

## Recommended analytical workflow

The workflow treats the corpus as a **dynamic, overlapping co-authorship network** enriched with paper-level topical content. It deliberately distinguishes:

- **Communities:** groups inferred from recurring collaboration structure, with authors allowed to belong to multiple communities.
- **Core and leaders:** authors with persistent, central membership within a community—not merely highly cited or prolific authors.
- **Research foci:** topic distributions derived from titles, abstracts, and keywords, then aggregated for each community and time period.
- **Evolution:** continuity, overlap, membership flows, and structural events such as merges, splits, emergence, and dissolution.

The analyst first creates annual or rolling multi-year co-authorship-network snapshots from disambiguated author IDs. A dynamic overlapping-community model then identifies communities while preserving membership strengths across time. The analyst iteratively evaluates whether the detected communities are structurally coherent, temporally plausible, and interpretable; they can adjust temporal resolution, community-resolution settings, overlap thresholds, and persistence regularisation.

In parallel, a topic model derives paper-level research-focus vectors from textual metadata. The analyst checks topic coherence and revises topic-model settings where necessary. Community-level topic profiles are then obtained by aggregating the papers and authors associated with each community in each time period.

The central visual interface combines:

1. **Temporal community evolution view:** alluvial/lineage view showing communities over time, member retention, merges, splits, emergence, and dissolution.
2. **Community network view:** node-link or NodeTrix-style view of selected community members and inter-community overlaps.
3. **Membership and leadership view:** ranked persistent-core members, bridge authors, centrality, productivity, citations, and membership strengths.
4. **Topic evolution view:** topic-by-community heatmaps and temporal profiles, linked to representative papers.
5. **Evidence view:** selected papers, abstracts, keywords, citation links, and author histories supporting interpretation.

The analyst iteratively selects suspicious or important transitions—such as a split, merger, or loss of a core—and investigates associated author movements and changes in topic distributions. This supports explaining *when* a change occurred, *which authors* drove it, and *why* it may have occurred substantively.

This design draws particularly on:

- **Dynamic Network Exploration / “Reducing Snapshots to Points”** for snapshot-based dynamic-network representations and temporal structural exploration.
  - **NodeTrix / “Hybrid Network Community Analysis”** for inspecting dense communities while retaining inter-community links and bridge actors.
  - **Human-Steered Topic Modelling / “Interactive Topic Modeling”** for iterative evaluation and refinement of interpretable research topics.
  - **EventThread Analysis** for representing branching, merging, and evolving group structures over time.
  - **Cluster-calendar workflow** for analyst-guided refinement of clustering parameters and interpretability assessment.
- 

## ATWL representation

atwl

```
workflow ieee-vis-community-evolution-analysis
```

```
template: define-unit (authors and temporal networks)
  characterise (network and text features)
  loop(build-model (overlapping communities)
    define-unit (infer communities)
    build-model (topics)
    characterise (community profiles)
    contextualise visualise
    abstract assess
    generate-knowledge (refine specifications))
  contextualise visualise
  loop(abstract define-unit (select transition)
    visualise abstract assess)
  generate-knowledge
```

```
description: "Identify, characterise, validate, and explain the evolution
of overlapping research communities in IEEE VIS publications from
1990–2024. Communities are inferred from dynamic co-authorship
```

structure, linked to research topics derived from paper text, and interpreted through coordinated visual exploration."

```
# =====  
# INPUT  
# =====
```

```
artifact D_papers : entities  
  origin: given  
  internal structure: elementary  
  embedment: time  
  features:  
    - id: f_paper_id  
      value structure: atomic  
      value type: reference  
      description: "Unique paper identifier"  
    - id: f_year  
      value structure: atomic  
      value type: temporal  
      description: "Publication year, 1990-2024"  
    - id: f_authors  
      value structure: list  
      value type: reference  
      description: "Disambiguated DBLP author identifiers with raw name"  
    - id: f_title  
      value structure: atomic  
      value type: text  
      description: "Paper title"  
    - id: f_abstract  
      value structure: atomic  
      value type: text  
      description: "Paper abstract"  
    - id: f_keywords  
      value structure: list  
      value type: text  
      description: "Author- and venue-provided keywords"  
    - id: f_citations  
      value structure: list  
      value type: reference  
      description: "Cited paper identifiers"  
  description: "IEEE VIS, InfoVis, VAST, and SciVis papers published bet  
    1990 and 2024, with disambiguated author identities, bibliographic t
```

keywords, and citation information"

artifact S\_temporal\_network : specification

origin: given

representation form: "temporal-network construction parameters"

description: "Initial choice of annual or rolling multi-year windows,  
edge weighting rule (e.g., co-authored-paper count with recency decay,  
minimum activity threshold, and inclusion of citation-derived relationships"

artifact S\_community\_model : specification

origin: given

representation form: "dynamic overlapping-community model parameters"

description: "Initial settings for number or resolution of communities,  
overlap regularisation, temporal persistence regularisation, minimum  
membership strength, lineage matching threshold, and core-member  
persistence threshold"

artifact S\_topic\_model : specification

origin: given

representation form: "topic-model parameters"

description: "Initial topic count, text preprocessing, term weighting,  
use of title/abstract/keywords, and topic-coherence criteria"

artifact K\_domain\_context : knowledge(D\_papers)

origin: given

representation form: "domain annotations and interpretation notes"

description: "Analyst knowledge about visualization research areas,  
historically important venues, known research programs, institutions  
and potentially meaningful author or topic changes"

# =====

# STEP 1: DEFINE AUTHORS AND TEMPORAL CO-AUTHORSHIP NETWORKS

# =====

transform T\_define\_authors :

intent: define-unit

manner: "group papers by disambiguated author identifier"

input: D\_papers

output: D\_authors

actor: machine

description: "Create one author entity per DBLP-disambiguated identifier  
retaining links to authored papers and activity history"

```
artifact D_authors : entities
  internal structure: elementary
  embedment: set
  features:
    - id: f_author_id
      value structure: atomic
      value type: reference
      description: "Disambiguated DBLP author identifier"
    - id: f_raw_names
      value structure: list
      value type: text
      description: "Observed raw author-name strings"
    - id: f_authored_papers
      value structure: list
      value type: reference
      description: "Papers authored by this researcher"
    - id: f_active_years
      value structure: list
      value type: temporal
      description: "Publication years in the IEEE VIS corpus"
  description: "Unique visualization-community researchers identified through
    DBLP-based author disambiguation"
```

```
transform T_construct_networks :
  intent: define-unit
  manner: "time-partitioning and co-authorship-network aggregation"
  input: D_papers, D_authors, S_temporal_network
  output: D_network_snapshots
  actor: machine
  description: "Construct annual or rolling-window weighted co-authorship
    network snapshots; nodes are active authors and edges represent
    collaboration strength"
```

```
artifact D_network_snapshots : entities
  internal structure: formation
  embedment: time
  features:
    - id: f_coauthor_edges
      value structure: relational configuration
      value type: {reference, numeric}
      description: "Weighted co-authorship graph for each temporal snapshot"
```

```

- id: f_active_authors
  value structure: list
  value type: reference
  description: "Authors active in the snapshot window"
- id: f_snapshot_period
  value structure: atomic
  value type: temporal
  description: "Year or rolling time-window interval"
description: "Temporal sequence of weighted co-authorship-network snapshots
representing evolving collaboration structure"

# =====
# STEP 2: DERIVE NETWORK AND TEXTUAL FEATURES
# =====

transform T_characterise_network :
  intent: characterise
  manner: "compute graph statistics and inter-author relations"
  input: D_network_snapshots, D_authors
  output: F_network_structure, F_author_centrality
  actor: machine
  description: "Compute node centrality, weighted degree, brokerage,
    collaboration persistence, pairwise collaboration similarity, and
    temporal author-activity descriptors"

artifact F_network_structure : feature(D_network_snapshots)
  value structure: relational configuration
  value type: {reference, numeric}
  representation form: "weighted dynamic co-authorship graph descriptors"
  description: "Structural descriptors of author relationships, including
    edge strength, persistence, neighbourhood similarity, and temporal
    continuity between snapshots"

artifact F_author_centrality : feature(D_authors)
  value structure: vector
  value type: numeric
  description: "Per-author, per-time-window structural measures including
    weighted degree, betweenness, eigenvector centrality, brokerage,
    publication count, and collaboration persistence"

transform T_encode_documents :
  intent: characterise

```

```
manner: "text encoding and term-document representation"
input: D_papers, S_topic_model
output: F_paper_text
actor: machine
description: "Preprocess titles, abstracts, and keywords and create
    weighted document representations for topic discovery"
```

```
artifact F_paper_text : feature(D_papers)
    value structure: matrix
    value type: numeric
    representation form: "weighted term-document matrix"
    description: "Textual representation of each paper based on title,
        abstract, and keywords"
```

```
# =====
# STEP 3: ITERATIVE MODEL CALIBRATION AND COMMUNITY DISCOVERY
# =====
```

```
loop L_model_refinement:
    purpose: "Refine dynamic overlapping community and topic models until
        communities are structurally coherent, temporally plausible, and
        interpretable, while topics are semantically coherent and useful for
        characterising community research foci"
    until: "Community boundaries, overlaps, cores, lineage events, and top
        descriptions are judged interpretable, stable enough for the research
        questions, and supported by visual and domain evidence"
    body:
```

```
        transform T_build_community_model :
            intent: build-model
            manner: "fit dynamic overlapping community model"
            input: D_network_snapshots, F_network_structure, S_community_model
            output: M_community_model
            actor: machine
            description: "Fit a temporally regularised overlapping-community model
                so authors may have multiple membership strengths while recurring
                collaboration structure is encouraged to persist across time"
```

```
        artifact M_community_model : model(D_network_snapshots,
            F_network_structure)
            model type: "dynamic overlapping community model"
            representation form: "time-indexed author-community membership matrix"
```

with temporal regularisation and community lineage correspondence  
description: "Model of evolving collaboration communities allowing  
partial and simultaneous membership in multiple communities"

transform T\_infer\_communities :  
  intent: define-unit  
  manner: "infer overlapping dynamic communities"  
  input: D\_network\_snapshots, M\_community\_model, S\_community\_model  
  output: D\_communities, F\_membership\_strength, F\_lineage\_links  
  actor: machine  
  description: "Infer communities in each time period, author member  
    strengths, and correspondence links between communities in adjacent  
    periods"

artifact D\_communities : entities  
  internal structure: group/cluster  
  embedment: time  
  features:  
    - id: f\_community\_id  
      value structure: atomic  
      value type: categorical  
      description: "Persistent or period-specific community identifier"  
    - id: f\_member\_authors  
      value structure: list  
      value type: reference  
      description: "Authors associated with the community"  
    - id: f\_time\_period  
      value structure: atomic  
      value type: temporal  
      description: "Snapshot window in which the community exists"  
  description: "Time-indexed collaboration communities whose members  
    may overlap"

artifact F\_membership\_strength : feature(D\_authors)  
  value structure: matrix  
  value type: numeric  
  description: "Author-by-community membership strengths over time;  
    authors may hold substantial membership in several communities"

artifact F\_lineage\_links : feature(D\_communities)  
  value structure: relational configuration  
  value type: numeric

representation form: "weighted community correspondence graph"  
description: "Continuity, split, merge, emergence, and disappearance  
links between communities in consecutive temporal windows"

transform T\_build\_topic\_model :

intent: build-model  
manner: "fit topic model"  
input: F\_paper\_text, S\_topic\_model  
output: M\_topic\_model  
actor: machine  
description: "Build a topic model identifying research themes from  
paper titles, abstracts, and keywords"

artifact M\_topic\_model : model(F\_paper\_text)

model type: "topic model"  
representation form: "topic-term and paper-topic weight matrices"  
description: "Model of latent research topics in IEEE VIS publicat

transform T\_characterise\_topics :

intent: characterise  
manner: "compute paper-topic, community-topic, and temporal topic  
input: D\_papers, D\_communities, F\_membership\_strength, M\_topic\_mod  
output: F\_paper\_topics, F\_community\_topics, F\_topic\_trends  
actor: machine  
description: "Assign topic weights to papers and aggregate them by  
community and time, weighted by authors' community memberships"

artifact F\_paper\_topics : feature(D\_papers)

value structure: vector  
value type: numeric  
description: "Topic-weight vector for each paper"

artifact F\_community\_topics : feature(D\_communities)

value structure: vector  
value type: numeric  
description: "Topic distribution for each community in each tempor  
window, calculated from associated papers and membership strengt

artifact F\_topic\_trends : feature(D\_communities)

value structure: matrix  
value type: numeric  
description: "Temporal topic-prevalence trajectories for each

community"

```
transform T_characterise_community_roles :  
  intent: characterise  
  manner: "derive cores, leaders, bridges, overlap, and stability measures"  
  input: D_communities, F_membership_strength, F_author_centrality,  
         F_lineage_links, F_community_topics  
  output: F_community_roles, F_community_stability, F_overlap  
  actor: machine  
  description: "Compute persistent core membership, leadership rank, bridge roles, community retention, topic stability, and pairwise community overlap"  
  
artifact F_community_roles : feature(D_authors)  
  value structure: vector  
  value type: numeric  
  description: "Per-author community-role measures: persistent core score, within-community centrality, leadership rank, bridge score, number of concurrent communities, and membership trajectory"  
  
artifact F_community_stability : feature(D_communities)  
  value structure: vector  
  value type: numeric  
  description: "Community-level retention, member turnover, core persistence, cohesion, temporal continuity, size change, and topic-change measures"  
  
artifact F_overlap : feature(D_communities)  
  value structure: matrix  
  value type: numeric  
  description: "Pairwise community overlap through shared or multi-membership authors, calculated per time period"  
  
transform T_arrange_community_projection :  
  intent: contextualise  
  manner: "projection-based community and author similarity layout"  
  input: D_communities, F_membership_strength, F_community_topics, F_overlap  
  output: D_community_projection_space, A_community_projection  
  actor: machine  
  description: "Create a 2D similarity space where proximity reflects similarity in collaboration membership, topical focus, and overlap"
```

```

artifact D_community_projection_space : entities
  internal structure: elementary
  features:
    - id: f_dimensions
      value structure: atomic
      value type: numeric
      description: "Two projection dimensions"
  description: "Computed 2D similarity space for communities and aut

artifact A_community_projection : arrangement(D_communities)
  context: D_community_projection_space
  principle: "projection preserving similarity in membership, overla
    and topic-profile structure"
  description: "2D arrangement exposing related, overlapping, or
    separated communities"

transform T_arrange_evolution :
  intent: contextualise
  manner: "time-based lineage arrangement"
  input: D_communities, F_lineage_links, F_community_stability
  output: A_evolution
  actor: machine
  description: "Arrange communities by time and connect them through
    weighted lineage links to expose continuity, splits, merges,
    emergence, and dissolution"

artifact A_evolution : arrangement(D_communities)
  context: D_network_snapshots
  principle: "horizontal time axis; vertically ordered communities;
    ribbon links encode membership continuity and lineage relations"
  description: "Temporal lineage arrangement of evolving overlapping
    communities"

transform T_visualise_model_results :
  intent: visualise
  manner: "coordinated community, evolution, membership, and topic v
  input: A_community_projection, A_evolution, D_communities,
    F_membership_strength, F_community_roles,
    F_community_stability, F_community_topics, F_topic_trends,
    F_overlap
  output: V_community_overview

```

```

actor: machine
description: "Display a coordinated dashboard: community projection
temporal lineage ribbons, membership-overlap matrix, leader/core
ranking, stability plots, and topic evolution profiles"

artifact V_community_overview : visualisation(A_community_projection
A_evolution, F_community_roles, F_community_stability,
F_community_topics, F_overlap)
layout: "coordinated dashboard with similarity projection, temporal
alluvial lineage view, overlap matrix, ranked role table, and
topic-profile panels"
form: "points and hulls, alluvial ribbons, matrices, ranked bars,
line charts, and topic heatmaps"
encoding: "community position from projection; ribbon width from
retained membership; ribbon branching/convergence from split/merge
colour from persistent community identity; node size from active
member count; rank bars from core and leadership scores; heatmap
intensity from topic prevalence; matrix intensity from overlap"
description: "Interactive visualisation for evaluating communities
membership overlap, core stability, leadership, topical focus,
and temporal change"

transform T_interpret_model_results :
intent: abstract
manner: "perception-and-interpretation of community structure"
input: V_community_overview, K_domain_context
output: P_community_structure, P_topic_structure
actor: human
description: "Interpret candidate communities, their stable cores,
bridge authors, topical identities, suspicious model artifacts,
and candidate merge or split events"

artifact P_community_structure : pattern(D_communities,
F_community_stability)
representation form: "interpreted community and evolution description"
description: "Tentative descriptions of major communities, stable
cores, leaders, overlap structures, bridge authors, and lineage
events"

artifact P_topic_structure : pattern(F_community_topics,
F_topic_trends)
representation form: "named research-topic descriptions"

```

description: "Tentative semantic descriptions of recurring research foci and their evolution across communities"

transform T\_assess\_models :

intent: assess

manner: "evaluate structural validity, temporal plausibility, and

input: V\_community\_overview, P\_community\_structure, P\_topic\_struct  
S\_community\_model, S\_topic\_model, K\_domain\_context

output: K\_model\_assessment

actor: human

description: "Judge whether community boundaries, overlap levels, lineage events, core definitions, and topic descriptions are meaningful enough for substantive analysis"

artifact K\_model\_assessment : knowledge(D\_communities, M\_topic\_model  
representation form: "quality judgment"

description: "Assessment of community coherence, temporal stability, interpretability of overlap and leadership, and semantic coherence of research topics; identifies required model refinements"

if K\_model\_assessment indicates model refinement needed:

then:

transform T\_refine\_models :

intent: generate-knowledge

manner: "adjust temporal, community, and topic-model specifica

input: K\_model\_assessment, V\_community\_overview,  
S\_temporal\_network, S\_community\_model, S\_topic\_model

output: S\_temporal\_network\_r, S\_community\_model\_r,  
S\_topic\_model\_r

actor: human

description: "Revise time-window size, edge weighting, community overlap and persistence settings, lineage thresholds, topic count, or text preprocessing based on visual assessment"

artifact S\_temporal\_network\_r : specification

representation form: "revised temporal-network construction pa

description: "Updated snapshot resolution, window width, edge weighting, or inclusion criteria"

artifact S\_community\_model\_r : specification

representation form: "revised dynamic overlapping-community pa

```

        description: "Updated overlap, persistence, resolution, core,
            and lineage-matching settings"

    artifact S_topic_model_r : specification
        representation form: "revised topic-model parameters"
        description: "Updated topic count, preprocessing, or coherence
            settings"

    assign:
        S_temporal_network := S_temporal_network_r
        S_community_model := S_community_model_r
        S_topic_model := S_topic_model_r

    else:
        exit loop L_model_refinement

end loop L_model_refinement

# =====
# STEP 4: DETAILED EXPLORATION OF EVOLUTION EVENTS
# =====

artifact P_evolution_findings : pattern(D_communities)
    representation form: "accumulated community-evolution interpretations"
    description: "Initially empty; accumulates interpreted community
        transitions, causes, leadership changes, topical shifts, and evidence
        from selected papers"

loop L_evolution_investigation:
    purpose: "Investigate consequential community transitions and membership
        changes to explain how communities evolved and why merges, splits,
        emergence, or dissolution occurred"
    until: "Major communities, their cores and leaders, membership overlap
        topical trajectories, and consequential evolution events have been
        sufficiently explained for the analyst's research questions"
    body:

        transform T_identify_transition :
            intent: abstract
            manner: "identify salient evolution events and anomalous changes"
            input: V_community_overview, P_community_structure,
                P_topic_structure, P_evolution_findings

```

```

output: P_transition_candidates
actor: human
description: "Identify a significant merge, split, emergence,
            dissolution, leadership transition, unusually high overlap, or
            topical shift for detailed investigation"

artifact P_transition_candidates : pattern(D_communities,
            F_lineage_links)
representation form: "candidate evolution-event descriptions"
description: "Candidate events requiring explanation, including the
            time interval, involved communities, and preliminary structural
            topical interpretation"

transform T_select_transition :
    intent: define-unit
    manner: "interactive selection of communities, authors, and time i
    input: P_transition_candidates, V_community_overview
    output: D_transition_focus
    actor: human
    description: "Select communities, involved authors, and time windo
            for detailed analysis of a particular evolution event"

artifact D_transition_focus : entities
    internal structure: group/cluster
    embedment: time
    description: "Focused subset of communities, authors, papers, and
            temporal windows implicated in a selected evolution event"

transform T_characterise_transition_evidence :
    intent: characterise
    manner: "derive membership-flow, publication, citation, and topic-
    input: D_transition_focus, D_papers, D_authors, D_communities,
            F_membership_strength, F_community_roles, F_community_topic
            F_topic_trends, F_lineage_links
    output: F_transition_evidence
    actor: machine
    description: "Compute member inflow and outflow, retained core,
            leadership change, shared-author bridges, representative papers,
            citation links, and topic-distribution differences before and af
            the selected event"

artifact F_transition_evidence : feature(D_transition_focus)

```

```

value structure: vector
value type: {numeric, categorical, temporal, reference}
description: "Evidence for explaining a selected event: author flow
    core retention, leadership changes, topic shifts, representative
    publications, and citation relations"

transform T_visualise_transition :
  intent: visualise
  manner: "linked event-detail views"
  input: D_transition_focus, F_transition_evidence, D_papers,
        F_paper_topics, F_community_roles
  output: V_transition_detail
  actor: machine
  description: "Show selected communities as linked NodeTrix or
    node-link detail, author membership trajectories, before-after
    topic profiles, membership-flow ribbons, and representative-paper
    evidence"

artifact V_transition_detail : visualisation(D_transition_focus,
    F_transition_evidence, D_papers, F_paper_topics)
  layout: "coordinated event-detail views: community network, author
    membership trajectories, before-after topic profiles, flow ribbons
    and representative-paper list"
  form: "NodeTrix matrices and node-link links, alluvial ribbons,
    ranked lists, line charts, topic bars, and paper cards"
  encoding: "matrix cells and links from co-authorship strength;
    author line thickness from membership strength; colour from
    community identity; topic bar length from prevalence; paper cards
    show title, keywords, abstract excerpt, citations, and authors"
  description: "Detailed visual evidence supporting explanation of why
    a community changed"

transform T_interpret_transition :
  intent: abstract
  manner: "explain merge, split, dissolution, or topic shift"
  input: V_transition_detail, K_domain_context, P_evolution_findings
  output: P_evolution_findings_r
  actor: human
  description: "Interpret the selected event using structural and
    topical evidence; distinguish substantive reorientation from a
    methodological artifact or short-lived collaboration fluctuation"

```

```

artifact P_evolution_findings_r : pattern(D_communities, D_papers)
  representation form: "accumulated evidence-grounded evolution expl
  description: "Accumulated explanations of community emergence,
    persistence, merging, splitting, dissolution, leadership change,
    bridge roles, and changing research foci"

transform T_assess_evolution_coverage :
  intent: assess
  manner: "evaluate completeness of community-evolution understandin
  input: P_evolution_findings_r, V_community_overview,
    V_transition_detail
  output: K_evolution_assessment
  actor: human
  description: "Judge whether enough major communities and evolution
    events have been examined, or whether another event, period, or
    community requires investigation"

artifact K_evolution_assessment : knowledge(P_evolution_findings_r)
  representation form: "quality judgment"
  description: "Assessment of whether the analyst has sufficient
    evidence to answer questions about community count, stable cores
    leaders, overlap, research foci, and evolution"

if K_evolution_assessment indicates sufficient understanding:
  then:
    assign:
      P_evolution_findings := P_evolution_findings_r
    exit loop L_evolution_investigation
  else:
    assign:
      P_evolution_findings := P_evolution_findings_r

end loop L_evolution_investigation

# =====
# STEP 5: SYNTHESIS
# =====

transform T_synthesise_findings :
  intent: generate-knowledge
  manner: "formulate statements and evidence-backed community profiles"
  input: P_community_structure, P_topic_structure, P_evolution_findings,

```

F\_community\_stability, F\_community\_roles, F\_overlap,  
K\_model\_assessment, K\_evolution\_assessment  
output: K\_community\_knowledge  
actor: human  
description: "Produce an evidence-backed account of the intellectual  
structure and evolution of the IEEE VIS community"

artifact K\_community\_knowledge : knowledge(P\_community\_structure,  
P\_topic\_structure, P\_evolution\_findings)  
representation form: "community profiles, evolution narratives, and  
methodological provenance"  
description: "Final analytical knowledge: number and identity of major  
collaborating communities; stable cores and leaders; bridge authors  
overlapping memberships; research foci and their temporal evolution;  
merges, splits, emergence, and dissolution events with supporting pa  
citation, membership, and topic evidence; plus documented modeling  
choices and assessment rationale"