

第 19 回アンケート結果（抜粋）

1. 一般講演 1：「既存プログラム関連資産共有のための機械学習の活用」

1. 「分布図の軸のばらつきが小さいこと」を実現する画一的な基準とは、具体的にどのようなものなのでしょうか？

⇒計算機上で実行される機械学習の客観性です。軸を生成する際には、計算機上の機械学習（アルゴリズム）によって「業務カテゴリ（受注、在庫管理等）」軸を生成しているため、主観や、属人性が存在しません。この点において、画一的な基準といえます。また、本手法では、「業務カテゴリ」軸を生成する情報として、設計書に含まれるテキスト情報を用いています。

2. モレがなく、ダブリがないようにすることが大切なのは、と思いましたが、機能の粒度をさらに細かいレベルにまで落とし込むと、プログラムの再利用に向けた動機づけにつながると思いました。

⇒アドバイスありがとうございます。クラスタリング対象となる「機能」のデータが増加すれば、より小さい粒度のカテゴリが設計できると考えております。現在はクラスタの粒度を小さくするとそのクラスタ内に含まれる「機能」数が少なくなってしまう、実現が困難です。「機能」データが増加するにつれて、粒度を細かくすることも視野に入れようと思えます。

3. 背景の説明を、もう少しゆっくり丁寧にしていただけると、理解を促進したと思えます。

4. ①想定する入力データ（ドキュメントの種類）は、どれが使えるのでしょうか？
②また、複数のプロジェクトを入れていったときに、検索することになると思いますが、クラスタ化したデータでどの程度、検索性を持たせることができますか？

⇒4-①. 本発表における可視化手法に関しては、設計書（Excel ドキュメント）のみがインプットとして利用可能です。本来の研究の範囲としては、あらゆる情報を用いるべきであると考えており、限定はしていません。傾向として非機能（性能等）に関する文書は業務に限らず似たような記述になる場合が多く、情報としての価値は低いと考えております。今後扱えるデータ（ドキュメント等）が蓄積され次第対応を考えます。

4-②. 検索性に関しては、今後実業務で評価していくほかないと考えておりますが、全文検索に加えてカテゴリを以て検索できると思いますので、その点、検索性が上がると考えております。

5. 講演が面白いです。分かりやすかったと思います。しかし、図の文字が小さいので、最後のところが良く見えませんでした。
6. 20 頁の「機能を削減することで対応した」ことが分かりません。増やすのではないのでしょうか？人によって分類の判断が違うのは、カテゴリ分けが独立していないからではないのでしょうか。

⇒クラスタに違和感がある「機能」が含まれる場合に、それらを除いて、有識者が違和感のないクラスタにするという意味です。カテゴリ分けを独立させることは非常に難しいです。論理的に完璧な独立を目指すと、検索の際に無意味なカテゴリになる場合があります。本来はカテゴリという概念より、ラベルのように複数個付与できるほうが良いのかもしれませんが。

7. いまは「業務カテゴリ」を階層化していない。関連資産の範囲が広がると対応できなくなると思います。カテゴリの階層化を考えた方が良さそうな気がします。
8. ナレッジの共有は難しいですが、早く完成するといいな、と思いました。
9. 機械学習の既存技術を使って、どこまで実務に応用できるか、という点にとっても関心があります。是非とも実評価の結果を知りたいです。