私が課題研究のテーマとして考えているものはビッグデータで何かできないかと考えています．

ビッグデータとは，従来のシステムでは記録、保管、解析が難しい巨大なデータで、リアルタイム性がある．

今までは管理しきれないので見過ごされていたのだが、これを解析することでビジネスに利用することができる．

（例）アマゾンや楽天などのオンラインショップでは、購買履歴やサイト内のアクセス情報などのビッグデータを基に、商品を購入する際に、他のおすすめ商品を表示している。ソフトバンクでも、同社が関わる検索サイト「ヤフー」から得られるビッグデータを積極的に利用して、他社から乗り換える可能性の高そうなユーザーを絞り込み、該当するユーザーにのみ乗換案内キャンペーンのバナー広告を表示させている。

そんで、何かないかと先生に聞いたらギットハブもビッグデータとなったので

ギットハブのデータで研究した参考文献についてこれから説明します．

参考文献にあったのはギットハブ上に挙げられているプロジェクトでどのプログラミング言語が人気かという集計がとられていました

集計を取るまでの手順は、GitHub Archive　ギットハブ　アーカイブを使い必要なデータを集めます．

GitHub Archive　は、公開されているGitHubのタイムラインを記録してそれを分析できるというもの

こうして集めたデータをGoogle BigQuery（スケールのでかいデータを分析するためのツール）これらを用いて調べた結果が先ほどの表になります．