Milestone 2 (Due: Wed, Apr 24 @ 11pm)

What progress has been made since Milestone 1 and any unexpected challenges you have encountered

- Following Milestone 1, we finalized our implementation of the resolver to actively measure and gather DNS records from the Tranco Top 1 Million dataset. To better understand and address IPv6 delegation issues, we initiated a data cleanup process. This effort aligns with the paper's examination of DNS resolution in an IPv6-only context, where it was found that a significant portion of zones could not be resolved due to incomplete DNS delegation chains, highlighting the need for comprehensive IPv6 support across the entire DNS infrastructure.

- We also implemented an `analysis.py' script to evaluate whether domains are resolvable over IPv6 by analyzing DNS response data from our initial resolver active measurement results. This script checks if a given IP address is IPv6. It also performs a recursive search through the DNS response data, checking if any IP addresses in the "Answer" section of the DNS response are IPv6. If not, it looks for further IPs to query in the "Additional" section and continues recursively until an IPv6 address is found or it reaches the maximum depth limit (to prevent infinite recursion). It does this for each domain processed, and verifies the domain's IPv6 resolvability status, indicating whether the domain can be resolved to an IPv6 address based on the available DNS response data.

- We also implemented a script to identify and record out-of-zone name servers for domains that cannot be resolved over IPv6, based on a dataset of DNS lookups. We hope to use this to further investigate DNS configuration issues, i.e., resolvability issues.

What tasks remain to be completed, who will complete these, and when they will be completed

- Currently, we're waiting for the resolver to finish running on the full Tranco dataset. Oliver and Jaanhvi are running the resolvers with the screen and updated error handling. We had hoped to complete the actual data collection by the end of milestone 2, but we got stuck due to insufficient error handling and some unexpected conditions that we weren't aware of.

- We need to clean the data, perform the necessary analysis, and create visuals to accompany the results for the final artifact submission. Additionally, we are in the process of starting to write the paper to go along with this.

- The data visualization and visual aspects can be tackled independently, depending on who is working on which part of the data analysis. The writing will mostly be a group effort, but luckily it can also be done asynchronously. In addition, we need to organize our GitHub and clean up our scripts with proper comments, which can also be done independently based on individual contributions to each script.

How the completed and remaining tasks align with the plan included in your project proposal, and any adjustments you have/will make

- Due to the unexpected challenges we've faced, we don't have a clear idea of how long data collection will take (which we are pretty confused about as to why it is taking so long and if there is any way to speed this process up). We hope to finish the analysis in 1 or 2 days unless we find more things to analyze. If needed, we'll consider stopping data collection early to start the analysis and continue collecting data in the background. We also need to move on to analyze all our artifacts for the final paper and create data visualizations.