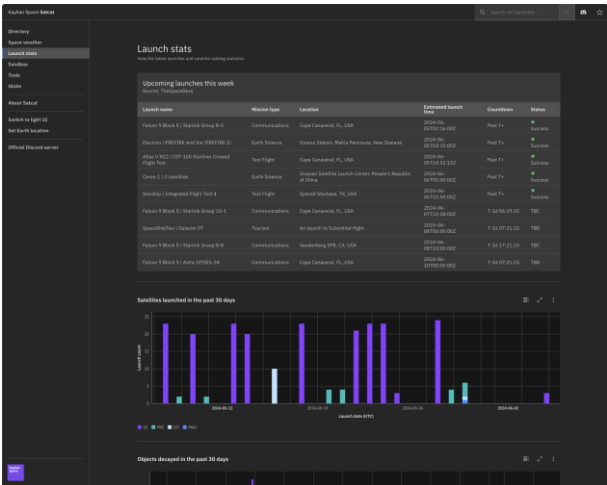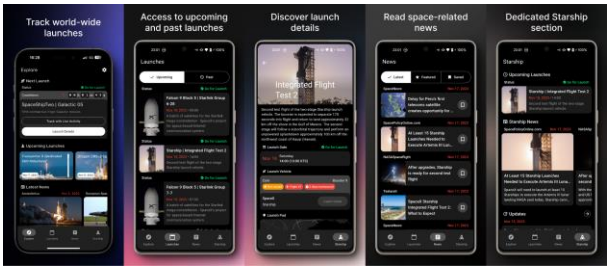Fig. A11. Kayhan Space's Satcat [35]
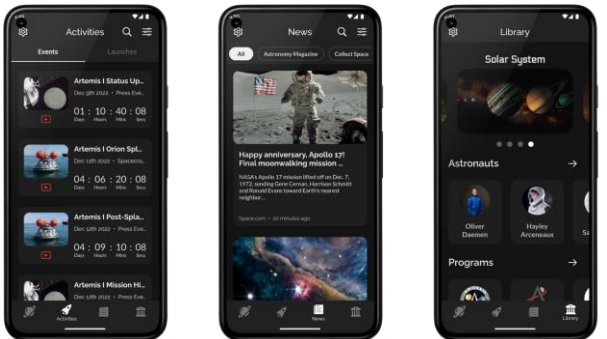

Fig. A12. StarCat iOS & Android apps [36]

Fig. A13. SpaceVoid Android app [23]
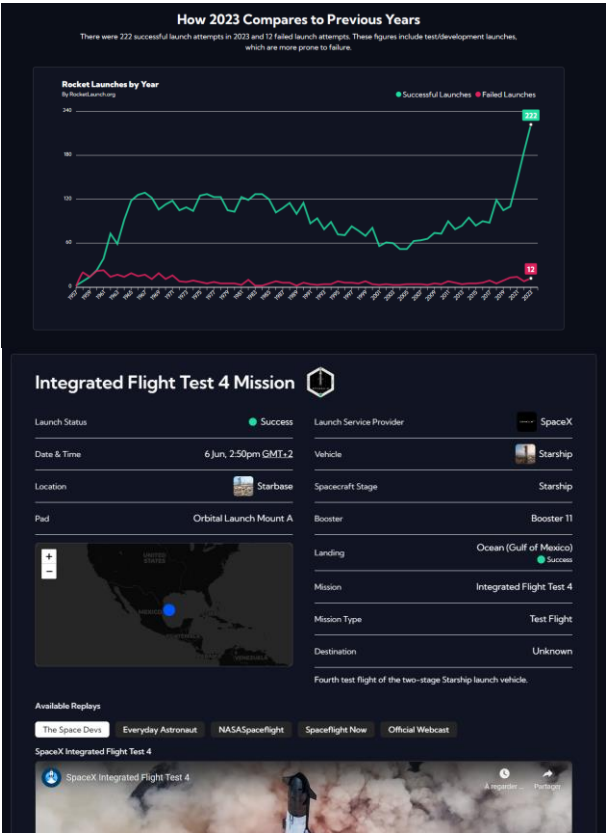
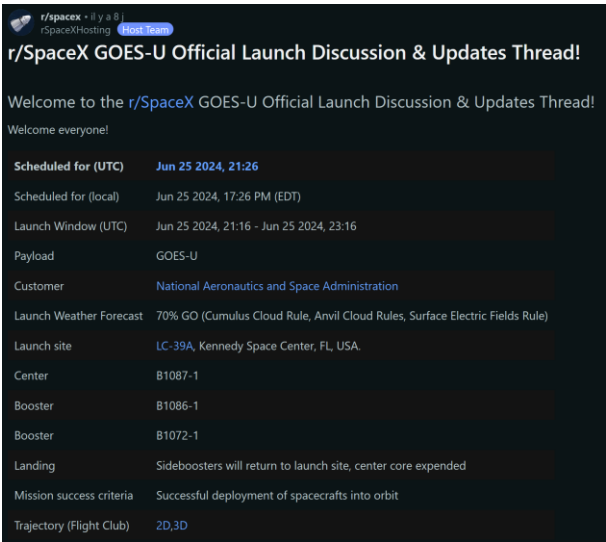Fig. A14. All-time launch statistics and launch page on RocketLaunch.org [15]

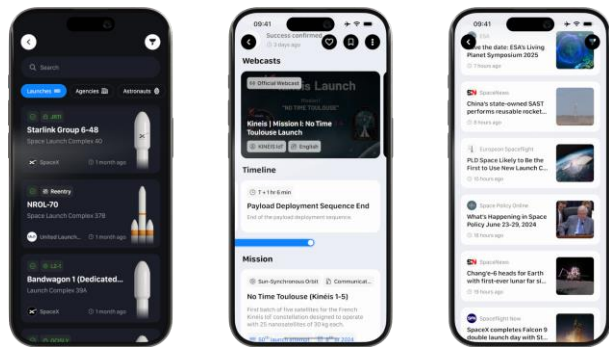Fig. A15. Automated launch thread on the r/SpaceX subreddit [7]
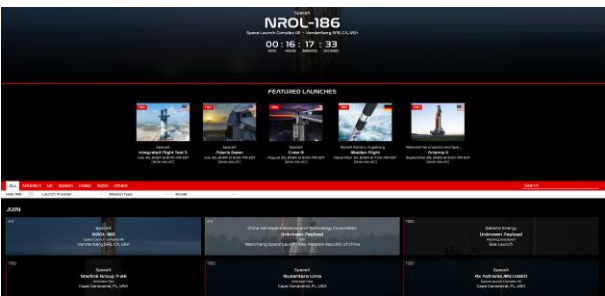
Fig. A16. T-Minus iOS app [18]

Fig. A19. Upcoming launches list on the website of The Launch Pad news network [38]
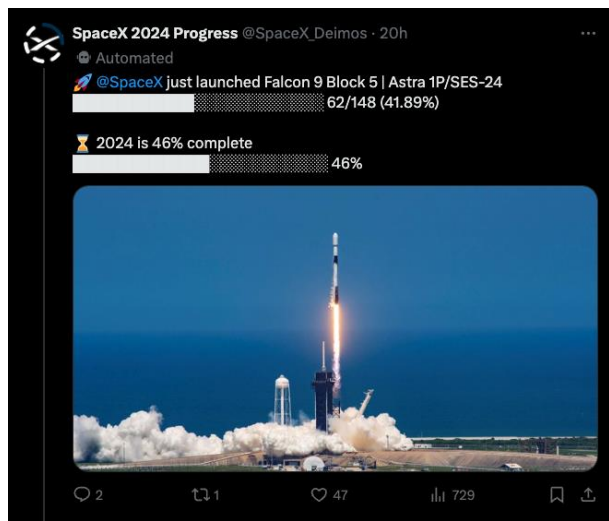

Fig. A17. Twitter bot tracking SpaceX' progress towards their yearly launch objective [25]
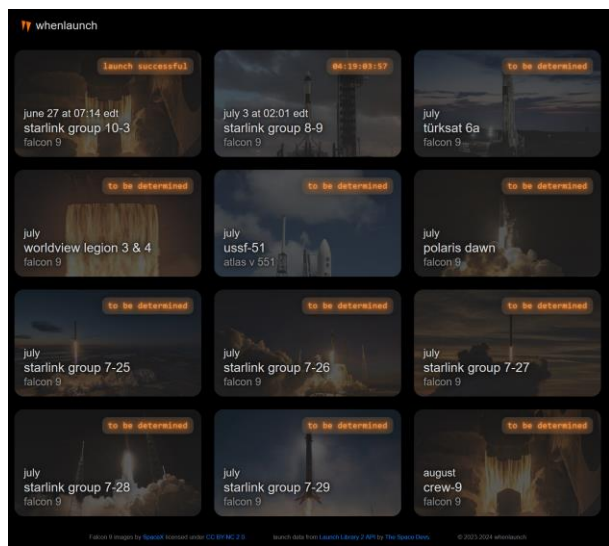

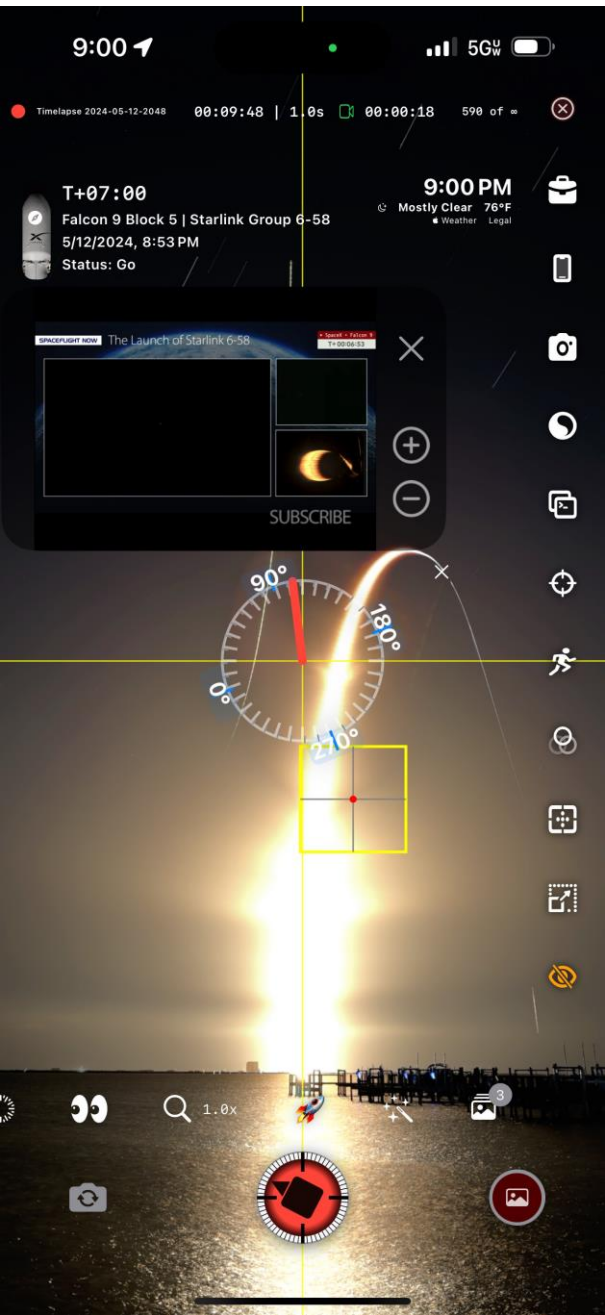Fig. A18. Upcoming launches list on whenisthelaunch.com [37]

Fig. A20. Timelapsed iOS app [39]

References

[1] NASA Space Science Data Coordinated Archive Master Catalog, https://nssdc.gsfc.nasa.gov/nmc/

[2] US Space Force 18th Space Defense Squadron, Space-Track, https://www.space-track.org

[3] T.S. Kelso, CelesTrack, https://celestrak.org/

[4] J. McDowell, General Catalog of Artificial Space Objects, https://planet4589.org/space/gcat/

[5] G. D. Krebs, Gunter's Space Page, https://space.skyrocket.de/index.html

[6] D. Murphy, Flight Club, https://flightclub.io/

[7] r/SpaceX, https://www.reddit.com/r/spacex/

[8] C. Jones, Space Launch Now, https://spacelaunchnow.me/

[9] CalliEve, OKTO, https://github.com/CalliEve/okto

[10] S. Jenner, SpaceX Launch Bot, https://spacexlaunchbot.dev/

[11] S. Goldebeld, LiveLaunch, https://livelaunch.juststephen.com/

[12] Australian Space Agency, Australian Space Discovery Centre, https://www.space.gov.au/australian-space-discovery-centre

[13] J. Rascagneres, Go4Liftoff, https://go4liftoff.com/

[14] Atlas Team, Horizon, https://horizonbot.xyz/

[15] RocketLaunch.org, https://rocketlaunch.org/

[16] Space Launch Schedule, https://www.spacelaunchschedule.com/

[17] Z. Hadjineophytou, SpaceX - Launch Tracker, https://play.google.com/store/apps/details?id=uk.co.zac_h.spacex

[18] TARS Technologies, T-Minus, https://apps.apple.com/us/app/t-minus-space-launch-tracker/id1496516813

[19] Flyoverbot, https://discord.com/invite/arHKCRR

[20] M. Zambianchi, ForumAstronautico, https://www.forumastronautico.it/

[21] V. Moreira, Liftoff: Rocket Launch Tracker, https://play.google.com/store/apps/details?id=com.viniciusmoreria.liftoff

[22] D. Śledź, Rakietomania, https://play.google.com/store/apps/details?id=com.danielsledz.rakietomaniaapp

[23] Limitnil, SpaceVoid, https://www.spacevoid.limitnil.com/

[24] Rock It!, https://rockit.010.one/

[25] Robin, SpaceX Progress, https://x.com/SpaceX_Deimos

[26] A. Muller, Nosu's LL2 Launch Stats, https://launch-stats.com/

[27] Museumis, Rocket Launch Space Race, https://play.google.com/store/apps/details?id=com.museumis.rocketlaunchspacerace

[28] J. Sørensen, Launch Library Home Assistant Integration, https://www.home-assistant.io/integrations/launch_library/

[29] J. Keegan, ApoBot, https://www.home-assistant.io/integrations/launch_library/

[30] J. Keegan, TheSpaceXFans, https://discord.com/invite/HWabhNtsr9

[31] M. Folbigg, Departing Earth – Spaceflight, https://apps.apple.com/gb/app/departing-earth-spaceflight/id1641494756

[32] T. Dodd, Everyday Astronaut, https://everydayastronaut.com/

[33] Primal Nebula, https://primalnebula.com/upcoming-launches/

[34] T. Isenman, Pocket Rocket, https://apps.apple.com/us/app/pocket-rocket/id1493916909

[35] Kayhan Space, Satcat, https://www.satcat.com/

[36] H. Jóźwiak, StarCat, https://www.producthunt.com/products/starcat

[37] Whenlaunch, https://whenisthelaunch.com/

[38] The Launch Pad, https://tlpnetwork.com/launches

[39] Shutter Speed Apps, TimeLapsed – Time Lapse Camera, https://apps.apple.com/us/app/timelapsed-time-lapse-camera/id6479436667