PNAS

RESEARCH ARTICLE | Social Network Analysis

The Architecture of Aspiration: A Network Perspective on Human Goals

Eve Carney[a,1], Rebecca Schmieder[a], and Nico Sherpa[a]

Goals guide human behavior and shape life outcomes, serving as the fundamental units of motivation and purpose. While psychological frameworks have long aimed to categorize and investigate these ambitions, research has largely relied on theoretical models and small-scale surveys, leaving the global topology and interconnectivity of human aspirations unmapped. Here we analyze the landscape of human ambition using digital trace data from the DayZero platform, constructing a co-occurrence network of goals to reveal how aspirations cluster and connect. We show that the network exhibits small-world properties and is governed by a steep power-law distribution. After establishing open-coding–inspired semantic themes and classifying our goals using LLMs, we find that structural communities do not align with semantic categories. Instead, users weave together diverse aspirations, creating strong cross-category ties between distinct themes like altruism and social connection, whereas some themes such as religion remain more tightly connected to themselves. These results demonstrate that human ambition is not compartmentalized but rather broad, heterogeneous, and deeply interlinked across domains. Our findings establish a data-driven approach to understanding the collective patterns of human aspiration, highlighting that the pursuit of a meaningful life is a highly interconnected and shared phenomenon.

Networks | Social Network Analysis | Life Goals | Digital Trace Data | Computational Social Science

Significance

We study the landscape of human ambition by analyzing a network of goals derived from digital trace data from the DayZero platform. While past research has focused on theoretical or small-scale empirical frameworks, we here show that aspirations form a cohesive small-world network driven by shared ”hub” experiences. We outline a computational pipeline using LLMs to characterize the structural properties of these ambitions, revealing that individuals frequently bridge distinct domains to create complex narratives. We utilize an array of network metrics to analyze key characteristics of the goal network. This work offers insights into fields such as computational social science, psychology, and network science by demonstrating that human goals are not isolated desires but elements of a globally interconnected architecture.

We all have goals in life. Whether written down on a bucket list, revealed in conversations with friends, or reflected in choices we make, goals guide both the small and large choices we make in life. While some people aim to be successful in their careers, wealthy, or famous, others want to see the world, live healthily, or help others. Despite a wealth of literature focusing on goal setting, categorization of goals, and impact of goals on life outcomes, there is little research mapping the goal landscape of humans on a larger scale. To explore this gap, our study uses digital trace data, retrieved from the DayZero project platform, to map the landscape of human aspiration and investigate the network of goals created based on it. This is guided by the research question: *What are humans aiming to achieve throughout their lives, and how is the network of their ambitions characterized, connected, or clustered? A case study of goals on the DayZero platform.*

As a widely studied subject, the definitions behind human goals and ambitions vary. To map the landscape of human goals, we define goals broadly as meaningful, future-oriented desires or needs directing people towards outcomes and even purpose in life (inspired by (1)). Human goals have been studied for a long time and can be understood in the context of fundamental psychological frameworks, such as McClelland's theory of human motivation (2), or Deci and Ryan's self-determination theory (3). Psychological research has investigated the motivation behind setting goals, as well as the influence of goals and goal achievement on life outcomes (4). Furthermore, a considerable amount of research is conducted in the context of disease (5) (6). Other popular foci of research focus on goals look at new years resolutions or bucket list (7). Past research has found goals related to social and romantic connection (8), physical fitness (7), physical health, personal growth and academic achievements (9) and tourism-related topics (10) to be among the most popular. Additionally, authors, such as Talevich et al., have aimed to understand the general landscape of goals, suggested the existence of nine clusters of human motives (1), based

Author affiliations: [a]Københavns Universitet

E.C., R.S., and N.S. designed research, performed research, and analyzed data. E.C. was responsible for characterization of the network and sentiment analysis. R.S. was responsible for scraping the data and analyzing clustering. N.S. was responsible for pre-processing the data, classification, and analyzing connectivity. All authors completed the writing of the paper together and contributed equally throughout the whole project.

The authors declare no conflict of interest.

[1]To whom correspondence should be addressed. E-mail: djp898@sodas.ku.dk

on a range of prior structures proposed.

The main methodological focus of existing goal and ambition related research is theoretical, based on literature reviews, qualitative studies or experiments and surveys. With advances in technology, we can instead utilize digital trace data and machine learning algorithms to map, understand and investigate the landscape and network of human ambition. Digital trace data is defined as passively collected, fine-grained records of user interactions with digital platforms, devices, or services, generated as a by-product of original system use rather than created for research purposes (11). For our project, we chose to use data from the platform DayZero project, a site made to create lists of goals and engage in a community of goal setters (12). To create these lists, users need to register and can choose from either pre-set or individually created goals and lists. Goals can be marked as *in progress*, or *done*, and can have notes or comments attached to them. Users can follow one another and interact with others' achievements. While the exact number of users is not publicly known, the platform, which has been in existence since 2009, claims it has 'the largest community of goal setters in the world' (12). Web traffic analysis tools reveal that most users of the platform come from the United States, followed by Australia, the UK and Canada (13). This is also reflected in English being the predominant language used on the platform.

In our results section, we describe how the network was created and the analyses performed on it. We examine characteristics of the network and explore connections and clusters within it, showing that human aspirations are complex, diverse and heterogeneous.

Results

Network Creation & Background. Due to the platform's nature, with no overview of all goals or users existing, the only way to access goals and users in a semi-structural way is by using the search function, in which one can search for users by username. We therefore started with a set of the 400 most common names from the 2000s in the US (14), and scraped the results page of usernames found for each search query, resulting in a list of 12,998 unique usernames. Next, we scraped the list of all goals available for each user, resulting in 231,269 unique goals. We then scraped the page of each goal to retrieve additional attributes available such as their IDs, titles, descriptions, comments, assigned tags, and the number of people who wanted to do the goal and who had completed it. To be able to perform complete semantic analyses on the entire network of goals, we decided to only keep those with a description available. Additionally, to ensure that goals, which are described differently but substantially the same, for example *Learn Spanish* and *Learn to speak Spanish*, are treated as the same goal, we used a sentence transformer model (all-MiniLM-L6-v2) to create embeddings and merge goals above a similarity threshold of 0.9. The process of merging goals, keeping only those where at least one description was available, resulted in reducing 866 goals into 283 representative merged goals. We decided for a higher threshold that emphasizes a low false positive rate, at the cost of missing some similar goals that are not detected. Overall, we are left with a network of 2890 nodes and 219130 edges. When necessary, we worked with the giant connected component of our network, consisting of 2860 nodes and 219126 edges.

A node in our network represents a goal, and an edge is placed between two nodes if the two goals co-occur on at least one user's list. Each node has a number of attributes, namely the goal attributes scraped from the websites as well as a popularity score indicating how many of the users in our sample wanted to or had completed the goal. Additionally, the edge weight indicates on how many lists the goals co-occur, determined by adding up all weights in one group if goals are merged. Our decision to build a network centered around goals rather than users is both based on our main research question as well as the lack of available information about users.

In addition to the pre-existing goal attributes retrieved from the platform, we decided to create a category attribute for each goal to classify our goals and later be able to compare our classification to structural communities. To create categories of goals, we first used SBERT embeddings of our goals' titles and descriptions together with k-means topic clustering to create an initial inspiration for possible categories. This process left us with 20 topics and a number of indicative keywords for each. We then conducted a manual open, qualitative coding process by exporting all goals, including the SBERT topic keywords, to excel and manually creating categories in an iterative process. After combining and validating our results, we identified 15 distinct categories, such as *Travel destinations*, *Health*, *Academic and professional achievements*, *Media consumptions* and *Religion*. As can be seen in Figure 1 travel related categories, such as *Travel destinations*, *Places of interest*, and *Nature*, make up about 75% of all goals, while other categories, for example *Health* and *Creativity* are very small.

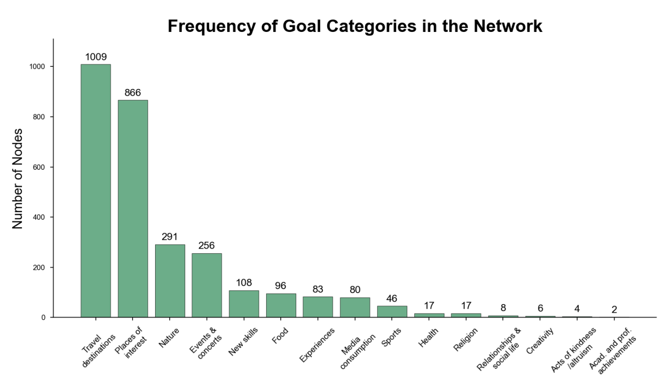

Fig. 1. Distribution of goal categories within the network. The bar chart displays the frequency of nodes assigned to each specific category, sorted in descending order. The number above each bar indicates the total count of nodes for that category. The data reveals a strong skew towards travel-related interests.

We then applied an LLM classification pipeline to obtain a structured, mutually exclusive category classification for each result, using the Gemma 3 (27B, IT) model with a low temperature (0.1) and top-p setting (0.3), as recommended by Törnberg (15). Additionally, we assigned a country, city, and other mappable locations to each result to the extent possible, using the same logic but a more capable model (Gemini 2.5 Flash-Lite).

Network Characterization. While our network is relatively sparse, with a density of 0.052 and an average degree of

151.647, its diameter of 5, the average clustering coefficient (0.776) and average shortest path length (2.051) indicate that the network possesses small-world properties. This means that due to a small number of high degree nodes and a small average shortest path length, the distance between any two goals is relatively small. We also conducted an analysis of the distribution of the degree of the nodes in our network, to observe the structure of our network. As shown in Figure 2A, the distribution of the degrees skews to the right with only a few high-degree nodes, such as *See the Northern Lights* (degree=2089), *Get a Tattoo* (degree=1767) and *See the Grand Canyon* (degree=1756). As these goals are shared by many users, they have a high relevance in connecting our network. Further investigation into the distribution, particularly in respect to Power-Law Exponent ($\gamma = 3.575$), highlights the steep degree distribution.

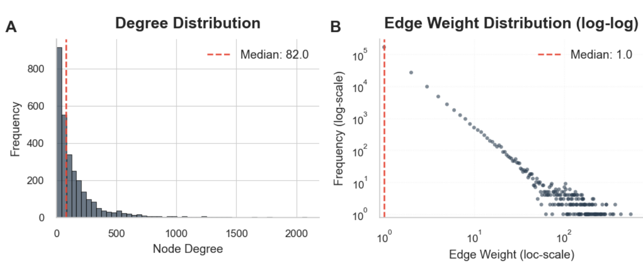

Fig. 2. Evidence of heavy-tailed distributions in the network topology. (A) The Degree Distribution histogram is right-skewed with a median degree of 82.0 (red dashed line), indicating that while most goals have fewer than 100 connections, a few "hub" goals have over 1000. (B) The Edge Weight Distribution is plotted on a log-log scale to visualize the frequency of edge weights (number of users sharing two goals). The linear trend on this scale suggests a heavy-tailed distribution, where the median edge weight is 1.0.

Utilizing degree, betweenness and eigenvector centrality, both to observe influence and whether some nodes act as bridges across our network, we find that *See the Northern Lights* has the highest centrality, regardless of method. On the opposite end of the centrality measures, our lowest scoring centrality nodes, are more specific and obscure goals, such as *Go hot-air ballooning at Château-d'Oex*, and *Visit Cathedral Caverns State Park*. Overall, our network has a similar structure to other social networks, insofar as having high average clustering, a few highly central nodes and small world properties.

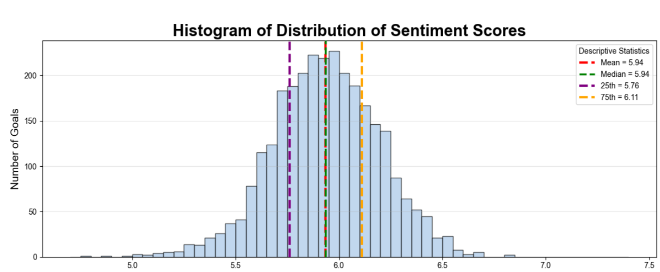

Fig. 3. Distribution of Sentiment Scores, based on the LabMT 'average happiness evaluation' score. Alongside a normal distribution curve, the median, mean and quartile ranges are also marked by color-coordinated dashed lines. These key values are markedly close to one another, especially that of the mean and median score.

When investigating the language of goals in both titles and descriptions, utilizing the LabMT wordlist (16), we find a more positive sentiment, following a normal distribution with a mean and median of 5.94, as shown in Figure 3. Further investigation into the nodes themselves shows that our highly positive nodes focus on topics such as baking, visiting gardens and love-related activities, with the top goal of *Bake a Chocolate Cake*, scoring 7.400. For our lowest scoring nodes, these focus on visiting locations related to war and violence, and also activities focusing on blood. Our lowest scoring goal was *Visit Gettysburg*, a civil war site in the US, scoring 4.793. Overall, the more positive sentiment within our network is expected, as these are undertakings that people actively want to pursue, rather than obligations in their daily lives.

Connected Goals. We now examine how goals connect to one another, both individually and through the patterns that emerge between categories. The distribution of edge weights, as shown in Figure 2B, reveals that most connections within the network are weak (median weight = 1; 95th percentile = 5). The three strongest connections, defined as the edges with the highest total weight between two goals, all involve *Donate blood* in combination with *Get a tattoo* (553), *Leave an inspirational note inside a book for someone to find* (429), and *See the Northern Lights* (371).

From an aggregated category perspective, most connections occur across categories (72.1% of edges; 77.0% of the total weight) rather than within categories (27.9% of edges; 23.0% of the total weight). The three most connected category pairs are *Places of interest* *Travel destinations* (37,433 edges; 56,802 total weight), *Nature* *Travel destinations* (12,609 edges; 21,153 total weight), and *Nature* *Places of interest* (13,447 edges; 19,929 total weight). These patterns are strongly influenced by the fact that these three categories are the most frequently assigned, which puts them in a favorable position to accumulate many connections. We therefore additionally considered measures that account for this size effect. Similar to Newman (17), who did this for individual node vertices rather than weights, we calculated the expected weight between two categories by multiplying the total weights of each category and dividing by the total weight in the network. This allows us to assess whether two categories are connected more strongly or more weakly than expected. Notably, every possible pair of categories is represented in the data, giving a comprehensive view of how themes relate to each other, although connections involving smaller categories still need to be interpreted with caution and are best seen as indicative patterns.

The heatmap visualizing the pointwise mutual information (PMI; where PMI ≈ 0 means about as often as random mixing, PMI ≈ 1 means about twice the expected value, and negative PMI means less than expected) shows that categories vary widely in terms of over and under expected connections. Some categories have PMI values close to zero for almost all cross category connections (that is, around the expected total weight), while goals in those categories are strongly over connected to goals from the same category. This is, for example, the case for *Religion* (within category PMI = 1.66) and *Events & Concerts* (within category PMI = 0.93). Other categories show stronger variability in their cross category connections. For instance, *Acts of kindness/altruism* is more strongly connected than expected to *Relationships & Social life* (PMI = 1.21), which is the highest cross category PMI overall, as well as to *Creativity* (PMI = 1.20) and *Academic and professional achievements* (PMI = 1.08). At the same

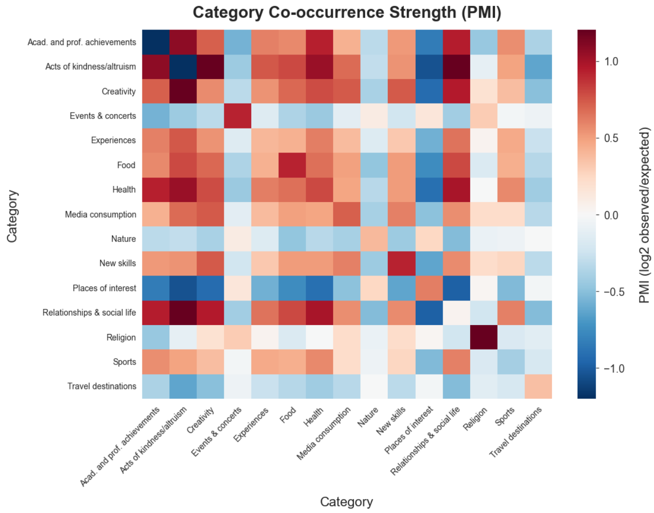

Fig. 4. PMI heatmap of goal category co-occurrences. Heatmap shows $\log_2$(observed/expected) co-occurrence ratios between categories. Red indicates positive associations (goals co-occur more than expected), blue indicates negative associations (goals co-occur less than expected).

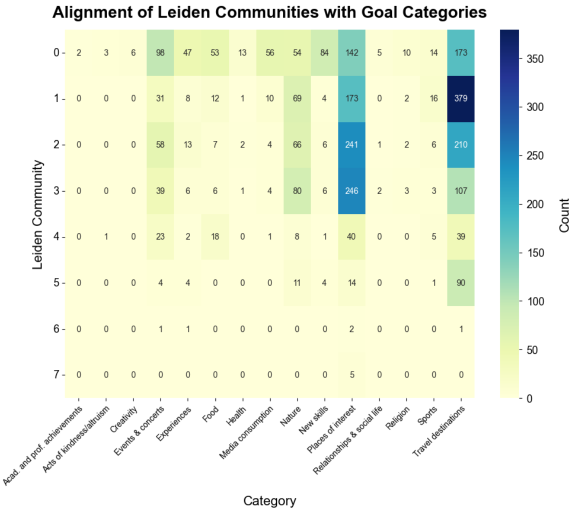

Fig. 5. Correspondence between structural network communities and semantic categories. A comparison of unsupervised community detection (Leiden algorithm) against manual topic labeling. The heatmap displays the overlap between communities detected via the Leiden algorithm (rows 0–7) and manually assigned goal categories (columns). The cell values (counts) and color intensity indicate the number of nodes shared between a specific structural community and a semantic category.

time, it has a strongly under expected connection to *Places of interest* (PMI = -0.92), which is the lowest cross category PMI overall.

Clustered Goals. Having examined the most prominent goals and their interconnections, we then investigated clusters in the goal network. To do so, we used both the manually developed categories described earlier and structural communities detected using the Leiden algorithm (18). The manual categorization produced 15 communities with a modularity of 0.07, indicating that these topic-based groups do not divide the graph in a structurally meaningful way. In other words, they contain only slightly more internal edges than would be expected at random, which is further supported by a neutral but slightly positive category-based attribute assortativity (0.075). In contrast, the Leiden algorithm identified eight structural communities, ranging in size from 5 to 760 nodes, with six exceeding 100 nodes. This partition achieved a modularity of 0.24, suggesting a moderately clear community structure.

The strong difference between the modularity of the Leiden partition, showing clear structural clustering, and the category based partition, having very low modularity, indicates that structural communities and topic based ones may not overlap much. This is furthermore confirmed by comparing the resulting communities based on categories with the ones based on the Leiden algorithm using the heatmap shown in Figure 5. Most topics, for example *Travel destinations*, are divided amongst many of the Leiden communities, confirming the meaning of human goals captured in the categories they are assigned to does not translate into structurally cohesive groups within our network. This could be due to several reasons, including the fact that semantic themes do not govern co-occurrence behavior, the limit of how many categories were created and how well they capture micro-behavior and ambitions, the relative sparseness of our network, and the overall set-up of our network and the biases introduced by it.

Overall, we find a goal landscape with several coherent themes, which, however, do not correspond to structural communities in our network. Furthermore, we find themes are interconnected and users tend to pursue goals spanning multiple categories rather than remaining confined within a single thematic domain.

Discussion

In our study, we constructed a network of goals to study the landscape of human ambition, using data from the DayZero project platform. We found that our network is relatively sparse, with a few high degree nodes dominating the degree distribution, which is in line with the expectation for social networks. It is important to note that goals like *Donate blood*, *Get a tattoo*, or *Visit the Grand Canyon* appear on the website's pre-made top-101 list, which likely increases the probability that users add them to their own lists and thus boosts their measured popularity. This suggests that the decision to include goals might reflect platform design rather than users' independent preferences.

Besides identifying popular and central goals, our analysis shows that the DayZero network is broadly interwoven, with mostly weak and few strong ties and links mainly between goals from different categories. By looking at the connections between different goals and the categories they were placed into, we could see that certain themes co-occur more frequently and might thus be more closely tied to one another. For example, we found stronger-than-expected links between *Acts of kindness/altruism* and *Relationships & social life*, and weaker links to *Places of interest*, indicating that altruistic ambitions are more closely tied to social meaning and self-development than to visiting specific places. We also found that there are differences between how strongly goals of the same category are connected to each other. For example, we found strong within-category connections in goals within *Religion* and *Events & concerts*, suggesting that some domains of ambition are pursued in a more focused, domain-specific way, pointing to coherent life themes within an otherwise broadly interconnected goal network.

Finally, we investigated goal communities, we found that structural and semantic communities showed limited overlap, likely due to differences between semantic similarity and behavioral co-occurrence, as well as the network's hub-dominated structure. All categories co-occurred within the network, suggesting that users pursue a wide mix of goals. Travel-related themes, such as *Travel destinations*, *Places of interest*, and *Nature*, were among the most common, alongside *Events & concerts*, and learning new skills. This aligns with previous work highlighting the prominence of travel pursuits on bucket lists (19) (10). In contrast, prior studies emphasizing social and romantic goals (8) are not reflected in our data. These patterns may stem from the nature of a public goal-sharing website, which shapes what users choose to reveal and may attract specific user groups. As personality traits have been shown to predict goal pursuit (20), the platform's user base could influence the overall distribution of goal types in the network.

The potential importance of user characteristic for better understanding the goal landscape portrayed in our network indicates one of the limitations of our data source. Most DayZero users are anonymous, leaving us without any user related data to explain our findings in the context of user characteristics, nor correlate these two. Sampling limitations also arise from the platform's structure. DayZero provides no overview of users or goals, and goals are scattered across individual lists. Because we used only goals with descriptions—less than 2% of all scraped goals—the sample is narrow and vulnerable to platform-driven biases. Our merging procedure helped consolidate semantically similar goals, but the conservative threshold likely excluded additional identical goals and cannot address distinct goals without descriptions. The sampling method used, resembling non-probability convenience sampling based on an initial set of 400 names, does not provide us with a representative sample. To inspect the impact of our sampling, we compared the top 101 most popular goals in our network (based on how many people want to do them) with the website's top 101 list. While we found only about 36% overlap, all top ten goals were included in both lists, indicating that our sampling process, though not fully representative, reliably captures the platform's most prominent goals.This limits the generalizability of our results, both within the platform and beyond. Besides constraints given by the structure of the platform, limits of available time and resources dictated the sampling method chosen.

Future studies could aim to use data sites which provide more insight about users in order to be able to investigate more user centered questions surrounding the topic of human ambitions. Additionally, our dataset contains a few variables not used in our primary analysis, including tags attached to goals by the platform itself and comments of users. These could also be used to gain additional insight about how users frame their goals, articulate motivations, and interact around shared ambitions.

Overall, our study mapped the landscape of human ambition by constructing a goal-to-goal network from DayZero users' digital traces; a novel approach in this context. We found that people pursue a remarkably wide range of aspirations: from seeing the northern lights to donating blood to visiting iconic places. Most individuals aim for many goals spanning diverse categories such as travel, learning, and personal experiences. Taken together, our network shows that human ambitions are broad, interconnected, and often shared. By striving for fun, meaningful, deep and challenging goals, humans aim to make the most out of their limited time on this planet.

Materials and Methods

Data Source. The data used was retrieved from dayzeroproject.com.

GitHub. An extensive description of the methods used, including all code and data needed for our analysis can be found in the explainer notebook linked here.

Ethics. The dayzeroproject website allows scraping according to its robots.txt and ToS. No individuals can be identified in our network and thus no GDPR relevant data is included.

Point Mutual Information (PMI). This method compares the probability of two events occurring together. $\mathrm{PMI}(x, y) = \log \frac{p(x,y)}{p(x)p(y)}$, where $p(x, y)$ is the observed joint frequency of x and y divided by the total number of observations, and $p(x)p(y)$ is the expected joint probability of x and y under the assumption that they are independent.

Leiden Community Detection. On graph $G = (V, E)$, the Leiden algorithm finds a partition $\mathcal{P}$ by greedily maximizing a quality function (e.g., modularity or CPM) such as $H = \sum_{c \in \mathcal{P}} \big(e_c - \gamma n_c(n_c - 1)/2\big)$, where e_c and n_c are the internal edge count and size of community c, and γ sets the resolution. It iterates three phases: local node moves, partition refinement, and graph aggregation, until convergence.

1. JR Talevich, SJ Read, DA Walsh, R Iyer, G Chopra, Toward a comprehensive taxonomy of human motives. *PloS one* **12**, e0172279 (2017).
2. DC McClelland, *Human Motivation.* (Cambridge University Press, Cambridge), (1987).
3. EL Deci, RM Ryan, The" what" and" why" of goal pursuits: Human needs and the self-determination of behavior. *Psychol. inquiry* **11**, 227–268 (2000).
4. EE Messersmith, JE Schulenberg, Goal attainment, goal striving, and well-being during the transition to adulthood: A ten-year us national longitudinal study. *New directions for child adolescent development* pp. 27–40 (2010).
5. K Ikeuchi, S Nishida, M Karikawa, Life goal domains, traits, and setting process in the collaboration between healthcare professionals and cancer survivors: a scoping review. *Prog. Rehabil. Medicine* **9**, 20240018 (2024).
6. J Avery, et al., Rewriting the script: How covid-19 affected the relation between intrinsic aspirations and depressive symptoms. *Pers. Individ. Differ.* **200**, 111869 (2023).
7. JM Dickson, A Hart, C Fox-Harding, CD Huntley, Adaptive goal processes and underlying motives that sustain mental wellbeing and new year exercise resolutions. *Int. journal environmental research public health* **20**, 901 (2023).
8. JF Benenson, H Markovits, Young adults' desired life tradeoffs: love first, sex last. *Sci. Reports* **14**, 19680 (2024).
9. M Oscarsson, P Carlbring, G Andersson, A Rozental, A large-scale experiment on new year's resolutions: Approach-oriented goals are more successful than avoidance-oriented goals. *PLoS One* **15** (2020).
10. S Zascerinska, R Sharpley, D Wright, Living life or denying death? towards an understanding of the bucket list. *Tour. Recreat. Res.* **49**, 221–234 (2024).
11. J Breuer, L Bishop, K Kinder-Kurlanda, The practical and ethical challenges in acquiring and sharing digital trace data: Negotiating public-private partnerships. *New Media & Soc.* **22**, 2058–2080 (2020).
12. DayZero, The online community of people all over the world who love creating lists, setting challenges, and making positive changes in their lives. create an action list not a bucket list! (n.d.).
13. SimilarWeb, Similarweb identity (https://pro.similarweb.com/#/digitalsuite/home) (year?) Accessed: 2025-12-06.
14. US Government, Top names of the 2000s (n.d.).
15. P Törnberg, Best practices for text annotation with large language models (2024).
16. PS Dodds, KD Harris, IM Kloumann, CA Bliss, CM Danforth, Temporal patterns of happiness and information in a global social network: Hedonometrics and twitter. *PloS one* **6**, e26752 (2011).
17. ME Newman, Modularity and community structure in networks. *Proc. national academy sciences* **103**, 8577–8582 (2006).
18. VA Traag, L Waltman, NJ Van Eck, From louvain to leiden: guaranteeing well-connected communities. *Sci. reports* **9**, 1–12 (2019).
19. VS Periyakoil, E Neri, H Kraemer, Common items on a bucket list. *J. Palliat. Medicine* **21**, 652–658 (2018).
20. BW Roberts, RW Robins, Broad dispositions, broad aspirations: The intersection of personality traits and major life goals. *Pers. social psychology bulletin* **26**, 1284–1296 (2000).