

This dataset consists of all starter packs and all following network data available on Bluesky in January/February 2025. Starter packs can be created by any Bluesky user. They are lists of users and curated feeds with a minimum of 6 and a maximum of 150 entities, curated by the starter pack creator, and users typically name them and provide a description. Other users can use a single click to follow all entities in the starter pack, or they can scroll through a specific starter pack to decide which entities to follow within that starter pack. In our dataset, all DIDs (persistent, unique identifiers) are anonymized via a one-to-one mapping between DIDs and integers in the range [0, num_users); users in the network, as well as users who created starter packs, or appear in starter packs, are identified by a unique, anonymous integer that maps to a single DID. Similarly, starter packs themselves are identified by anonymous unique integer IDs in the range [0, num_starter_packs).

First, we include the Bluesky following network as it appeared in late January/early February 2025. This shows all available directed following relationships on Bluesky. We also include a network dataset of starter packs with information on creators and starter pack members; this is intended for users who wish to undertake a computational analysis of the networks created by starter packs or starter packs' influences on networks. Additionally, we include the starter pack network in two file formats that facilitate analysis of the starter pack hypergraph for research on higher-order network interactions.

Filename	Description	Data Structure / Types
deidentified_starterpack_blob.json	JSON blob (list) of all starter packs found by scraping all starter packs for all users in the Bluesky PLC directory as of January 2025. Contains pack membership and metadata. All timestamps are rounded to the nearest day to preserve anonymity.	List of dictionaries, each with the following keys and values: pack-id : integer; anonymized identifier of the starter pack list creator-id : integer; anonymized identifier of the starter pack's creator creator-date-created : date; the date on which the creator's account was created date-created : date; the date on which the starter pack was created members : list of members' anonymized integer identifiers
deidentified_starterpack_hif.json	HIF-compliant JSON file. For details on the HIF standard, see the associated repository .	In the output, nodes are anonymized integer identifiers and edges are starter packs.

deidentified_starterpack_edgelist.csv	A hyperedge list stored as a CSV file. Each line in file corresponds to a starter pack and each entry in the line is a member of the starter pack.	Each line contains entries separated by commas and corresponds to a starter pack. E.g., "1,2,3,4" indicates that users with anonymized identifiers 1, 2, 3, and 4 belong to the same starter pack.
deidentified_follows_edgelist.csv	CSV that delineates who follows whom on Bluesky.	An entry of A, B indicates that user A (i.e. the integer identifier corresponding to DID A) follows user B (i.e. the integer identifier corresponding to DID B).