دادههایی که استفاده شده سخنرانیهای شاه و امام در سال ۱۳۵۷ است.

برای دادههای مربوط به سخنرانیهای شاه از لینک زیر استفاده شده و ۳۰۹۲۸ کلمه است. های\_محمدرضا\_شاه\_پهلوی\_آریامهرE2%80%8Cسخنرانی=https://mashruteh.org/wiki/index.php?title

برای دادهّای مربوط به سخنرانیهای امام از صحیفهی نور استفاده شده و ۶۶۶۴۴ کلمه است.

دادهها به صورت دستی گردآوری شده است. و wordmapها با ماژول persian\_wordcloud کشیده شده است و حروف اضافه مانند با، تا، که، در و ... در آن حذف شده است و در هر تصویر ۱۰۰ کلمه کشیده میشود.

برای wordmapهای اختلاف و شباهت ابتدا یک دیکشنری از تمامی حروفی که هر دو شخص به کار بردند به وجود آوردیم و سپس برای اختلاف از حروفی که فقط در یک دیکشنری وجود داشت یک دیکشنری ساختیم همراه با تعداد آن و برای کشیدن wordmap استفاده کردیم و برای شباهت کلماتی که در هر دو دیکشنری وجود داشت را در یک دیکشنری جدید ریختیم و برای تعداد آن از مقدار مینیمیم بین دو نفر استفاده کردیم.

حال برخی کلماتی را که به نظرم جالب بود را میآورم:

بسم الله الرحمن الرحيم و انشاءالله در تقريبا تمامى سخنرانىهاى امام وجود داشت و كلمهى انشاءالله فقط در سخنرانىهاى امام بود.

فاسد، مسلمین، غلط فقط در سخنرانیهای امام وجود داشت.

آموزش ، پرورش ، دموکراسی ،کارگران ، آینده و اصول فقط در سخنرانیهای شاه بود.