

USAGE-BASED BILLING



GitHub 最新情報ウェビナー

~ Usage Based Billing ~



WILLIAM ZHANG

AI Go-To-Market Lead | GitHub Copilot

GitHub

Agenda Session

- 01 Usage-Based Billing (従量課金制) への移行
- 02 料金プランの紹介
- 03 AI クレジット共有プール
- 04 Budgetの予算管理



Usage Based Billing (従量課金制)

6月1日

⇔ 変更内容

6月1日より、プレミアムリクエストユニット (PRU) は GitHub AI クレジットに置き換わります。

クレジットは、入力・出力・キャッシュされたトークンを含むトークン使用量に基づき、各モデルの公開 API レートに従って消費されます。

🕒 タイムライン

- 4月27日 – 発表
- 5月 – 請求プレビュー提供開始
- 6月1日 – 従量課金制開始

なぜ従量課金制に移行するのか

6月1日

- ✔ AIコンピューティングの価格はトークンベースになります。
- ✔ リクエストベースの価格とAIコンピューティングコストの乖離
- ✔ Usage Based Billingはモデルや価格の変化に適応するよう設計されています。
- ✔ 従量課金により、従来の価格モデルでは提供できなかった機能（例：より大きなコンテキストサイズ）を提供できるようになります。
- ✔ 実際の使用状況をより正確に反映するために進化しています。

従量課金制の仕組み

GitHub Copilotに高度なAI機能の提供が増えたことに伴い、従量課金制はAIの使用量とコストを結びつけます

- AIの使用量は、各インタラクション中に処理されたトークン数で測定されます
- トークンにはプロンプト、レスポンス、キャッシュされたコンテキストが含まれます
- 付帯する使用量: GitHub Copilotライセンス間で共有されます
- 超過使用量: 予算とコントロールで管理可能です
 - 超過使用量の予算は米ドルで設定し、使用量はGitHub AIクレジットで表示されます。
 - GitHub AIクレジットは固定レート（1 AIクレジット = \$0.01 USD）で予算から消費されます。例えば\$10の予算は1,000 AIクレジット分に相当します。

トークンの説明

トークンはAI処理の単位です(入力、出力、コンテキスト)



入力トークン:

送信するもの:

- プロンプトと新しいコンテキスト
- 大きなファイルやコードベースに応じて増加



出力トークン:

受け取るもの:

- AIが生成する応答
- 通常、最もコンピューティングコストが高い



キャッシュトークン:

再利用されるもの:

- 過去のインタラクションからのコンテキスト
- 速度と効率を向上

📌 **ポイント:** すべてのインタラクションは、入力、出力、コンテキストに基づいてトークンを消費します。



従量課金制の課金 (UBB)

課金対象トークン = 入力 + 出力 + キャッシュトークン



課金の仕組み:

- Copilotの利用量は、各インタラクションで処理されたトークンの合計で測定されます。
- トークンはAIクレジットを消費します([モデルレート](#)に基づく)
- AIクレジットが課金の基準となります

📌 **ポイント:** より大きなプロンプト、より長い応答、より多くのコンテキストは、処理されるトークンとクレジット消費の増加につながります。



使用量はタスクや ワークフローによ って変動します

モデル

モデルごとに異なるモデルレートカードがあります。

プロンプトサイズ

コンテキストが増えるとトークン使用量が増加します。

レスポンスサイズ

出力が長くなるほどトークン使用量が増加します。

コンテキスト キャッシュ

コンテキストの再利用で効率が向上します。

エージェントワークフロー

複数ステップのワークフローはより多くのトークンを消費する可能性があります。

MCP / スキル

スキルやMCPサーバー経由で接続されるツール



料金モデルの概要

GitHub Copilot 各プランと含まれる価値

Copilot Pro	Copilot Pro+
月額\$10 / ユーザー (変更なし)	月額\$39 / ユーザー (変更なし)
含まれる \$10 の トークン価値 (1,000 AIクレジット)	含まれる \$39 の トークン価値 (3,900 AIクレジット)

仕組み

- 処理されたトークンに基づいて消費
- 年次プランのユーザーはプランの有効期限まで
現行のプレミアムリクエストベースの料金体系が維持

Copilot Business	Copilot Enterprise
月額\$19 / ユーザー (変更なし)	月額\$39 / ユーザー (変更なし)
含まれる \$19 の トークン価値 (1,900 AIクレジット)	含まれる \$39 の トークン価値 (3,900 AIクレジット)

Copilotプランに含まれるAIクレジットはエンタープライズ全体で共有され、トークンの処理に応じて消費されます。

仕組み

- 処理されたトークンに基づいて消費
- クレジットはエンタープライズレベルで共有
- 個々のユーザーに紐付かない

📌 **ポイント**：顧客は実際の利用量に応じて支払います—ユーザー単位ではなく—エンタープライズ全体で共有されるクレジット



従量課金制 プランの紹介(プロモーションオファリング)

プランと含まれる価値

Copilot Business	Copilot Enterprise
月額\$19 / ユーザー	月額\$39 / ユーザー
含まれるトークン価値 \$19 (AIクレジット1,900)	含まれるトークン価値 \$39 (AIクレジット3,900)
共有クレジットプール	共有クレジットプール

プロモーションオファリング

2026年6月、7月、8月分の利用へ適用

Copilot Business	Copilot Enterprise
月額\$19 / ユーザー	月額\$39 / ユーザー
含まれるトークン価値 \$30 (AIクレジット3,000)	含まれるトークン価値 \$70 (AIクレジット7,000)
共有クレジットプール	共有クレジットプール



注意点

無償モデル

現在、PRUを使い切ったユーザーはコスト0のモデルにフォールバックして作業を継続できます。

新しい課金モデルでは全てのモデルにモデルレートが設定されています。

コード補完機能

コード補完機能とNext Edit Suggestions（次の編集候補）は従来通り無償で提供されます。

Auto modeモデルの 機能改善

6月1日

- **Auto mode**は自動的に適切なモデルへのルーティングを行います
- シンプルなタスクは**軽量モデル**
- 複雑なタスクは**高性能モデル**
- **不必要な**高性能モデルの使用を削減します。
- 複数のモデルプロバイダーが提供する**各種モデル**へ動的にアクセスが可能です。

 デフォルトでコストを最適化

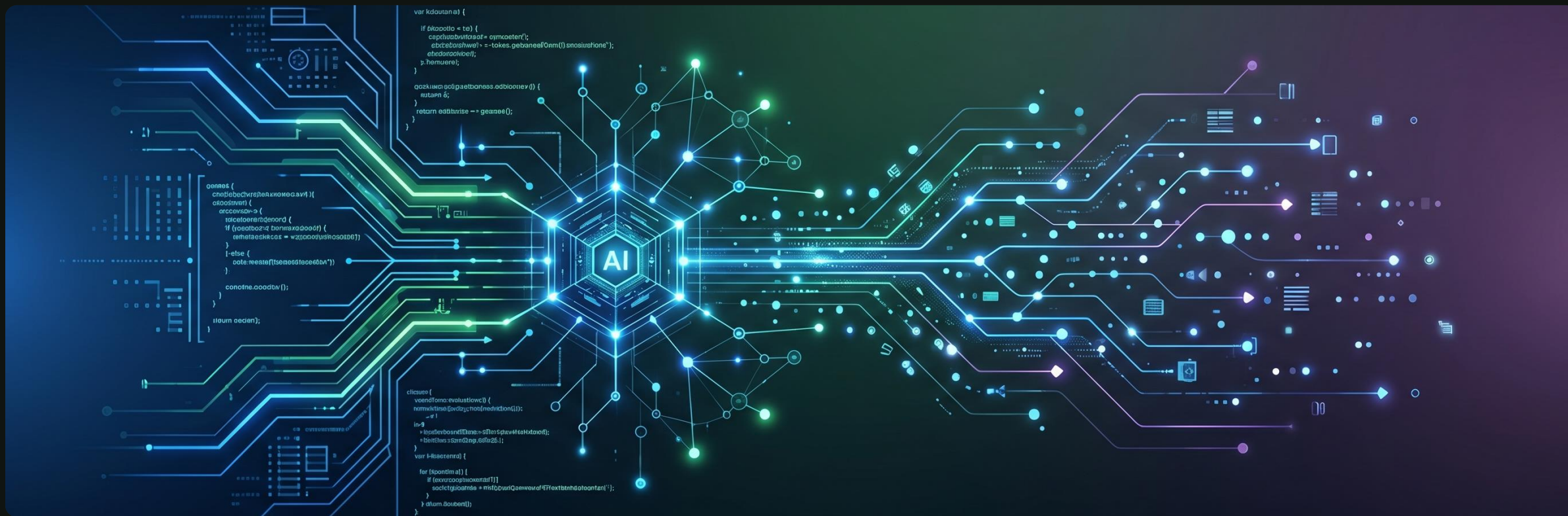
従量課金制 vs プレミアムリクエスト

項目	PRU	UBB
計測単位	[各 プレミアムプロンプトまたはセッション] × [モデル係数]	インタラクションで処理されたトークンの合計 (AIクレジット換算)
コストの要因	使用した機能 + 選択したモデル	選択したモデル、プロンプトサイズ、コンテキスト長、レスポンス長、キャッシュトークン
予算管理 / プーリング	ユーザーごとの月間エンタイトルメント	共有プール
コストフレームワーク	使用量はリクエスト数に応じてスケール	使用量は実際のAIワークロードに応じてスケール

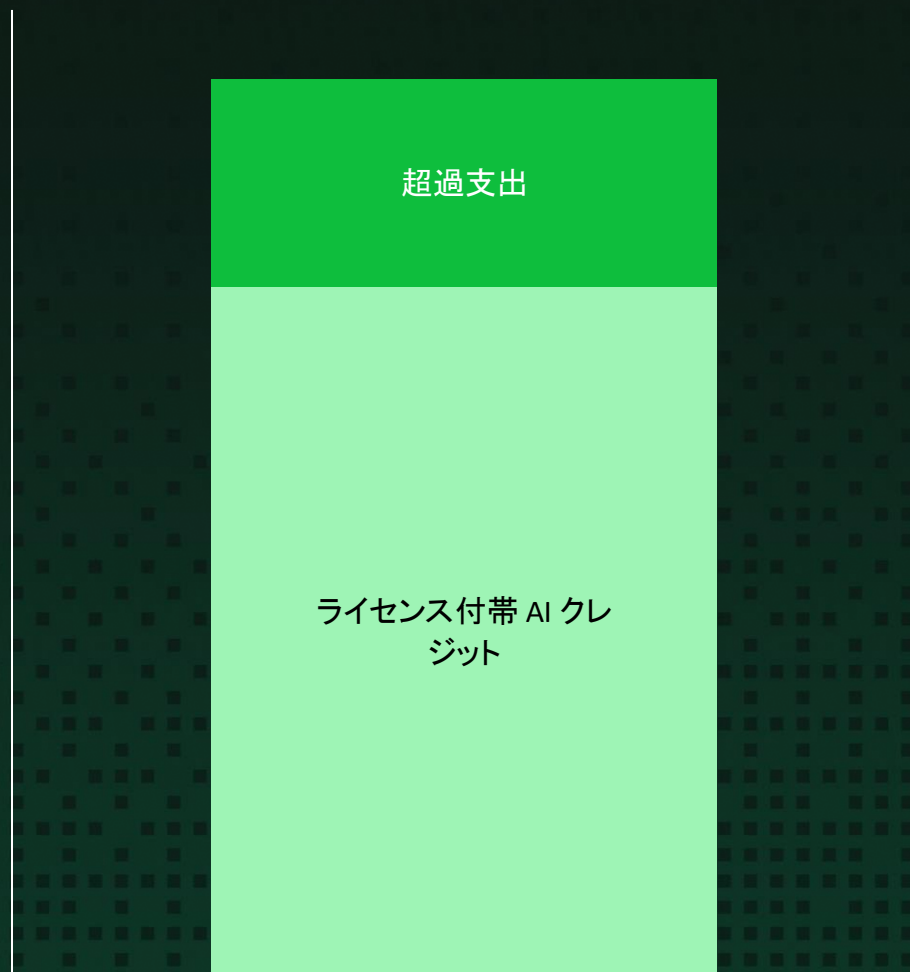
📌 **主な違い:** PRUはCopilotの使用頻度を測定しますが、UBBは各インタラクションでどれだけのAI処理が行われるかを測定します。



AI クレジット共有プール



■ ライセンス付帯 AI クレジット ■ 超過支出



総 AI クレジット

AI クレジット共有プール

GitHub Copilot ライセンスに付帯するAIクレジットは**単一エンタープライズの共有プール**に合算されます。

ライセンスを持つすべてのユーザーは、まずこのプールから消費します。

プールのクレジットがすべて使い切られた後に発生するAIクレジットの使用分は、**追加費用**となります。

AIクレジット 共有プール

エンタープライズ内での利用分布



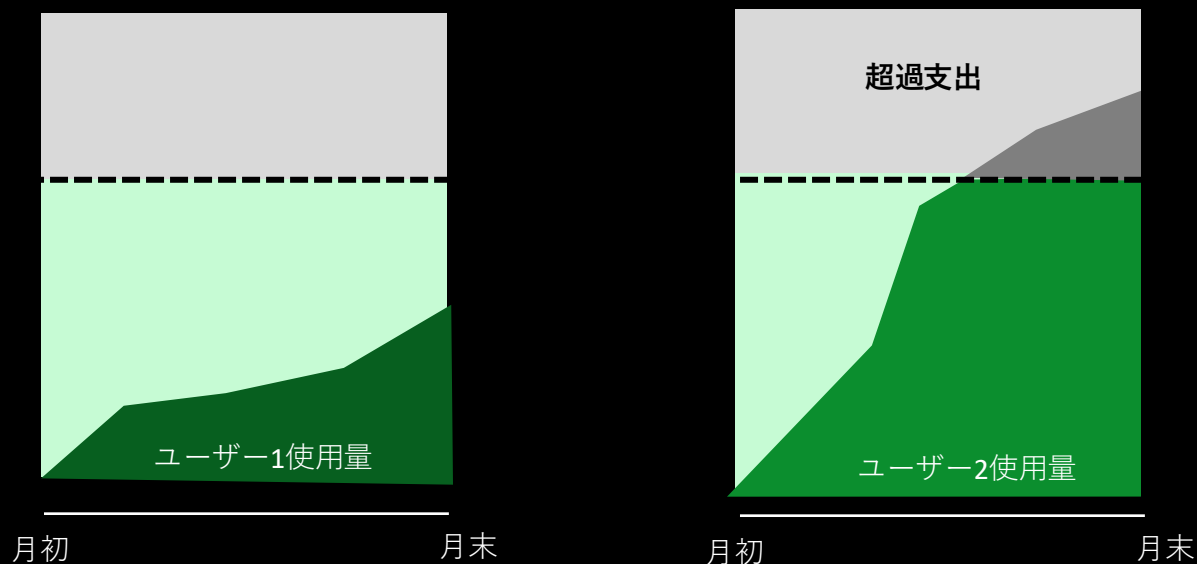
✓ 全ユーザーが同じプールから利用—ただし利用率は異なる



AI クレジット 共有プール

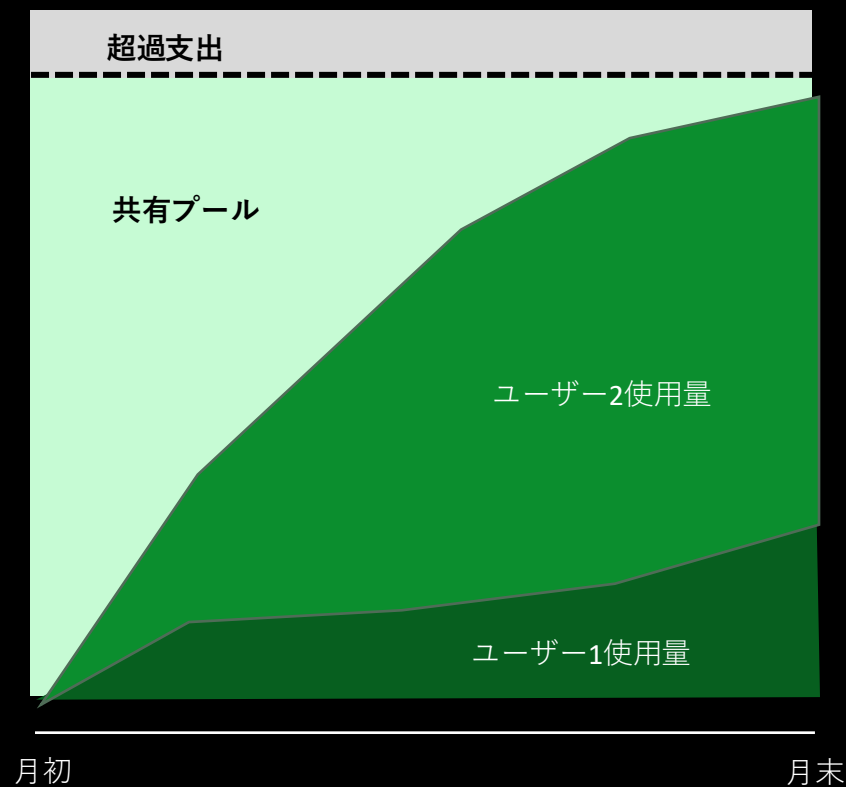
PRU モデル

メリット：あるユーザーが他ユーザーの使用量を使い切ることがない
デメリット：他ユーザーにまだ残量があっても、超過使用が発生する可能性

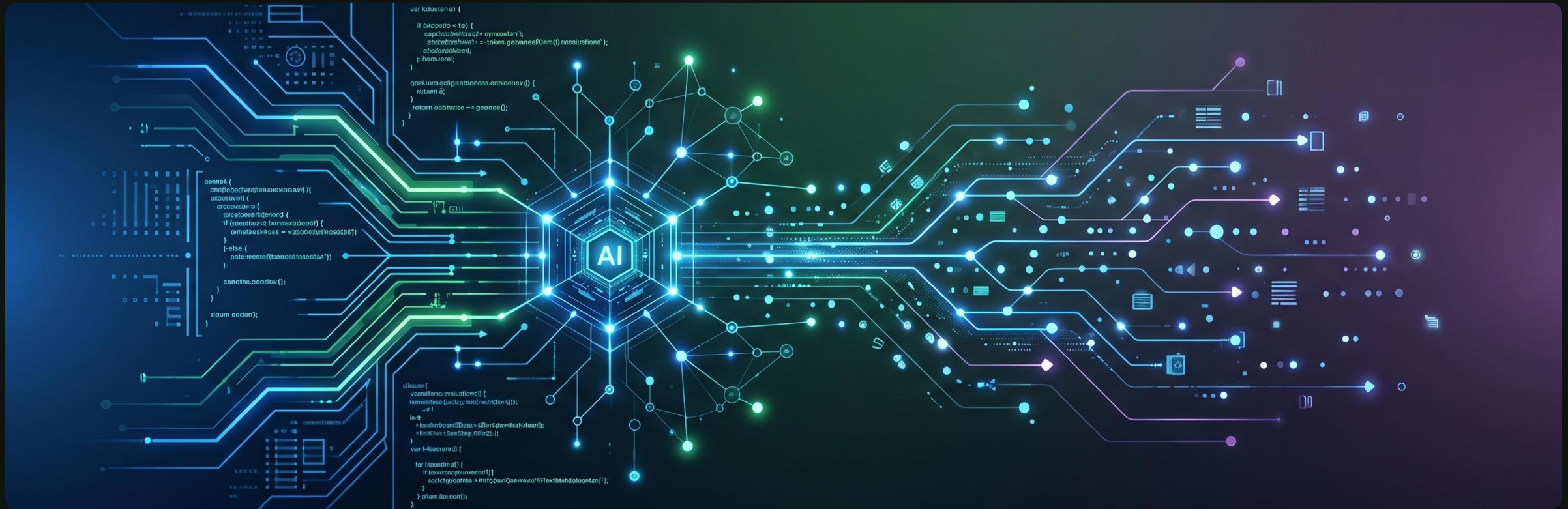


AI クレジットモデル

メリット：無駄が少ない
デメリット：ユーザーごとに消費量が偏る可能性



Budgetの管理



ユーザーレベル Budget (new)

請求期間中にユーザーが消費できるAIクレジットの合計数（共有プール使用分と超過支出の両方を含む）を管理します



共通デフォルトユーザーレベルBudget

すべてのユーザーに適用されるデフォルトのユーザーごとのAIクレジット予算

新規ユーザーが基本的な支出権限を持つことで、意図しない過剰支出を防ぎます



個別ユーザーレベルBudget

パワーユーザー上書き設定

共通デフォルトを上書きする、ユーザーごとの個別AIクレジット予算です。

Budgetのオプション

【新】ユーザーレベルBudget

共通デフォルト

個別*

超過支出

超過支出

共有プール

共有プール

*共通デフォルト設定を上書きします

組織レベルBudget

組織の超過支出を制限

超過支出のみ

コストセンターレベルBudget

コストセンターの総超過支出を制限

超過支出のみ

エンタープライズレベルBudget

エンタープライズの総超過支出を制限

超過支出のみ

↑
----- コストセンターの使用量を
除外できます



Budget 階層

重要ルール: いずれかのBudgetが\$0になると使用が停止されます

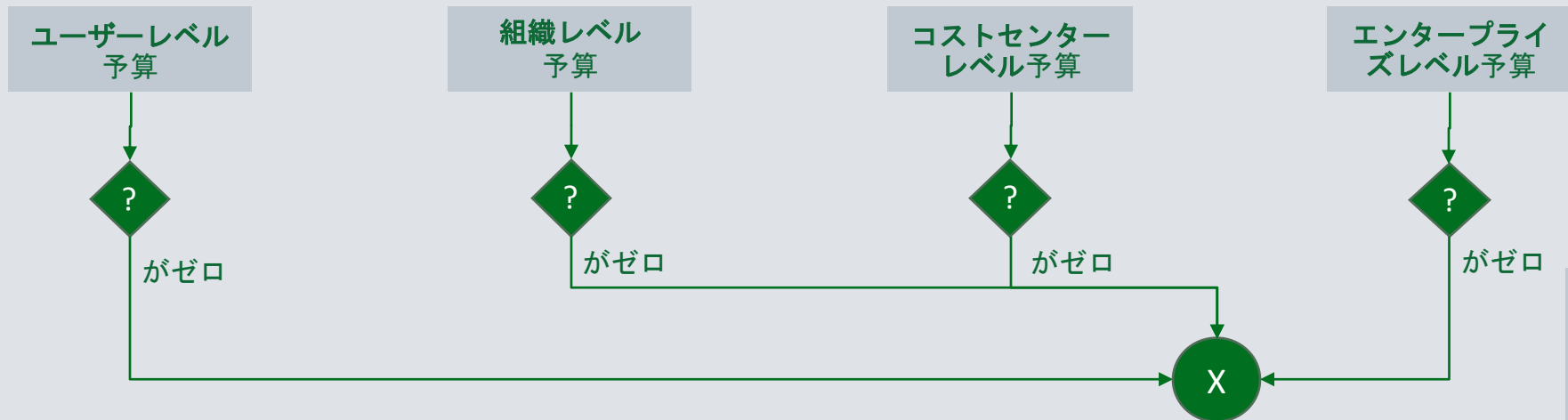
フェーズ1. 共有AIクレジットプール

全ユーザーが共有プールから消費し、ユーザーのULBのみがチェックされます



フェーズ2. 共有AIクレジットプールの枯渇

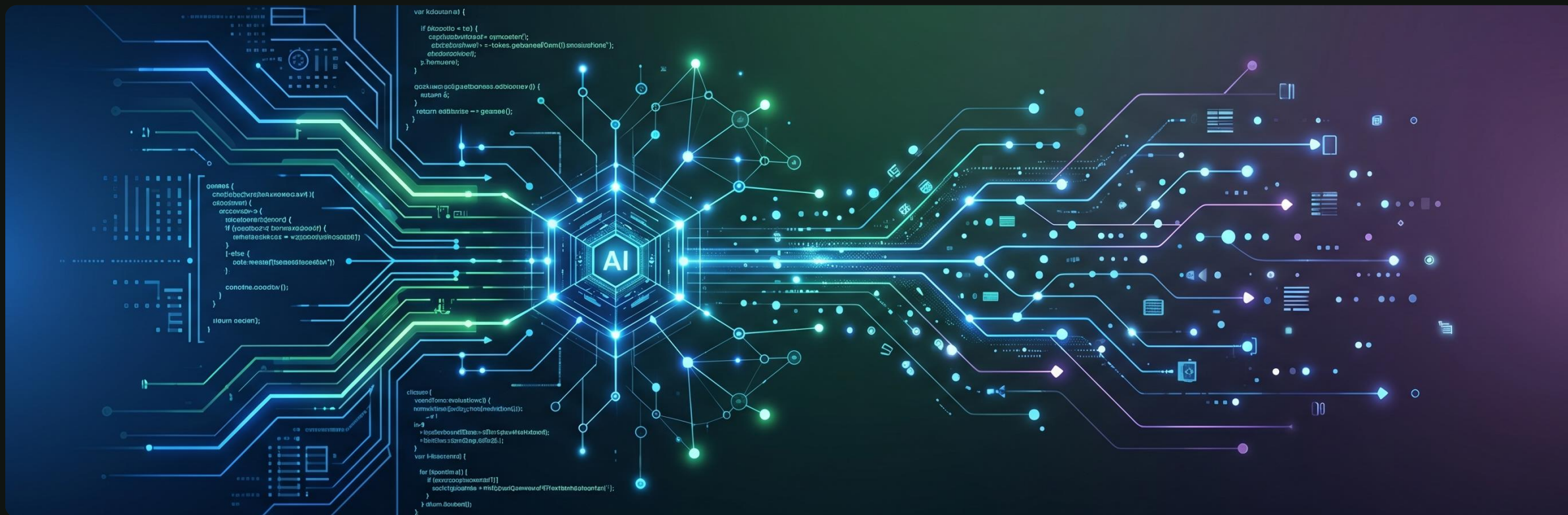
超過支出分については、以下の予算をチェック



オプションポリシー: コストセンターをエンタープライズ予算から除外することで、エンタープライズ予算が\$0になってもコストセンターのユーザーは引き続き使用できます。



Budgetの設定画面



従量課金制移行初 日 (6/1)

- サプライズ請求を防ぐため、予算は自動的にAIクレジットに移行されました
- 請求UIがAIクレジットを示すよう更新されました（モデルおよびユーザーによるフィルター付き）
- ユーザーレベルBudgetは設定されていません

The screenshot shows the GitHub Billing & licensing interface for Avocado Corp. The main section is titled "AI usage" and displays the following information:

- Included usage:** 6,470 / 30,000 AI units. A progress bar shows that approximately 21.6% of the monthly limit has been used.
- Overage spend:** \$0 / \$500 budget. A progress bar shows that no additional credits have been used.
- Usage grouped by model:** A bar chart for the period Jun 1 - Jun 2, 2026, showing AI units used for three models: Claude Sonnet 4.5 (approx. 3,470 units), Gemini 3 Pro (approx. 2,150 units), and Claude Sonnet 4 (approx. 850 units).

Model	Included credits	Included usage	Additional credits	Overage spend
Claude Sonnet 4.5	3470	\$34.7	0	\$0.00
Gemini 3 Pro	2150	\$21.5	0	\$0.00
Claude Sonnet 4	850	\$8.50	0	\$0.00

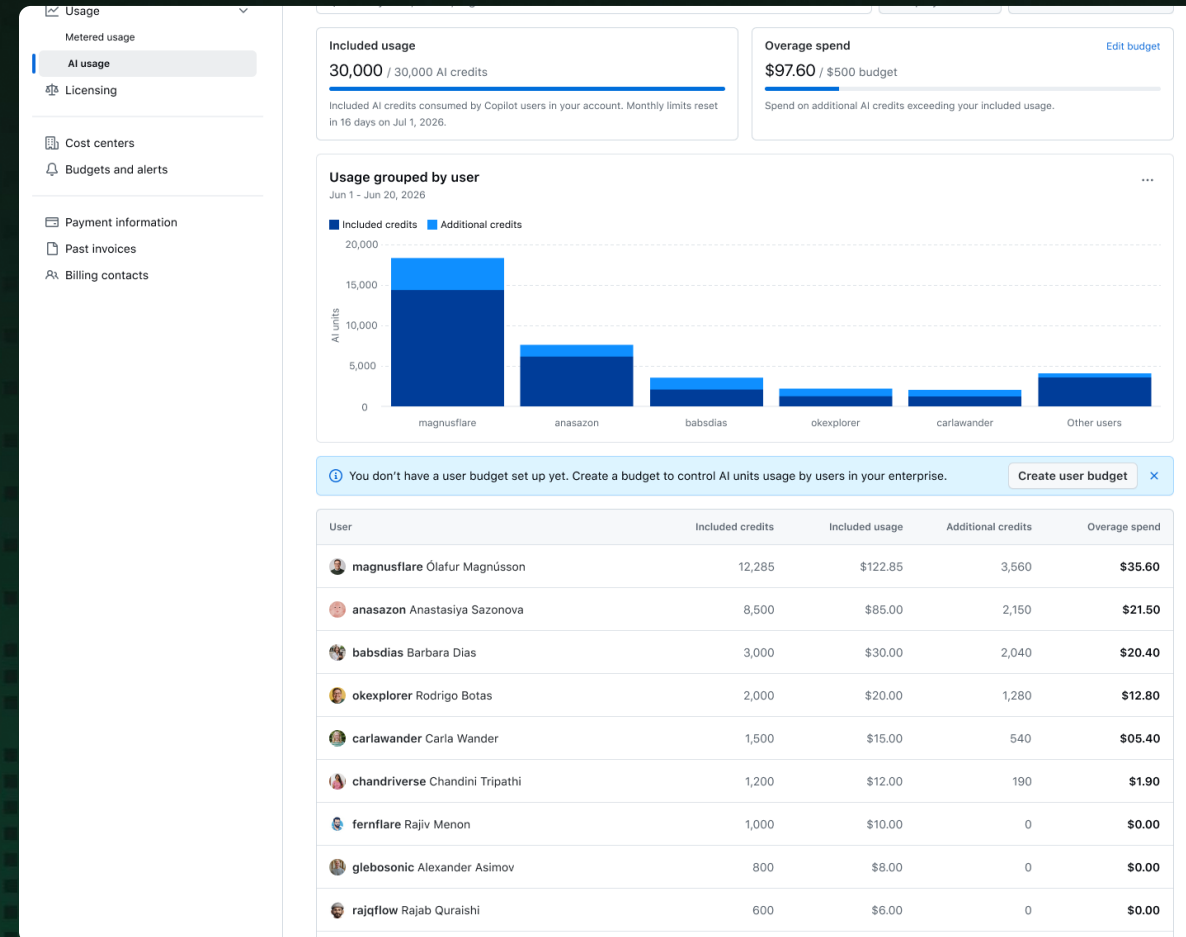
Each GitHub AI unit costs \$0.01.

© 2025 GitHub, Inc. | Terms | Privacy | Security | Status | Docs | Contact GitHub | Pricing | API | Training | Blog | About

設定: ユーザーレベルBudget

01 / 03

- ユーザー別の使用状況を閲覧し、必要に応じてユーザーレベル予算の適用を検討します
- ユーザーレベル予算は、共有プールと超過支出の両方を追跡し、1人のユーザーが共有プールを使い切ったり、予期せぬコスト超過が発生するのを防ぎます。



設定: ユーザーレベルBudget

02 / 03

- 管理者はまず共通デフォルトユーザーレベル予算(ユーザーごとに同一予算)の設定を促されます。

Licensing

- Cost centers
- Budgets and alerts**
- Payment information
- Past invoices
- Billing contacts

1 Budget type
Set budget for either a product or individual SKU within a product.

- Product-level budget
- SKU-level budget
- Bundled AI credits budget**
Set a budget for all SKUs that consume AI credits, where [override spend is enabled](#).
Supported SKUs:
 - Copilot AI credits
 - Coding agent AI credits
 - Spark AI credits

2 Budget scope
Select the scope of spending for this budget.

- Enterprise**
Scoped budgets track spending for all organizations, repositories and cost centers in the enterprise
- Organization**
Scoped budgets track spending for all repositories in the organization
- Cost center**
Scoped budgets track spending for a single cost center
- Users**
Each user in your enterprise gets their own individual budget.

3 Budget amount
Set a budget amount to track your spending on a monthly basis.

ⓘ This is the total usage allowance for a user covering both included usage and overages consumed. [Learn more](#)

Budget amount
\$20 per user
2,000 AI credits (\$0.01/credit)
100 currently assigned user licenses

設定: ユーザーレベルBudget

03 / 03

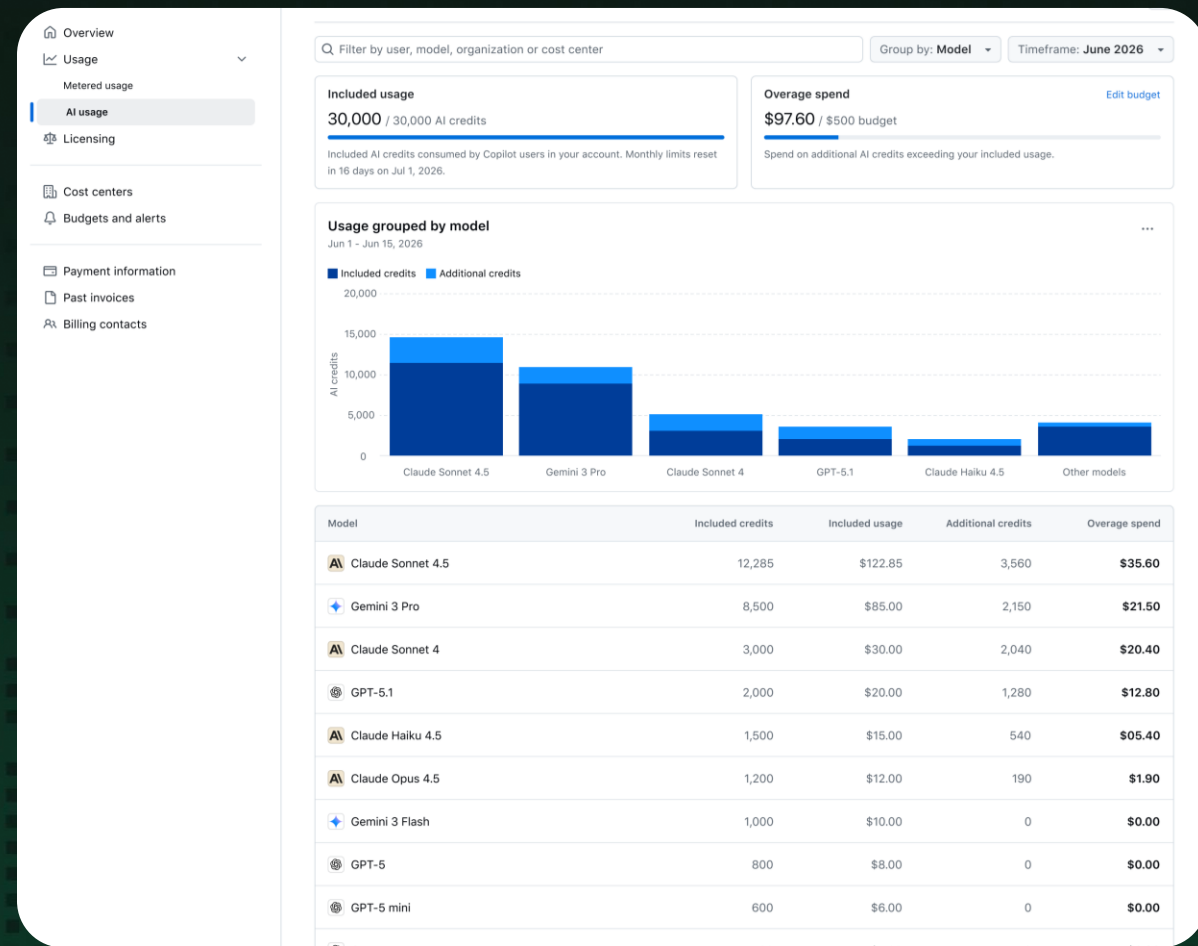
- 管理者はユーザー別の使用状況を確認し、必要に応じて個別に更新できます。
- 請求UIでは、全ユーザーにおける「上限到達」「上限間近」「75%未満」の使用状況の内訳を提供します。

The screenshot displays the 'AI credits user budgets' interface in the Azure portal. At the top, there are navigation tabs for 'People', 'AI Controls', 'Policies', 'GitHub Connect', 'Code Security', 'Billing & licensing', 'Settings', and 'Compliance'. The main content area is titled 'Budgets and alerts / User budgets' and includes a 'Get usage report' and 'Edit budget' button. Below this, a search bar allows filtering by user or budget threshold. A usage summary shows '100 licensed users' with a progress bar indicating usage levels: '< 75% budget used 28', 'Near limit 57', and 'At limit 4'. A table below lists individual users with their budget and usage details.

User	Budget	Used
Sebastian Holt	\$10.00	\$10.00
Owen Harrington	\$10.00	\$10.00
Henry Fletcher	\$10.00	\$10.00
Ethan Caldwell	\$10.00	\$10.00
Jason Nguyen	\$10.00	\$9.54
Diego Fernández	\$10.00	\$9.51
Nora Müller	\$10.00	\$9.51
Grace Choi	\$10.00	\$9.51
Lucas Silva	\$10.00	\$9.51
Amina Khan	\$10.00	\$9.51

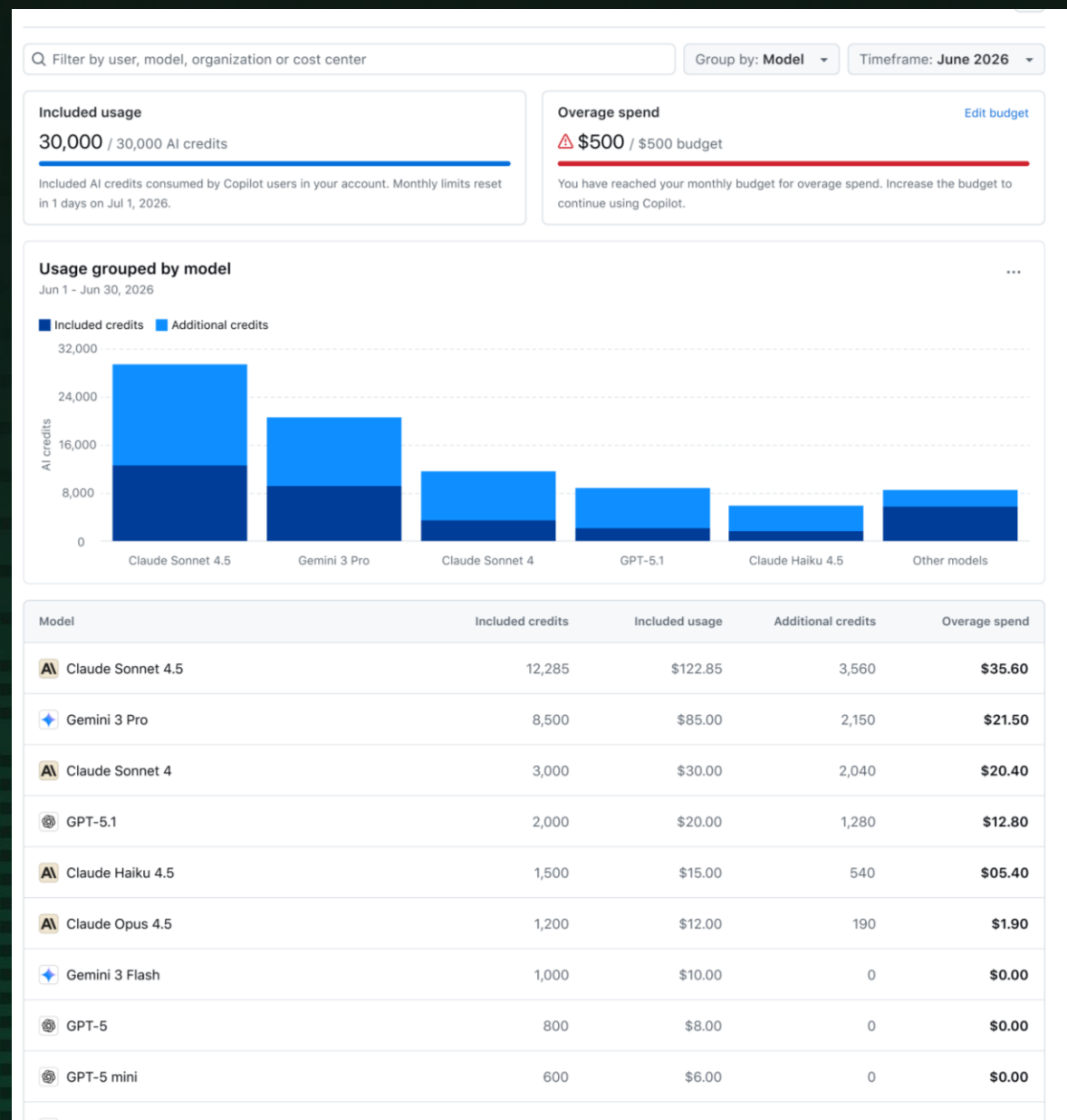
請求サイクル途中

- エンタイトルメントプールが先に消費されます。
- 超過予算は追加の支出を適切に追跡します。
- 請求UIには、含まれる使用量と追加支出が表示されます。



Budget 到達

- エンタープライズ予算に到達。
- 管理者は使用量が75/90/100%のしきい値に達すると警告メールを受信します。
- エンドユーザーは次の請求サイクルまで、または予算が増額されるまでブロックされます。



シンプルな予算設定の流れ

1. 可視化から始める

モデル、ユーザー、
コストセンター別に使用状況を確認

2. エンタープライズの ガードレールを設定

エンタープライズレベルで追加使用を制御

3. コストセンター予 Budgetを追加

使用状況をチームや事業部門に整合

4. 共通デフォルトユーザーレベルBudgetを追加

全ユーザーに適用される共通予算を追加

5. 個別ユーザーレベルBudgetを設定

パワーユーザーや重要なチームに個別の上限を追加

6. 導入後に見直し

導入状況や使用パターンが明確になるにつれて予算を調整



最適な分配

(均等 vs 80/20、など)

既知のパワーユーザー

使用率の高い組織

使用率の高いコストセンター

モデルの好み

ユーザー成熟度

Budget設定時の 考慮事項

以下を参考に、どのBudgetを、どのように組み合わせるかを考えます。

- 全員の公平な使用を確保するために共通のユーザーレベル予算を使いますか?
- 一部のパワーユーザーに高い予算を割り当てますか?
- 特定のチームに追加支出を許可するためにコストセンターを追加しますか?

AIクレジットの可視化

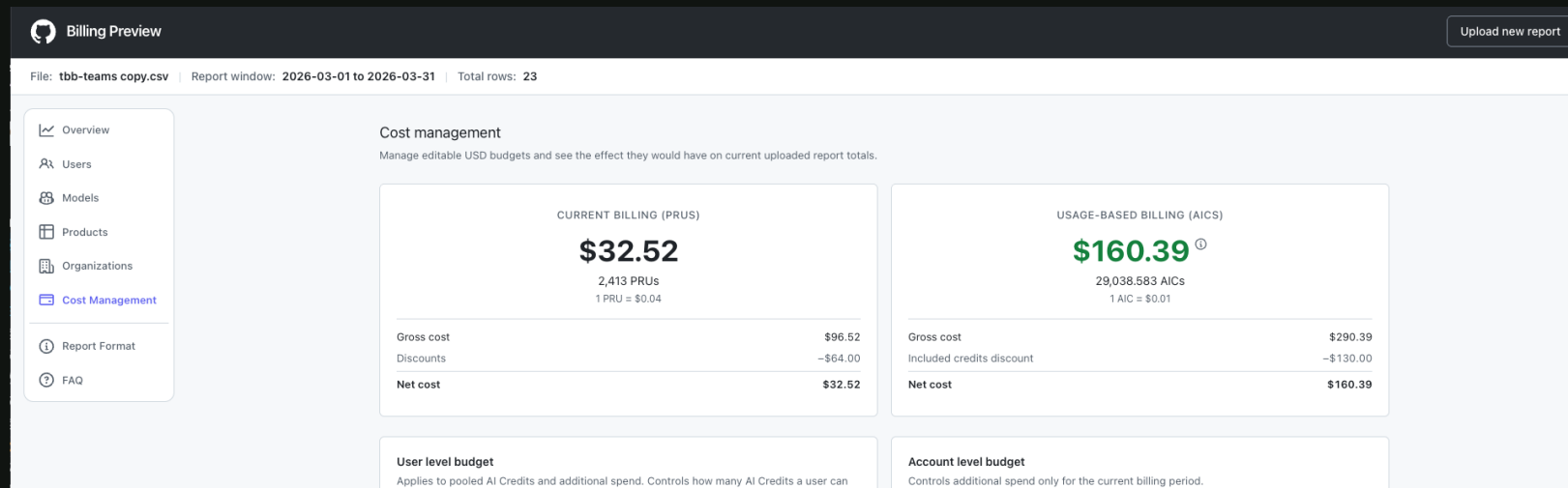
AIクレジット CSV Usage Reportのダウンロード:

Billing and licensing → Preview your billing impact → Download CSV

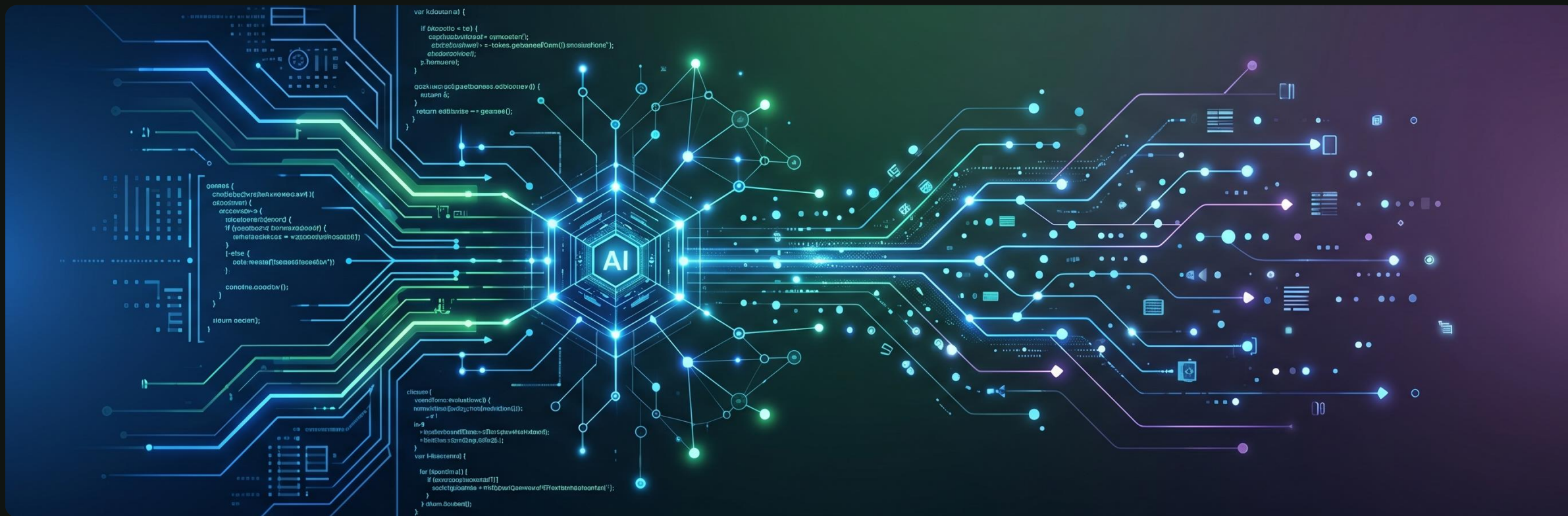
Usage Report分析ツール:

Billing and licensing → Preview your billing impact → View full billing preview

<https://copilot-billing-preview.github.com/>

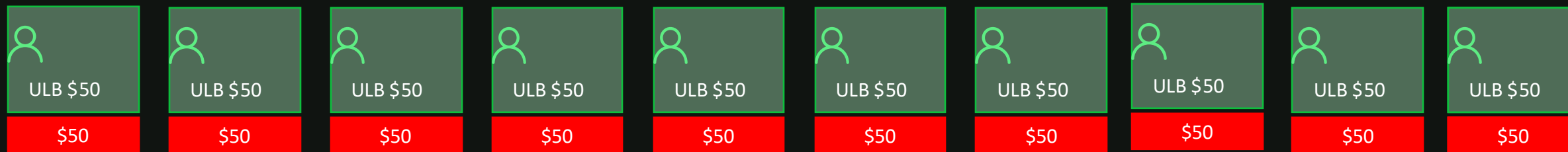


Budgetのシナリオ



均等利用シナリオ

共有プール – 上限到達 (\$190)



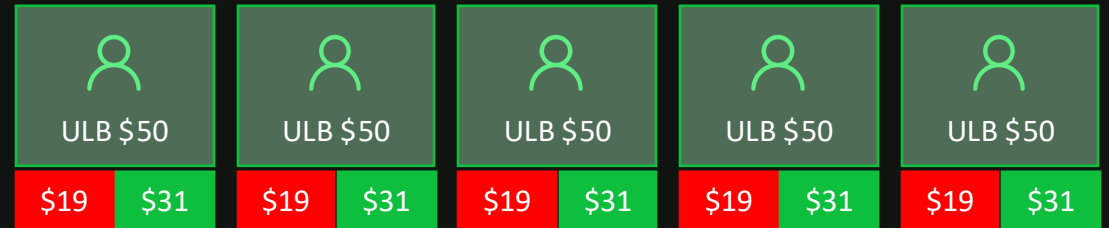
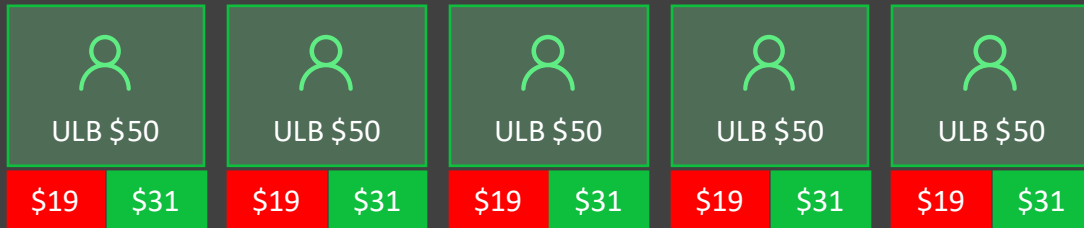
エンタープライズレベル Budget- 上限到達 (\$310)



コストセンターシナリオ

共有プール – 上限到達 (\$190)

COST CENTER A BUDGET = \$160



エンタープライズレベル Budget - (\$310)

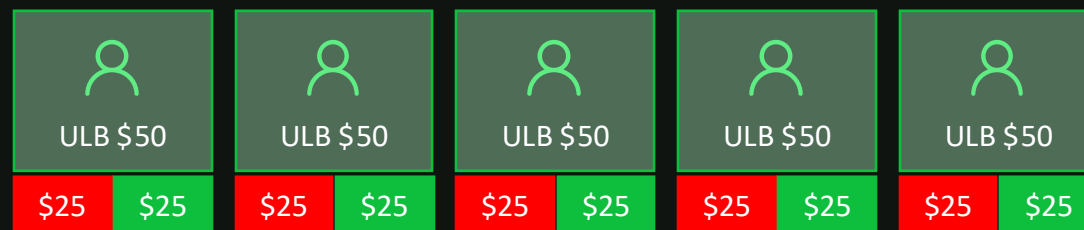
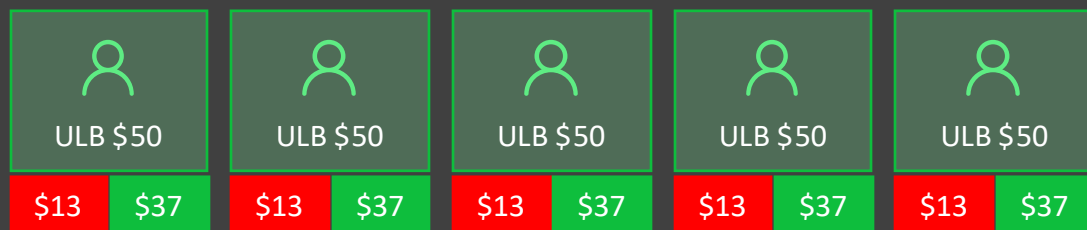
コストセンターレベルBudgetは\$160, このシナリオでは上限到達しません.
ユーザーは\$50のユーザーレベルBudgetにブロックされます。



コストセンターシナリオ – ブロック

共有プール – 上限到達 (\$190)

COST CENTER A BUDGET = \$160



エンタープライズレベル Budget - (\$310)

ユーザーレベルBudgetの残高が残っているものの、コストセンターAの超過支出が\$160に達した時点でブロックされます。



組織間のコスト按分について

コストセンターを使用すると、超過使用量を特定の Azure サブスクリプションに割り当てることができます。

ただし、共有プールに含まれる AI クレジットの按分をできない点にご注意ください。

トークン単位までコストを按分する場合は、使用状況 CSV の活用を推奨いたします。これには使用された共有プール内のトークン、追加支出、コストセンターでの支出が含まれます。こちらのデータをもとに皆様の組織内でコストを按分いただく必要があります。



Next Steps



Next Steps



自身のエンタープライズの利用状況を確認する



予算管理と今後の使用量の認識合わせをする

監視の責任者は誰か、レビュー頻度はどうするか

予算はどこから始めるべきか

誰にガードレールが必要か



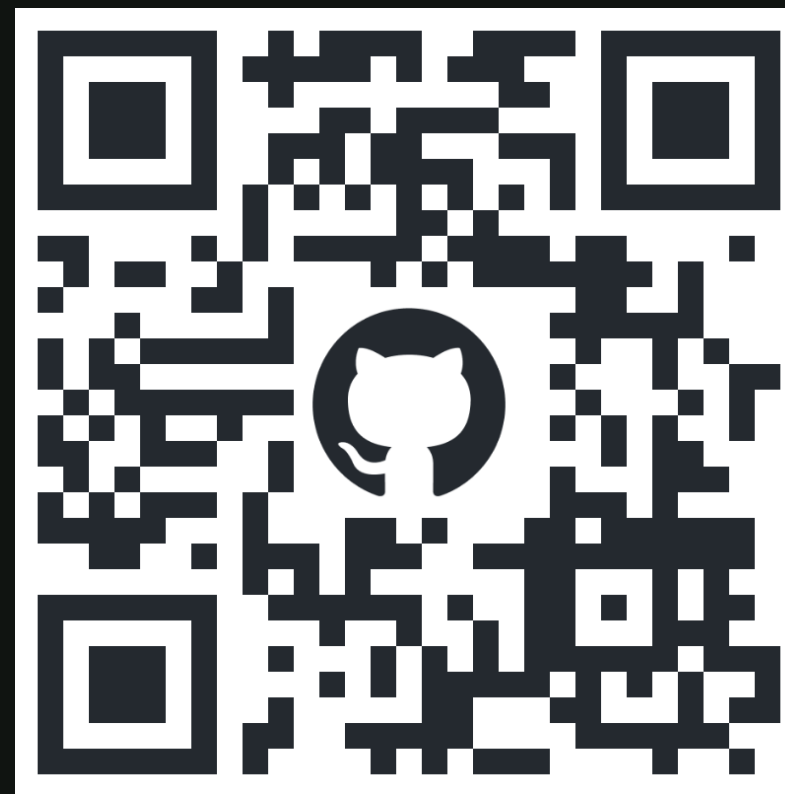
開発チームと連携して使用状況を最適化する

コンテキストサイズ削減

プロンプトを小さくする

デフォルトをAuto Modeに

イベント案内：6月23日火曜開催 GitHub After Dark Tokyo





Thank you