

# ピックアップ流データ活用基盤の これまでとこれから

PIXIV TECH SALON



pixiv.inc  
minamitary

pixiv

今日話すこと

# データ活用基盤

# 自己紹介

@minamitary / 森田洋介

2015年中途入社

pixiv運営本部 + データ駆動推進室

エンジニア + UX リサーチャー

領域横断的な動き方をすることが多い

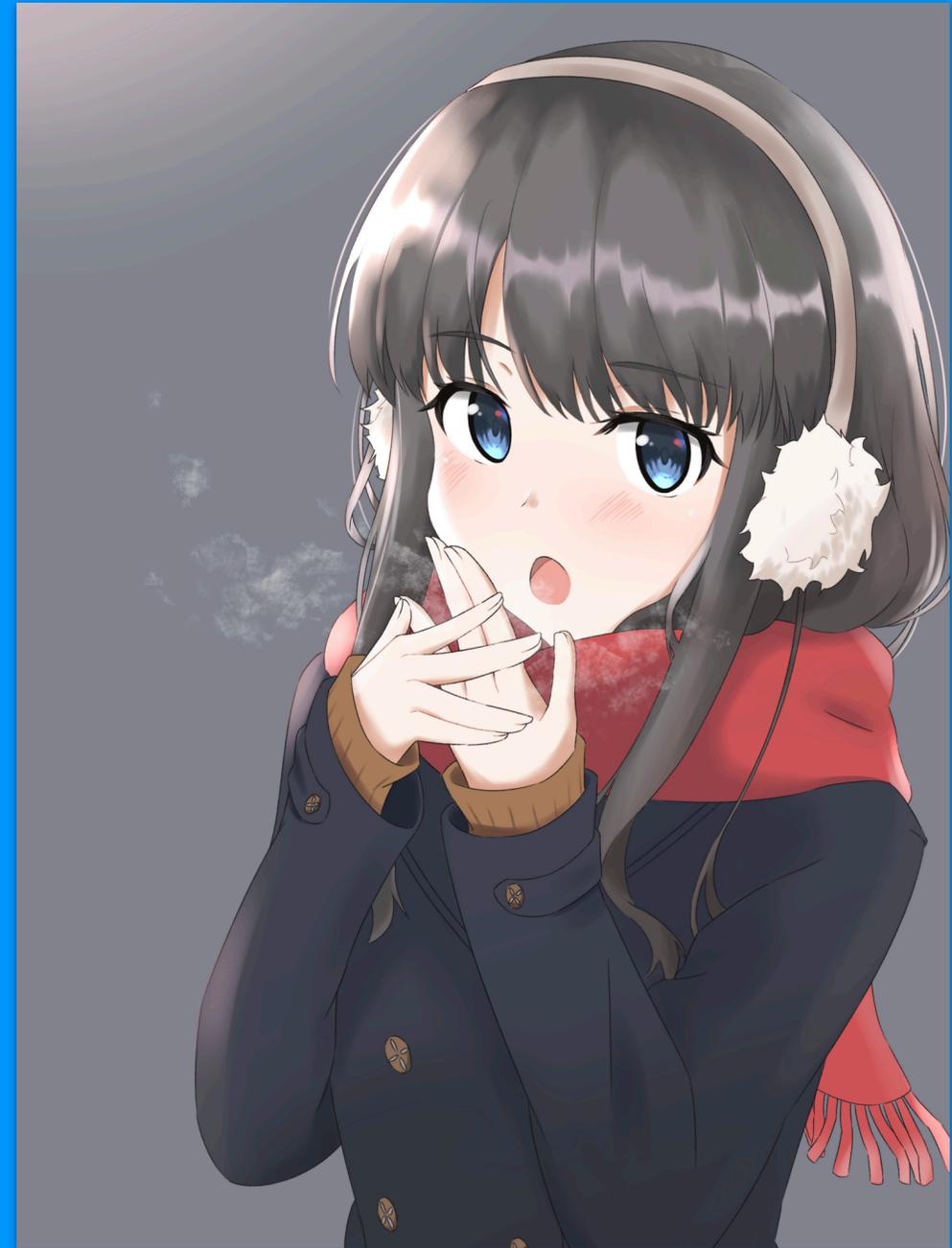
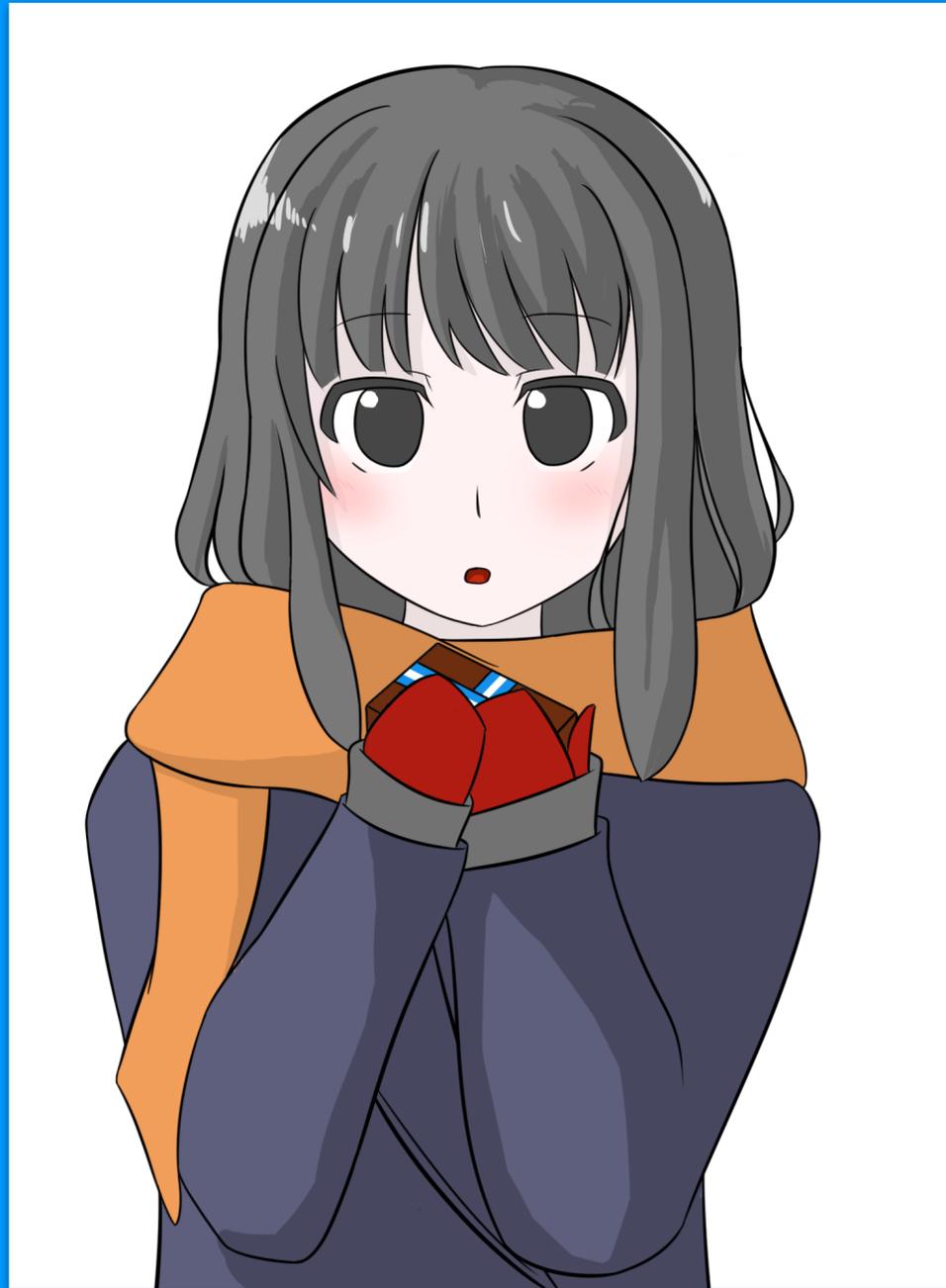


# 自己紹介

- 2015年：入社後「sensei」の立ち上げにリードエンジニアとして参加
- 2016年：senseiのリリース後グロースを担当、その後「pixiv」へ
- 以降クリエイターを中心としたUXリサーチを主に担当
- 定量的リサーチでBigQueryを叩き始めて約3年



# お絵かき歴3年



今日話すこと

# データ活用基盤

# ピックアップにおけるデータ活用基盤

≡ BigQuery



# ピックアップにおけるデータ活用基盤

- BigQueryにデータを集めただけではダメ
- **使われなければデータ「活用」基盤ではない**

「活用」にどう立ち向かうか

# ピクシブにおけるデータ活用基盤

- 「活用」を巡って、3年の間に色々な変化があった
- 技術的な変化だけでなく、組織的な変化も大きかった

# 今日話すこと

- この3年でデータ活用基盤に生じた変化
  - ▶ 技術的な変化
  - ▶ 組織的な変化
- ピクシブにおけるデータ活用基盤のこれから

「オーナーシップ」

# オーナーシップ？

担当するプロダクトに対するオーナーシップ

# オーナーシツプ？

「俺がこのプロダクトを世界一にするんや！」

# 前提：ピクシブの企業文化

- 全職種がユーザーを中心にみて動く
- ユーザー最優先 = 職種間の垣根が低い
- プロダクト愛・カルチャー愛が強いメンバーが多い

「オーナーシップが強い」

というのがピクシブにおける大前提

# オーナーシップはなぜ重要？

ドメイン理解がデータ活用の精度を高める



ドメイン理解の深い人がデータを活用していくべき

# オーナーシップはなぜ重要？

オーナーシップがドメイン理解を深める



つまりオーナーシップがデータ活用の精度を高める



オーナーシップの強い開発者がデータを活用していくべき

基本原則：

データ活用は  
オーナーシップとセット

# もう一つのピクシブの特徴

# 多種多様なプロダクト

pixiv

PIXIVコミック

PIXIVJAIL

ピクシブ文芸

PIXIV FANBOX

BOOTH

pixiv FACTORY

pixiv PAY

SKETCH

sensei

PixiVision

ピクシブ百科事典

Pawoo

Pawoo Music

pixiv chatstory

ImageFlux  
Powered by pixiv

draur

MRoid

MRoidHub

Palcy  
パルシィ

# ピクシブの抱えるプロダクト

- ドメイン領域はすべて「**創作活動**」
- サービス形態は異なりつつも、ユーザー層は同じ
- ユーザー単位での結びつけが大事！

# それゆえに.....

- 「データ活用はオーナーシップとセット」は維持
- データ活用の「橋渡し」も必要

二つのポイント：

「オーナーシップ」

「橋渡し」

1. この3年で生じた変化 ～技術編～
2. この3年で生じた変化 ～組織編～
3. これからの話

時は遡ること3年.....

# 技術編 -3年前-

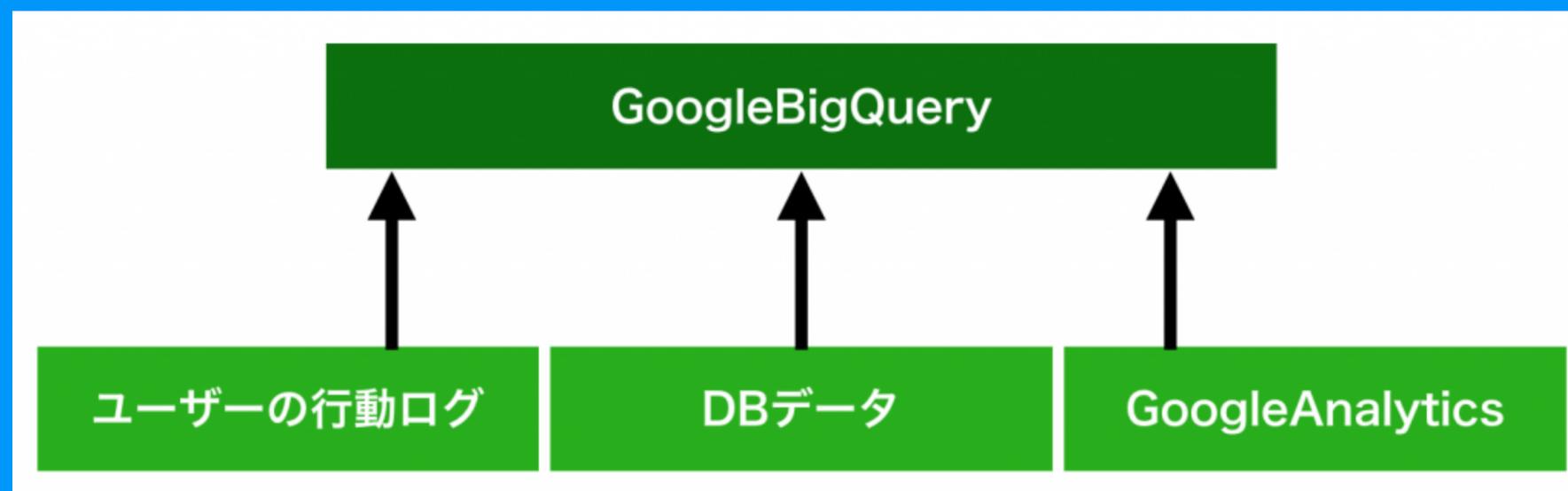
- データは既にBigQueryに乗ってた
- 利用開始は2014年！

# 「活用」は限定的.....

- DBレプリカを都度叩く
- Google AnalyticsのAPIを叩いて集計
- 秘伝のスプレッドシートの乱立
- データも部分的にしか上がっていない！

3年後

# データはBigQueryに集約



JOINし放題！

# 分析手段もBigQueryに統一

- 秘伝のスプレッドシート：ほぼ消滅！
- 秘伝の集計スクリプト：ほぼ消滅！
- 管理画面上に点在していたグラフ群：ほぼ消滅！
- 全てのデータ処理がSQLで記述される世界に

# 全てがSQLで記述される世界

- WITH句やビューで前処理を簡略化
- 使い回しやすいWITH句を資産として保存・シェア
- SQLの共通資産化！

# 全てがSQLで記述される世界

- SQLさえ覚えればデータを活用できる世界
- BigQuery上でデータ処理が完結
- マシンリソースを意識する必要がない

# 全てがSQLで記述される世界

- 「SQL便利」「SQLイケてる」
- 結果ビジネス側の人間にもSQL利用が拡大
- オーナーシップの恩恵！

「橋渡し」しやすい環境ができた！

# 技術的な変化まとめ

- すべてのデータがBigQueryに集約された
- 分析手段がSQLに統一された
- 「橋渡し」がやりやすくなった

1. この3年で生じた変化 ～技術編～
2. この3年で生じた変化 ～組織編～
3. これからの話

再び遡ること3年.....

# 組織編 -3年前-

- 各部署がそれぞれで頑張る
- ビジネス系の部署を中心にデータが活用されてはいた
- ダッシュボード、管理画面、スプレッドシートの乱立

# 背景：増えゆくプロダクト

2016/01

2019/01

9<sub>個</sub>



20<sub>個</sub>

この3年間でプロダクトの数は2倍以上に

# 背景：増えゆくプロダクト

- それに応じてチームの数も増加
- チームにデータ得意マンがない
- 片手間データエンジニアリング

分析専門の部署を作るべき？

# 答えはNO！

- 基本原則「データ活用はオーナーシップとセットで」
- 「データはあの部署の人が見てくれる」という意識は避けたい

分析担当者は各部署内にいるべき

とはいえ横の繋がりは必要.....

# 対策

# 対策

Slackチャンネルを作っただけ

# 経緯

- 社内勉強会用のSlackチャンネルを作る
- 勉強会終了後、主催メンバーが結託
- 「データに興味ある人はここに来てくれ！」 アピール

# 経緯

- チャンネルに人が集まりはじめる
- 多数の相談が寄せられる
- 他部署の詳しい人がそれに応じる文化
- 「**ここに行けば解決する**」 さらに人が増える！

# 経緯

- 週次で分析担当が集まるように
  - ▶ 新規データセット追加の連絡
  - ▶ 新しいダッシュボードの共有
  - ▶ BigQueryの新機能紹介
  - ▶ 長期的な方針決め
  - ▶ などなど



専門部署駆動ではない  
自発的な「活用」 基盤の改善

**Why?**

# Why?

- 「**オーナーシップ**」というピクシブの企業文化ありき
- 「プロダクトを成長させる上で、必要だからやる」

# Why?

- オーナーシップ + **ベストプラクティスの同期**
- 「これがベストプラクティス」が同期されることで.....
  - 「ちょっと手が空き気味なのでガッと改善しておきます」
  - 「更にラクにできる方法を思いつきました」
  - 「このBigQueryの新機能は便利そうなので使ってみましょう」

# Why?

- 「**オーナーシップ**」の発揮先を同期する必要性
- そのための、Slack等における「**橋渡し**」の設計

# 組織的な変化まとめ

- Slackチャンネル作成が変化のきっかけ
- チャンネルに人が集まり、横の連携が強化
- 「橋渡し」によるベストプラクティスの同期
- 「オーナーシップ」ベースで横断的な基盤改善が可能に！

1. この3年で生じた変化 ～技術編～
2. この3年で生じた変化 ～組織編～
3. これからの話

その前に

# 振り返ってみると.....

- 技術的な変化だけを進めていたら.....？
- 組織的な変化だけを進めていたら.....？
- **どちらも「活用」は成り立たない！**

# 振り返ってみると...



基本原則：

データ活用は  
オーナーシップとセット

もう一つの基本原則：

技術的変化は  
組織的変化とセット

1. この3年で生じた変化 ～技術編～
2. この3年で生じた変化 ～組織編～
3. これからの話

# 昨年末



CTO

部署作っちゃお！

昨年末

「データ駆動推進室」の誕生

# データ駆動「推進」室

- 基本原則「データ活用はオーナーシップとセットで」
- 分析の実務は担当しない！
- 各部署のデータ活用をサポート・推進する役割
- ビジネス面・UX面・技術面すべてが対象！



データ活用を  
サポート  
→

サービスを 担当するチーム	サービスを 担当するチーム	サービスを 担当するチーム
デザイナー	デザイナー	デザイナー
ディレクター	ディレクター	ディレクター
エンジニア	エンジニア	エンジニア
エンジニア 兼 分析担当	ディレクター 兼 分析担当	デザイナー 兼 分析担当
● ● ● ●	● ● ● ●	● ● ● ●

よりアクロバティックに  
「動ける」  
ように部署化

# 具体的なアプローチ

ユーザーに価値を最速で届けるために。  
ピクシブの「データ民主化」に向けた挑戦

<https://inside.pixiv.blog/jaggy/6421>

# 具体的なアプローチ

- 技術的なトピックもたくさん
  - ▶ BigQuery上で大規模なデータを扱う際のノウハウ
  - ▶ BigQuery周辺で利用しているツール群・具体的な実装内容
  - ▶ ピクシブならではのデータ活用事例あれこれ
  - ▶ などなど.....
- 詳しく聞きたい方は懇親会で！

# まとめ

- データ活用はオーナーシップとセット
- 技術的变化は組織的变化とセット
- データ駆動推進室ができた

社員もプロダクトも増え続けるけど、

「創作活動がもっと楽しくなる場所を創る」気持ちは変わらない。

ピクシブはデータの手で創作活動を支えていく。