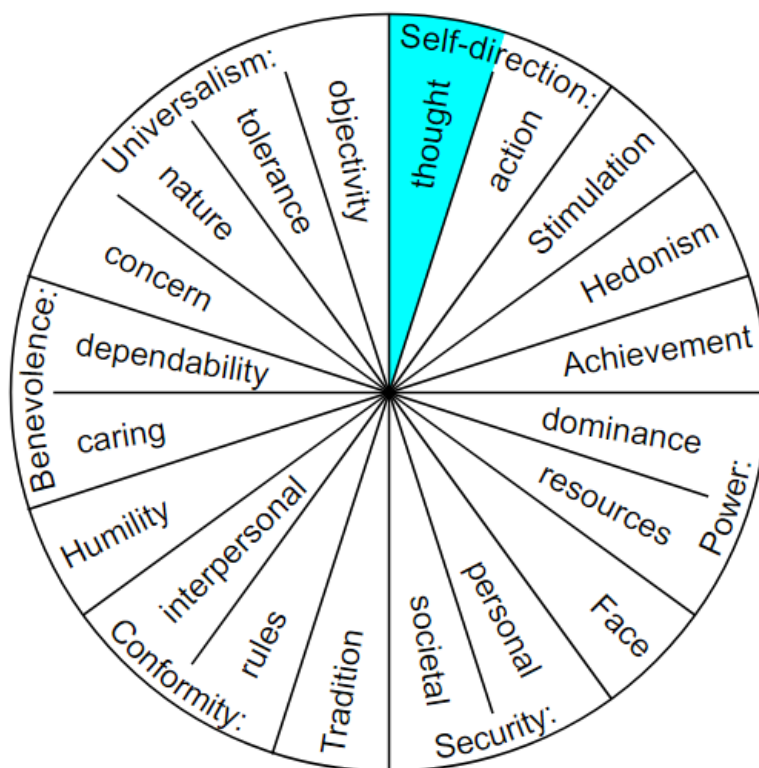


### SemEval 2023 – Task 4 – Human Value Detection behind Arguments

موضوع اصلی این پروژه استخراج ارزش های انسانی پشت یک بحث می باشد. چگونه ممکن است که مردم در مورد بهترین مسیر پیشروی در موضوعات بحث برانگیز با هم اختلاف نظر داشته باشند، با وجود اینکه آن ها از منابع اطلاعاتی مشابهی برای شکل دادن به نظر خود استفاده می کنند؟ یک راه برای دستیابی به عمق چنین اختلاف نظرهایی این است که مکرراً از آنها بپرسیم که چرا چیزی را برای خود مطلوب می بینند. ما مشاهده می کنیم که مردم اعتقادات و اولویت های متفاوتی در مورد چیزهایی که ارزش تلاش کردن دارن، دارند. یکی از اهداف اصلی این پروژه تلاش برای استخراج این ارزش ها از طریق پردازش بحث ها و نتیجه گیری بحث هاست.

دیتاست مورد استفاده در این پروژه متشکل از سه ستون اصلی (مقدمه، نتیجه گیری و جهت گیری) در ورودی و بیست کلاس خروجی است. هر یک از کلاس ها خروجی نماینده یک ارزش انسانی است. زبان دیتاست انگلیسی است اما دیتا ها لزوماً از منابع انگلیسی زبان جمع آوری نشده اند. دیتاست دارای متن رسمی می باشد. در شکل زیر بیست کلاس تشکیل دهنده دیتاست را مشاهده می کنید.



جذابیت اصلی این دیتاست برای بنده، عدم اشاره مستقیم بحث ها به ارزش مشخصی است. به عبارت دیگر باید مدلی آموزش داده شود که بتواند بر اساس الگو های درون هر بحث و کلمات کلیدی اشاره شده، ارزش یا ارزش های مورد نظر آن بحث را استخراج کند.

این دیتاست توسط تیمی از دانشگاه Bauhaus-Universität Weimer جمع آوری شده است. منابع اصلی این دیتاست از کشور های مختلف در نظر گرفته شده است تا بتوان توانایی مدل در دسته بندی داده های کشور های مختلف را به چالش کشید.

این دیتاست بخشی از وظیفه شماره چهارم از ورکشاپ SemEval 2023 می باشد. هدف از این وظیفه افزایش دقت مدل دسته بندی بر روی دیتاست داده شده است.

در این پروژه ما به دنبال finetune کردن یک مدل Roberta بر روی این دیتاست هستیم. از طرف دیگر برای بهبود دقت دسته بندی، دیتاست مورد نظر را Augment می کنیم.

## References:

[Identifying the Human Values behind Arguments](#)