

データサイエンス

すいすい会 #37

Pythonプログラマから
プロダクトオーナーへの道！

Streamlit を利用した
高速プロトタイピング入門



すいすい会の紹介

データサイエンス、データドリブン組織 etc.

データ利活用をがんばるデータの猛者たちのための雑談会です！

悩み
を共有する

知見
を共有する

理解
を高める



実践を重視するGRIの考え方を雑談しながらお届けします

資料: [GRIホームページ](#)

Slack: [ForecastFlowチャンネル](#)

本日のナビゲータ

大友 祐一
株式会社GRI
データサイエンス事業部 取締役



略歴

- 2015年: 東京工業大学大学院
理化学研究所 仁科加速器研究センター
- 2015年: 株式会社GRIに入社
AI基盤 ForecastFlowの開発などAIの社会実装に従事
- 2022年: 株式会社GRI 取締役就任

過去のすいすい会の動画 (YouTube)

<https://www.youtube.com/c/GRIinc>

チャンネル登録 / いいね！ / 高評価
よろしくお願いします！



今回はどんな回？

Pythonプログラマからプロダクトオーナーへの道！ Streamlitを利用した高速プロトタイピング入門

どんな人に向けて？

Pythonプログラマ

アプリ開発者

データサイエンティスト

どんなことを話す？

プロダクト開発の考え方

- ・GRI流のプロダクトの育て方
- ・プロダクトオーナーの役割

Streamlit の紹介

- ・プロトタイプ開発での有効性
- ・その強みと限界



[#すいすい会](#) で質問/感想どんどんツイートしてください！

本日の内容

思想

- プロダクトが花開くまでのフィットジャーニー
- プロダクトオーナーというお仕事
(マネージャとの違い)

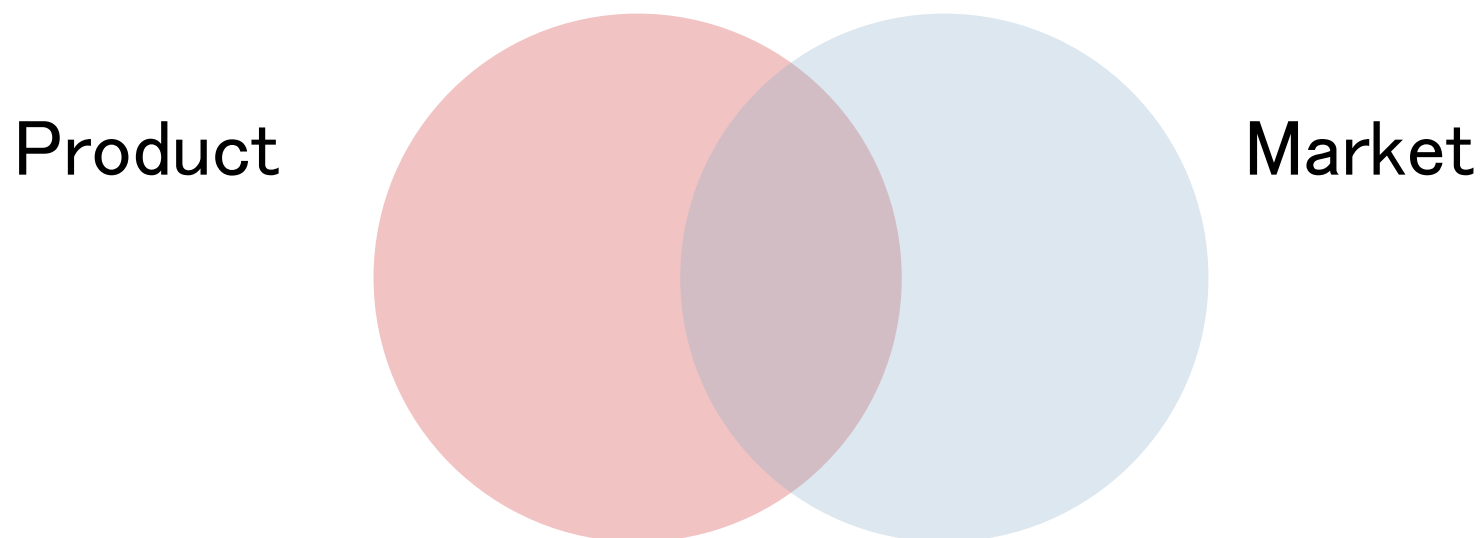
技術

- Streamlitとは？
高速プロトタイピングのためのツール紹介
- Streamlitアプリのクイックな開発方法
- 試作品から本開発への道

プロダクトのフィットジャーニー

PMF

プロダクト・マーケット・フィット



PMFとは



Marc Andreessen

wikipediaより写真参照

米国のベンチャーキャピタル
Andreessen Horowitz (a16z)
の共同創業者



Part 4 The only thing that matters

Posted on June 25, 2007

This post is all about the only thing that matters for a new startup.

But first, some theory:

If you look at a broad cross-section of startups—say, 30 or 40 or more; enough to screen out the pure flukes and look for patterns—two obvious facts will jump out at you.

First obvious fact: there is an incredibly wide divergence of success—some of those startups are insanely successful, some highly successful, many somewhat successful, and quite a few of course outright fail.

スタートアップの成功で最も重要なことはPMFすること

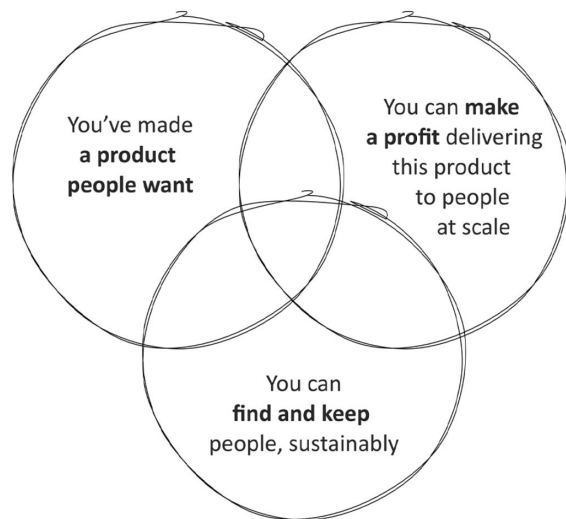
PMFとは、良い市場にいて、かつ、その市場を満足させられるプロダクトを持っていること

PMFを達成するとは、どんな感じ？

TECHBLITZ

FEATURE CATEGORY REPORT EVENT

When do I have
True Product-
Market Fit?



Netflix、Uber、Airbnb、Dropboxの
創業者たちが語る
「プロダクトマーケットフィットの瞬間」

Dropbox

「高速で走る車や離陸する飛行機に乗っている時に、
座席の後ろに押し付けられる感じ」

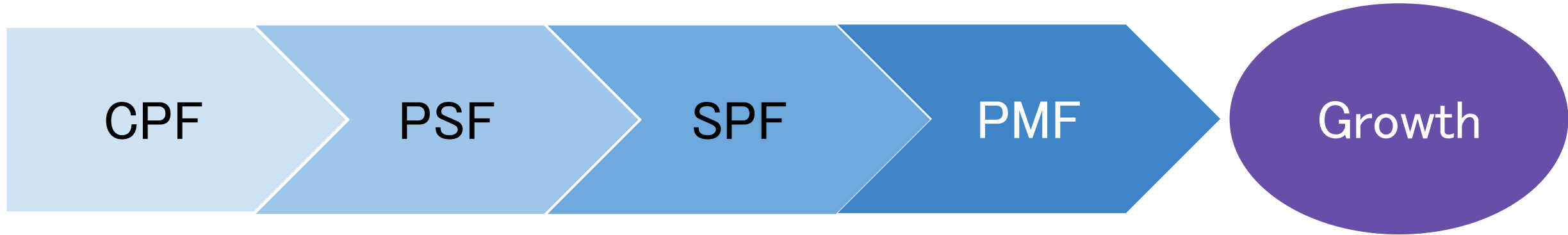
Netflix

「以前はトラフィックを得るのに苦労していたが、
突然、対応しきれないほどのボリュームに到達。
驚異的なほど余っていた在庫が、突然足りなくなった。
すべてがうまくいき始めた。」

Github

「驚いたことに、ユーザーから
『これにお金を払ってもいいですか？』と言われた。
これがうまくいきそうな最初の兆候だった。」

PMFに到達するまでのフィットジャーニー



CPF: Customer Problem Fit
顧客課題の検証

PSF: Problem Solution Fit
解決策の検証

SPF: Solution Product Fit
プロダクトの実装検証

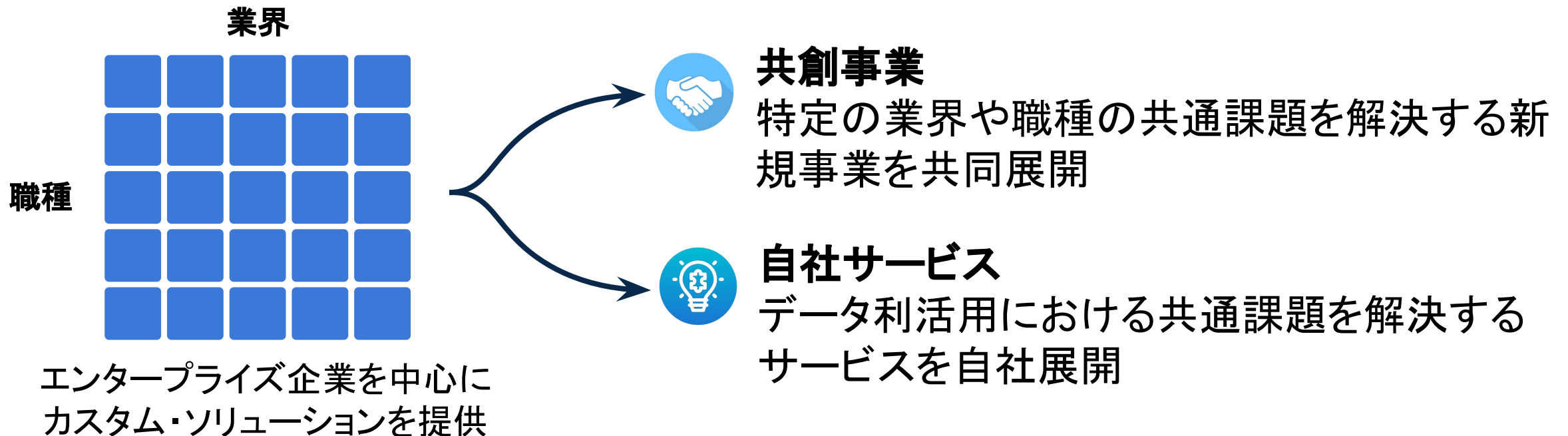
PMF: Product Market Fit
顧客の定着化の検証

GRIの場合

株式会社GRIについて



データで新たな事業を開発していく
カンパニー。



事業コンセプト

世界中のデータを資産化する

売上拡大

- ・ 既存サービスの改善
- ・ AI/BIによる営業効率の向上
- ・ データ販売

データ資産により創出された
経済的な成果を最大化

データ利活用コスト

- ・ 人件費/外注費
- ・ 人材採用/育成コスト
- ・ データ取得および分析コスト
- ・ システムの維持管理コスト

新たな付加価値の創出方法は千差万別のため、
個別提案による**コンサルテーション**を提供

コスト構造は業界や業種を問わず似ているので、
SaaSプロダクトや**データ人材の研修講座**として提供

自社サービス

AutoML(自動機械学習)サービス



- 機械学習モデルを構築するために最も重要なことは、いかに良い訓練データを構築するか
- 機械学習モデルのための環境構築や決まりきった前処理、パラメータチューニングは自動化
- 表形式データの予測のみに絞る、コンセプト検証から組織的な本運用まで一貫して利用できる機械学習プラットフォームはコストが高いという常識からの解放

分析拡張データの提供サービス



- 従来高価だった地理データを始め、オープンデータをGeoJSON等のデータ分析に適した構造で提供
- 郵便番号を基準に自社保有データと統合を可能にすることで、日本の名寄せの闇からの解放
- 専用ソフトを前提とせず全社共有も簡単で、更にデータ提携企業として他社販売も可能

GRIにおけるフィットジャーニー

	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6
	課題発見	価値検証	プロトタイプ 開発	PMF 試行錯誤	スケール 検討	グロース
旅程	お金を支払ってでも解決したい 課題かどうかの検証		プロトタイプをク イックに開発 本開発するか 投資判断	最低限の機能を備 えたプロダクトを ローンチ バリュー・プロポジ ションの確立 カスタマー サクセスの追求	カスタマー サクセスの パッケージ化 採用・育成 方法の検討	チャネル多角化 採用・育成のドラ イブ
手段	受託開発		受託開発 社内勉強会 すいすい会	受託開発 展示会 オープンイノベー ション	パートナー との提携	営業組織の構築

GRIにおけるフィットジャーニー

本日のトピック

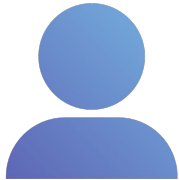
	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6
	課題発見	価値検証	プロトタイプ開発	PMF 試行錯誤	スケール検討	グロース
旅程	お金を支払ってでも解決したい課題かどうかの検証		プロトタイプをクイックに開発 本開発するか投資判断	最低限の機能を備えたプロダクトをローンチ バリュー・プロポジションの確立 カスタマーサクセスの追求	カスタマーサクセスのパッケージ化 採用・育成方法の検討	チャネル多角化 採用・育成のドライブ
手段	受託開発		受託開発 社内勉強会 すいすい会	受託開発 展示会 オープンイノベーション	パートナーとの提携	営業組織の構築

フィットジャーニーとチーム構成

Step 1	Step 2	Step 3	Step 4
課題発見	価値検証	プロトタイプ開発	PMF 試行錯誤

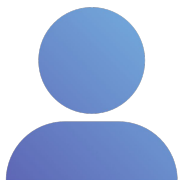


顧客企業



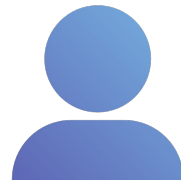
プロジェクト
マネージャ

- ・課題整理
- ・要件定義
- ・提案
- ・タスク管理



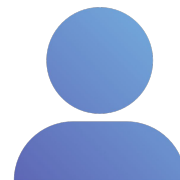
アナリスト

- ・実験デザイン
- ・手法選定
- ・結果解釈



エンジニア

- ・環境構築
- ・開発
- ・実験実施
- ・精度評価



プロダクト
マネージャ

- ・市場調査
- ・競争優位性
- ・予算/人材管理
- ・広報戦略

フィットジャーニーとチーム構成

Step 1	Step 2	Step 3	Step 4
課題発見	価値検証	プロトタイプ 開発	PMF 試行錯誤

受託開発

本日の主役

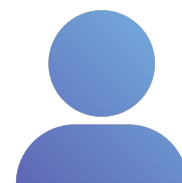
開発

顧客企業



プロジェクト
マネージャ

- ・課題整理
- ・要件定義
- ・提案
- ・タスク管理



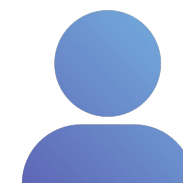
アナリスト

- ・実験デザイン
- ・手法選定
- ・結果解釈



エンジニア

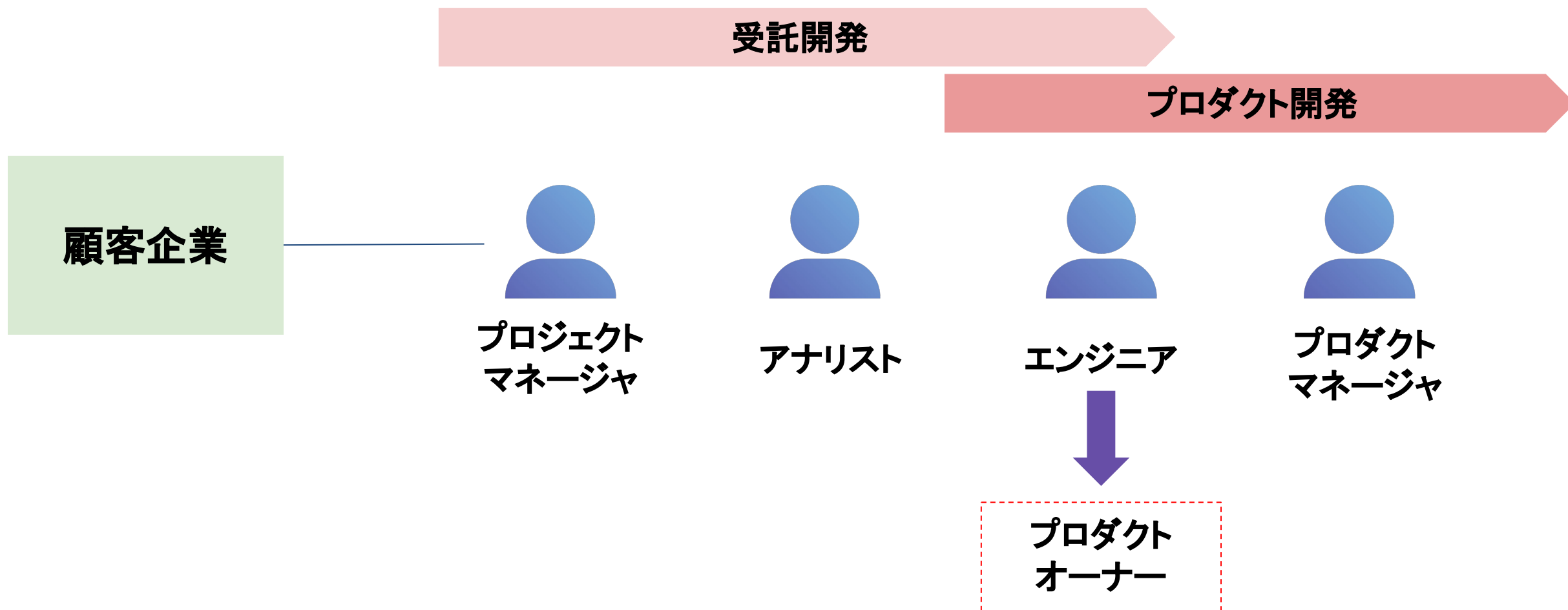
- ・環境構築
- ・開発
- ・実験実施
- ・精度評価



プロダクト
マネージャ

- ・市場調査
- ・競争優位性
- ・予算/人材管理
- ・広報戦略

プロダクトオーナーになるためには？



プロダクトオーナーの役割

プロダクトオーナーとプロダクトマネージャの違い

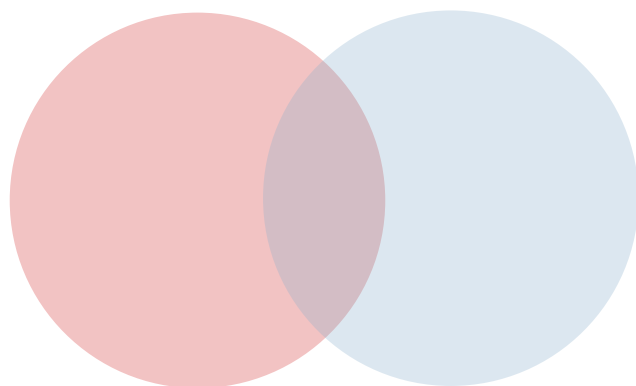


プロダクト開発の進め方に関して責任を持つ



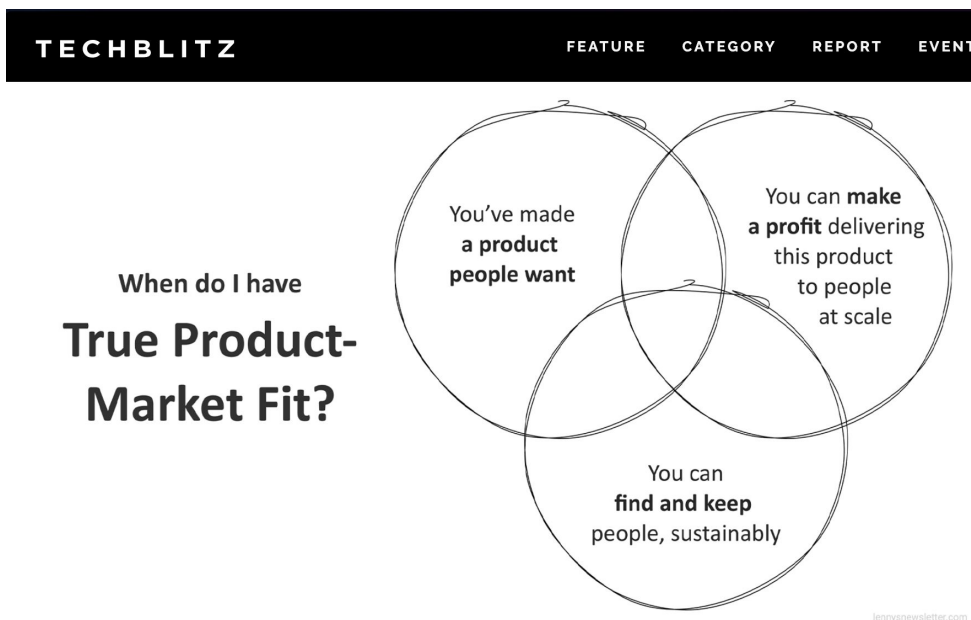
市場や顧客ニーズを理解し、全体戦略を策定する

技術シーズ指向



顧客ニーズ指向

PMFまでの期間



PMFまでに必要な期間は、
その時の市場規模と
製品の素晴らしさに依存

Netflix: 1.5年

Segment: 1.5年

Airbnb: 2年

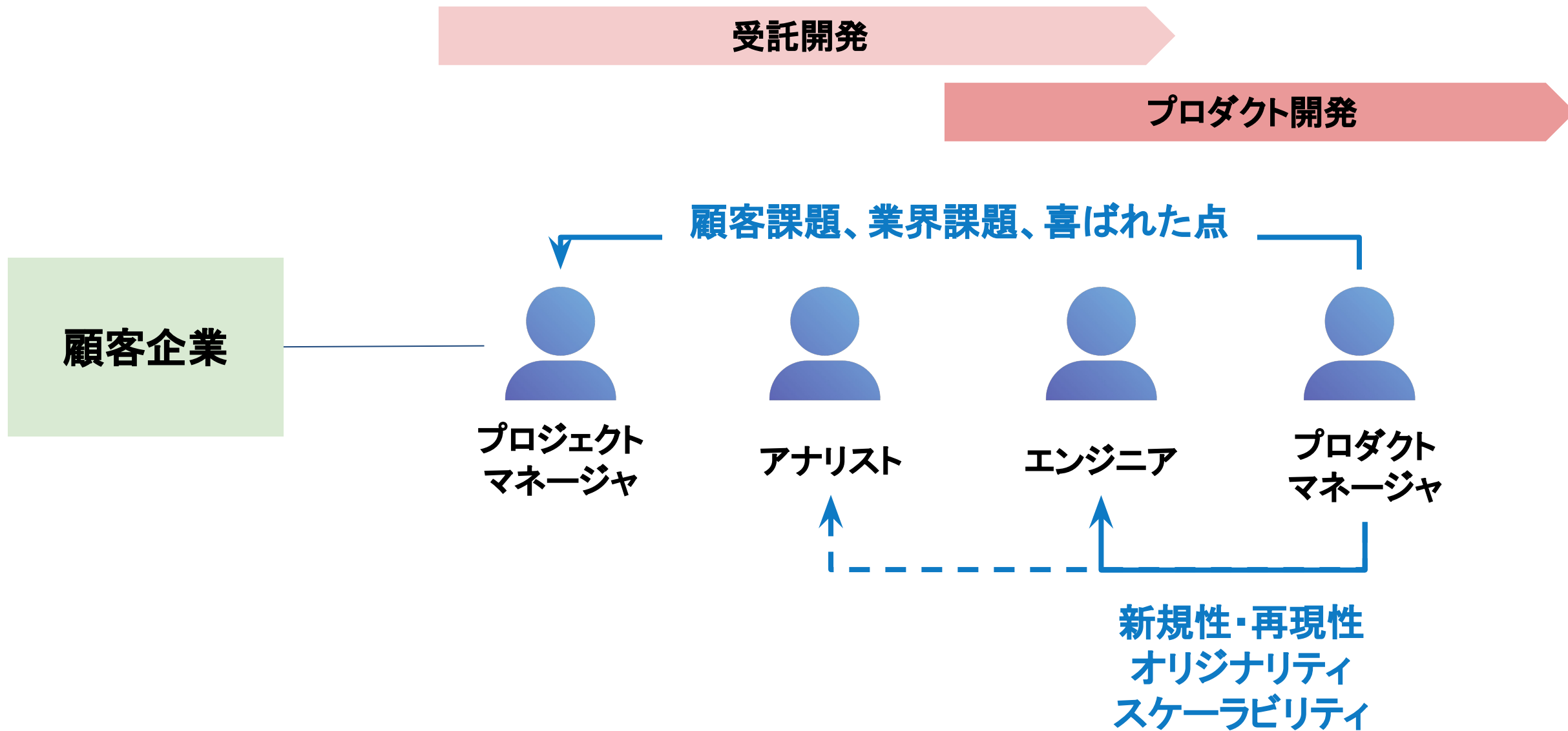
PagerDuty: 2年

Superhuman: 3年

Amplitude: 4年

短期／中期／長期などの各テーブルで
プロダクト開発のポートフォリオ形成

受託開発後の社内コミュニケーション



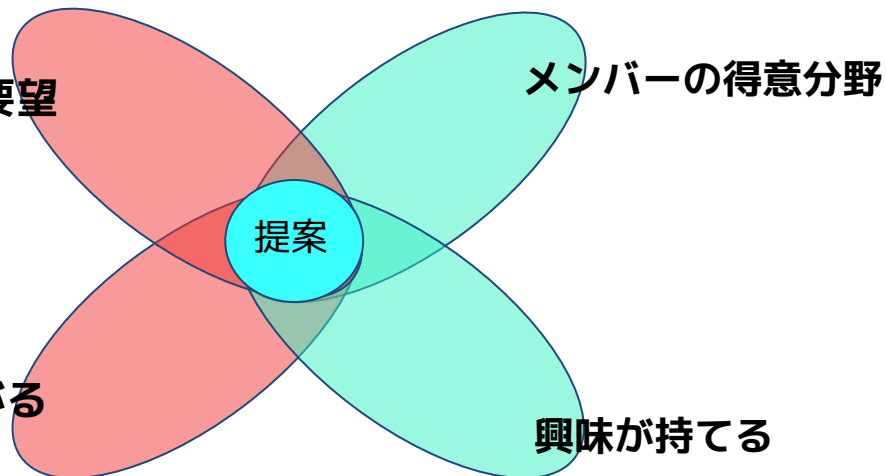
プロダクト開発する前提の受託提案のつくりかた

- プロジェクトマネージャは、顧客満足度とGRIメンバーの満足度をなるべく同時最適にする提案を模索
- 取り組む意義やチャレンジする内容を狩野モデルなどを利用して明示
 - 特に、魅力的品質は差別化要因になりやすい

顧客

おぼろげな要望

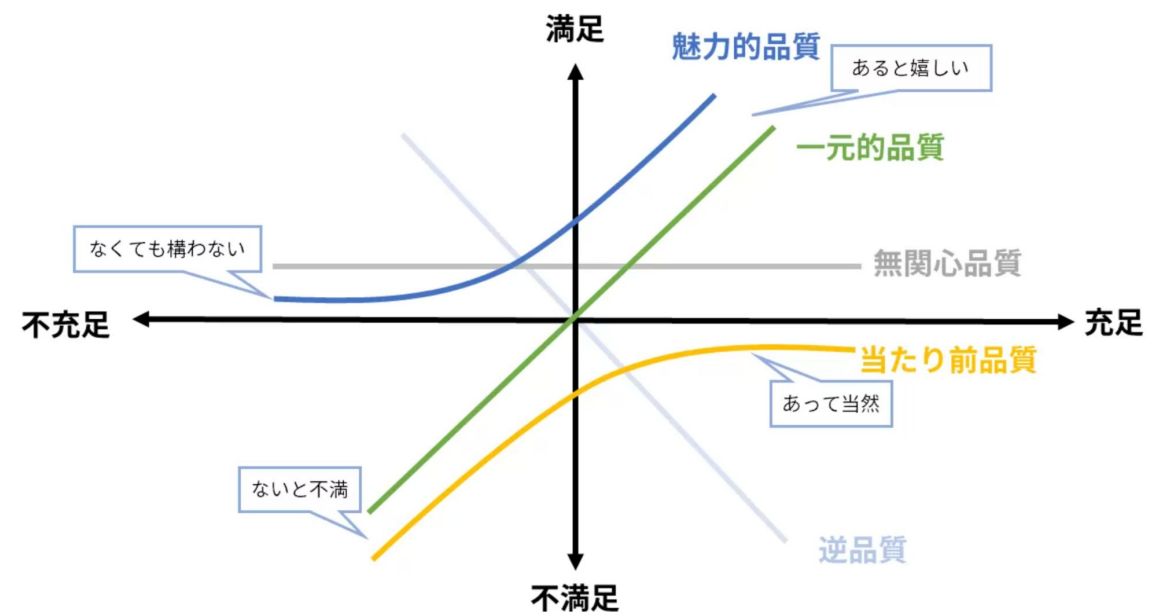
利益に繋がる



GRI

メンバーの得意分野

興味を持てる



GRIの場合の、プロダクトオーナーへの道

- プロトタイプからMVP (Minimum Viable Product)を開発し、PMFを目指すかどうか、他のメンバーと協議が必要
- 生々しくお話しすると、受託開発にアサインすれば短期的に売上につながるエンジニアをプロダクト開発にアサインして投資するか
- 他のメンバーが任せたい・託したいと思える情熱や説得力を感じられるかが重要
- 各マネージャとコミュニケーションをとりながら進めるのが重要

GRIの場合の、プロダクトオーナーへの道

HACKY MONDAY

隔週月曜日のランチタイム開催

受託開発等で得られた知見から
開発したプロトタイプを社内共有

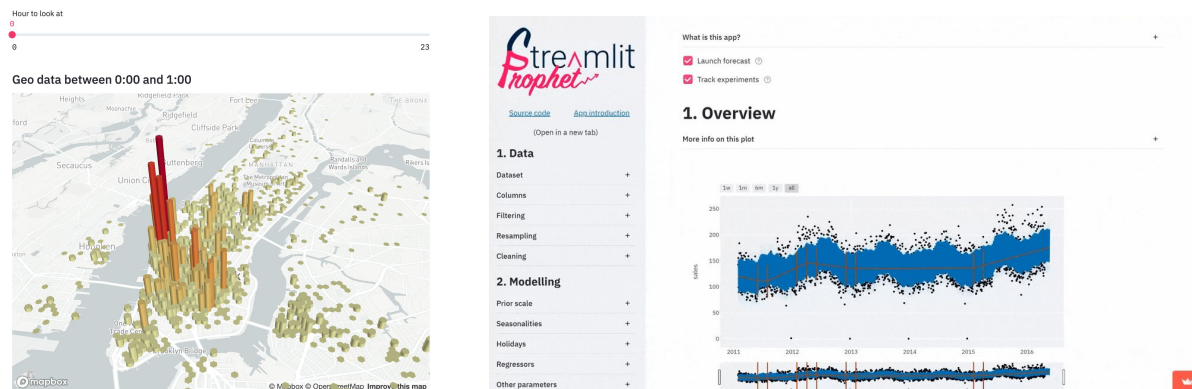
外部の企業・大学から招待し、
パートナー提携や協業検討することも

筑波大学 N氏	物質科学とMLOps入門
GRI 大友	Streamlitの紹介
Google Cloud I氏	データメッシュ アーキテクチャ
愛媛大学 T氏	自然言語処理における 言い換え表現の研究
GRI 大友	プロトタイプ紹介
GRI 松澤氏	プロトタイプ紹介
データX E氏、K氏	b-dash紹介

Streamlitを利用した高速プロトタイピング

Streamlit でできることのイメージ

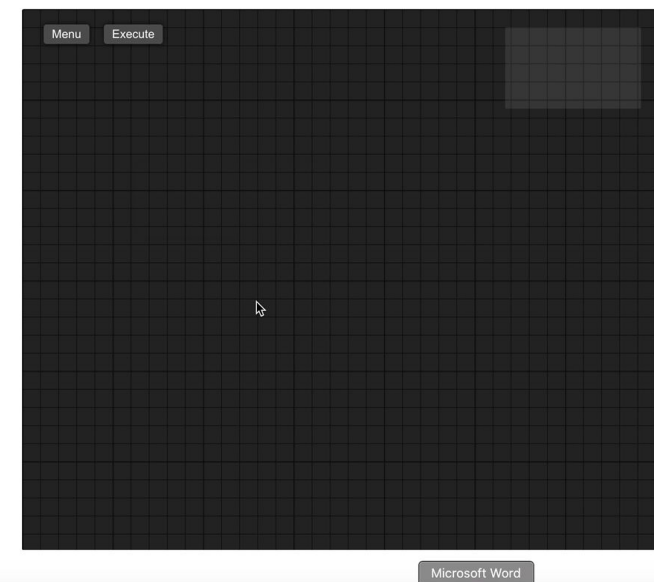
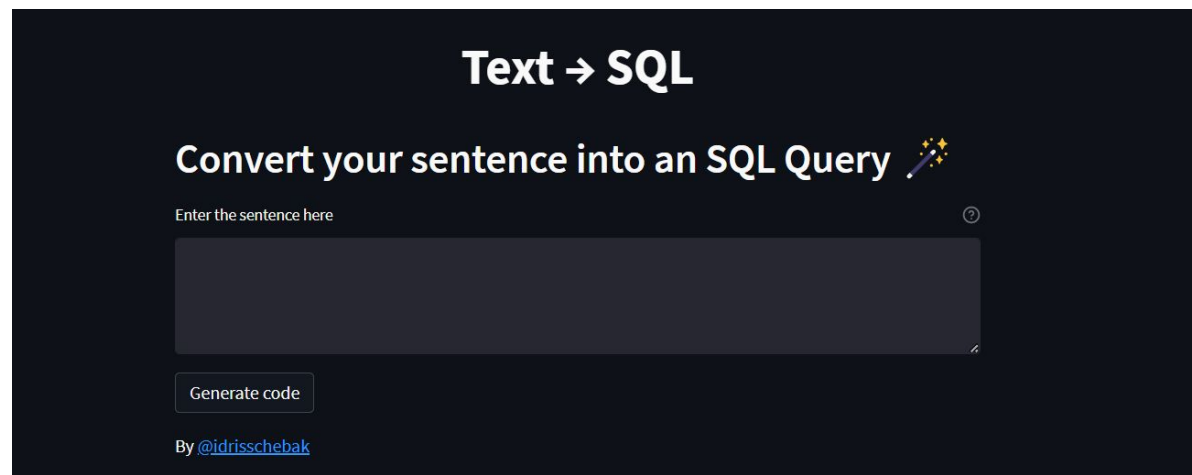
分析結果の可視化



データの作成と編集

	name	age	gender	is_active	status	homepage
	Kelly Kelley	75	female	<input checked="" type="checkbox"/>	0.71	http://edwards.com/
	Nicole Nguyen MD	9	other	<input checked="" type="checkbox"/>	0.47	https://www.cole.net/
	Ethan Turner	39	male	<input type="checkbox"/>	0.6	https://www.baird-garner.info/
	Todd Burton	28	female	<input checked="" type="checkbox"/>	0.26	https://www.porter.biz/
	Justin Garcia	89	other	<input type="checkbox"/>	0.9	http://ward-romero.org/

訓練済みモデルへの I/O

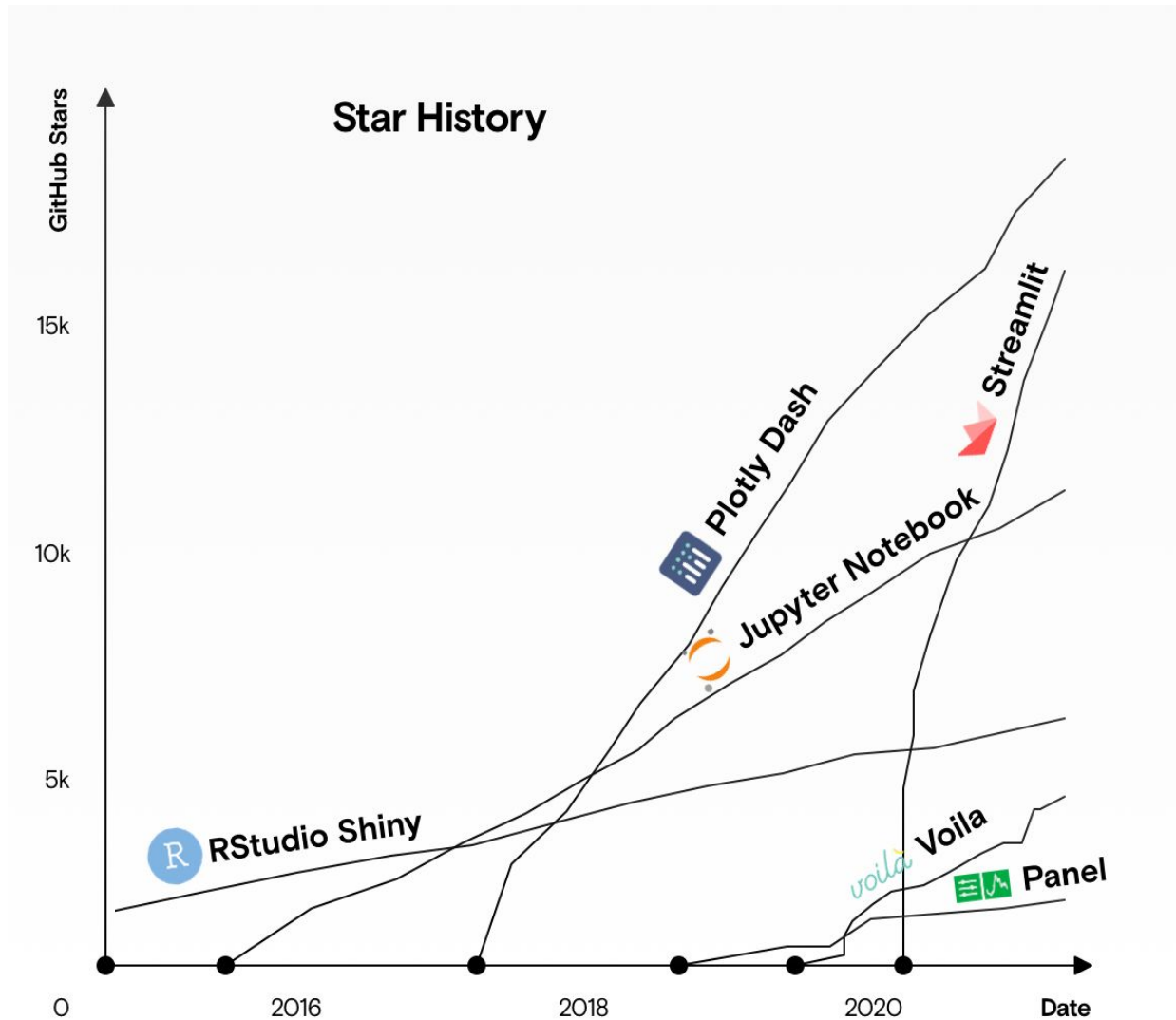


Streamlit ここが便利

- 開発する際に、ソースコードを編集すると即座にアプリ画面に反映されるので、開発やデバッグが高速（いちいちビルドしなくて良い）
- 良く使うUIコンポーネントが標準ビルトイン
 - 選択ボックス、ラジオボタン、ファイルアップロード、*etc.*
- 主要なデータ可視化ライブラリもすぐに連携可
 - matplotlib, plotly, seaborn, bokeh, graphviz, map

クイックにプロトタイプを構築可能な要素を多く含んでいる

データサイエンスのダッシュボードOSS の比較



[Data dashboarding tools](#)

[Streamlit v.s. Dash v.s. Shiny vs. Voila vs. Flask vs. Jupyter](#)

Streamlit Community Cloud

Cloud Gallery Components Generative AI Community Docs Blog

Sign in Sign up

App Gallery

These are some of our favorite apps created by Streamlit users and hosted on Streamlit Community Cloud. Try them out, browse their source code, share with the world, and get inspired for your own projects 🥳

Want to build your own? [Get started today!](#)

CATEGORIES

Favorites

Trending

LLMs

Snowflake powered

Data visualization

Geography & society

Sports & fun

Science & technology

NLP & language



Streamlit cheat sheet

daniellewisd1

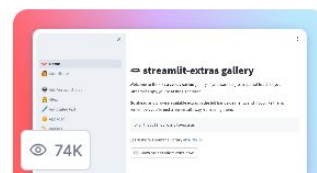
[View source →](#)



Roadmap

streamlit

[View source →](#)



Streamlit extras

arnaudmiribel

[View source →](#)

- コミュニティのパワーでサンプルアプリが豊富
- Githubと連携することで、プロトタイプ公開も可能
 - パブリックなアプリ
 - 無制限
 - ※Github RepoもPublicに
 - クローズなアプリ
 - 1アプリまで
 - メアドで閲覧可能な範囲を制限
 - 各アプリにつき1GBメモリ

Snowflake x Streamlit

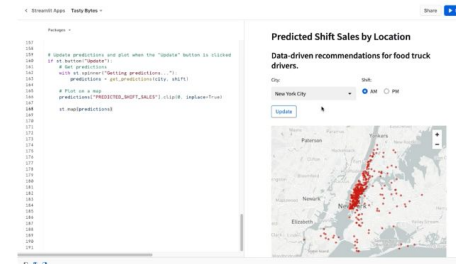


A FASTER WAY TO BUILD, DEPLOY AND SHARE DATA APPS

Data scientists and Python developers can now combine Streamlit's component-rich, open-source Python library with the scale, performance and security of the Snowflake platform.

*Streamlit is now featured in public preview.

[Build](#) [Iterate](#) [Share](#)



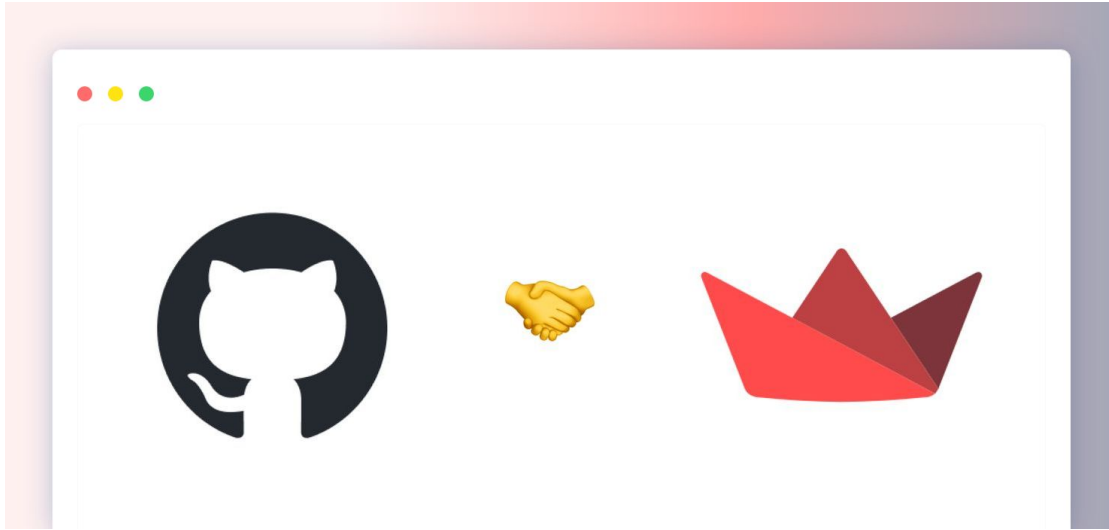
TURN PYTHON SCRIPTS INTO WEB APPS

Define widgets—filters, graphs, sliders and more—as variables to interact with your data and models

[TRY NOW](#)

22年3月頃、SnowflakeがStreamlitを\$800Mで買収。
将来的にはSnowflakeのコンソールから簡単にStreamlitのデプロイが可能になりそう。
(現在、パブリックプレビュー)

Github Codespaces x Streamlit Community Cloud



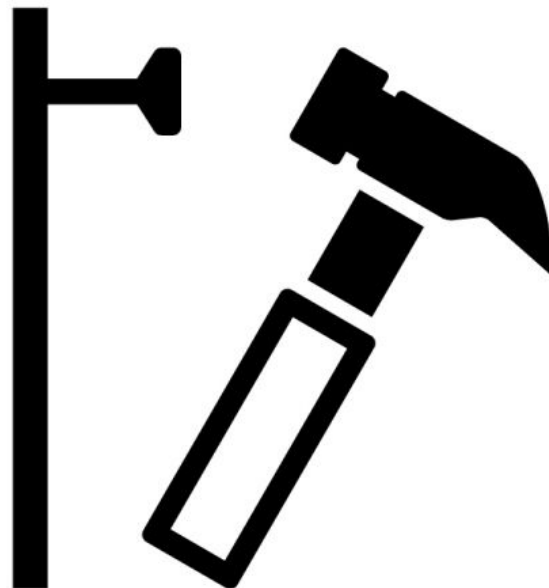
[Develop Streamlit apps in-browser with GitHub Codespaces \(2023/9/14\)](#)

- Github Codespaces と連携できるようになり、コードエディタも不要でブラウザですべて完結
- Streamlitによるプロトタイピングが自分に合っているかチェックする際おすすめ
- 「Streamlitで行く！」と決めたなら、Google CloudやAWSなどのパブリッククラウドで自前で構築するのがオススメ

プロトタイプから本開発に向けて

注意したいこと

StreamlitやPythonはすぐれたツールで、
慣れたんだ本開発も全部をつくりたくなるが。。。



Streamlitで向いてないこと

- ビッグデータの処理や可視化
 - ビッグデータのETL処理は、BigQueryなどの列指向DBなど任せる
 - ビッグデータの可視化は、TableauやLooker Studioに任せる
- 不特定多数が利用するプロダクト
 - ユーザ管理や認証機能は甘い
- 複雑で非同期な画面レンダリング

【参考】プロダクトの本開発でよく利用するツール

フロント



バックエンド



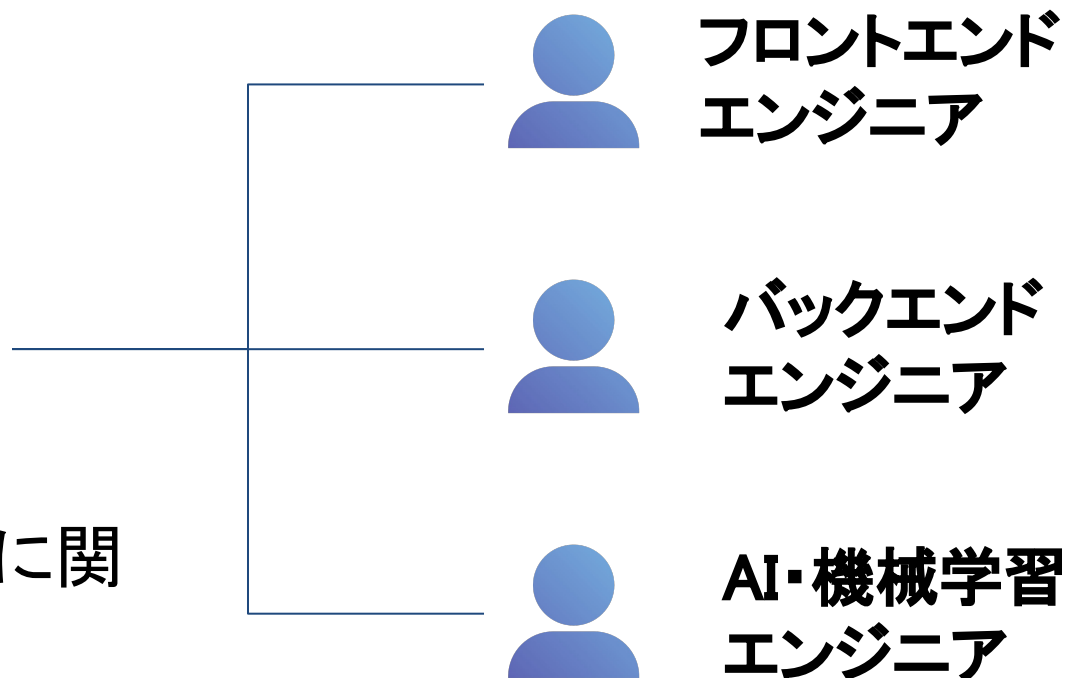
大規模データの非同期処理などが多いので、
GRIでは上記のような構成になりがち

プロダクトオーナーとして、プロダクトの魂をメンバーに伝える



プロダクト開発の進め方に関して責任を持つ

メンバーを熱中させるために、
自分が一番熱中する



⋮

【参考】ForecastFlowのシステム構成

データサイエンスの専門企業が
Google Cloud を使う理由



- 2021年頃のForecastFlowのシステム構成
- Google Cloud Dayでの登壇動画

さいごに

本日の内容

- エンジニアがプロダクトをつくる時のPMFの考え方を、GRIでの取り組みと共に紹介
- Streamlitを利用したプロトタイピングのイメージや本開発に移行する際の注意点

もっとGRIの取り組みを知りたくなった方



GRI HONDA

ここでしか聞けない
リアル限定イベント

自動車業界のトップランナーに聞く
データサイエンスでつくる
誰もが「安心」して移動できるモビリティの未来とは？

2023.10.18.wed
OPEN 18:30 / START 18:40
📍 東京都 港区

参加を申し込む



Saito Daisuke

齋藤 大介 氏

本田技研工業株式会社

電動事業開発本部
BEV開発センター
ソフトウェアデファインドモビリティ
開発統括部



Ohtomo Yuichi

大友 祐一

株式会社GRI

取締役
機械学習システムの社会実装を専門とし、AutoML基盤やMLOps基盤の開発および統括を行う。理化学研究所を経て、2015年より現職。

モビリティのセンシング
データ解析技術で
ワクワクする未来を創りたい



このプロジェクトでも
Streamlit利用



いっしょに働きたいメンバーも募集中

世界中を飛び回るデータにのって縦横無尽に活躍したい人、募集中



<https://gri.jp/recruit>