

KH Coder で国会議事録の テキストを分析してみる

情報探索および情報編集スキルを有する方々が
コードを書かなくても/書けなくても
ハッカソンの場に参加/活躍する方策の試作として



高橋 陽一

アーバンデータチャレンジ実行委員 (2013～)
Wikipediaタウン 二子玉川 (2013～)
GeoNames.jp (2015～)
LODチャレンジ実行委員 (2011～2015)

本編集著作物は クリエイティブ・コモンズ 表示 4.0 国際 ライセンスの下
に提供されています。ただし、本編集著作物に含まれる各素材につい
ては、個別に指定がある場合は、各々のライセンスに従うものとします。



KH Coder で国会議事録のテキストを分析

- 設定した分析テーマ
 - 戦後の「国土計画」に関する国会討議の焦点の変遷を、国会議事録のテキスト分析(共起ネットワーク分析)から探ってみる
- 対象
 - キーワードとして「国土計画」を含む発言を対象に、各「全総」策定年月日から遡及して各2年分を抽出期間として設定
 - 共起ネットワーク分析設定では対象を「名詞、サ変名詞、形容動詞、組織名、タグ(国土計画)、動詞、形容詞」に絞った
- 使用したデータとツール
 - [国会会議録検索システム検索用API](#)
 - Excel2013 (上記APIを使用し分析対象データを取得)
 - [KH Coder Ver. 3.Alpha.4](#) (上記で取得した発言テキストを分析)

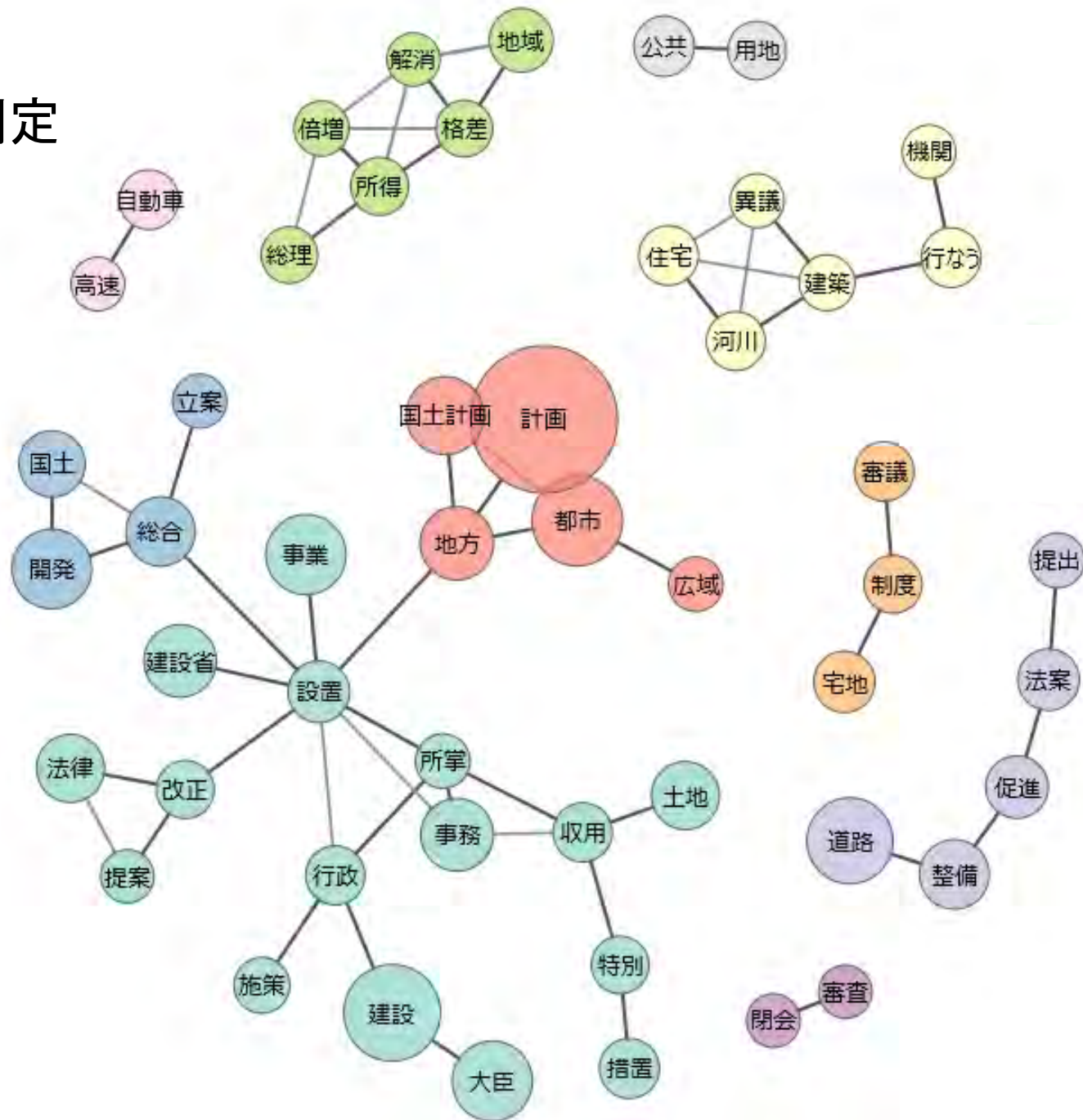
一全総

1962年10月5日制定

1. 高度成長経済への移行
2. 過大都市問題、所得格差の拡大
3. 所得倍增計画(太平洋ベルト地帯構想)

地域間の均衡ある発展

- 都市の過大化による生産面・生活面の諸問題、地域による生産性の格差について、国民経済的視点からの総合的解決を図る。



1962

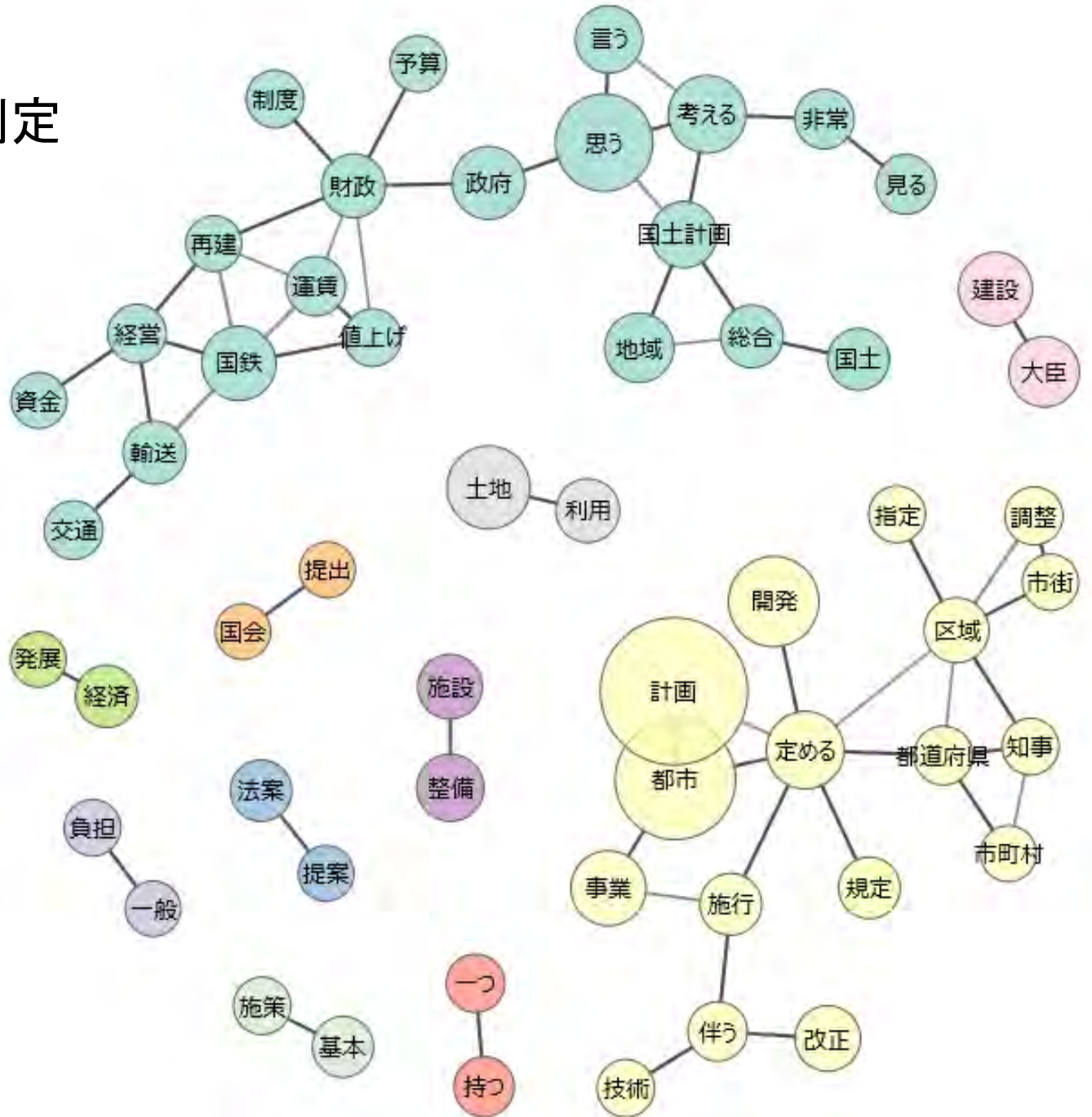
新全総

1969年5月30日制定

1. 高度成長経済
2. 人口、産業の大都市集中
3. 情報化、国際化、技術革新の進展

豊かな環境の創造

- 基本的課題を調和しつつ、高福祉社会を目指して人間のための豊かな環境を創造する。



1969

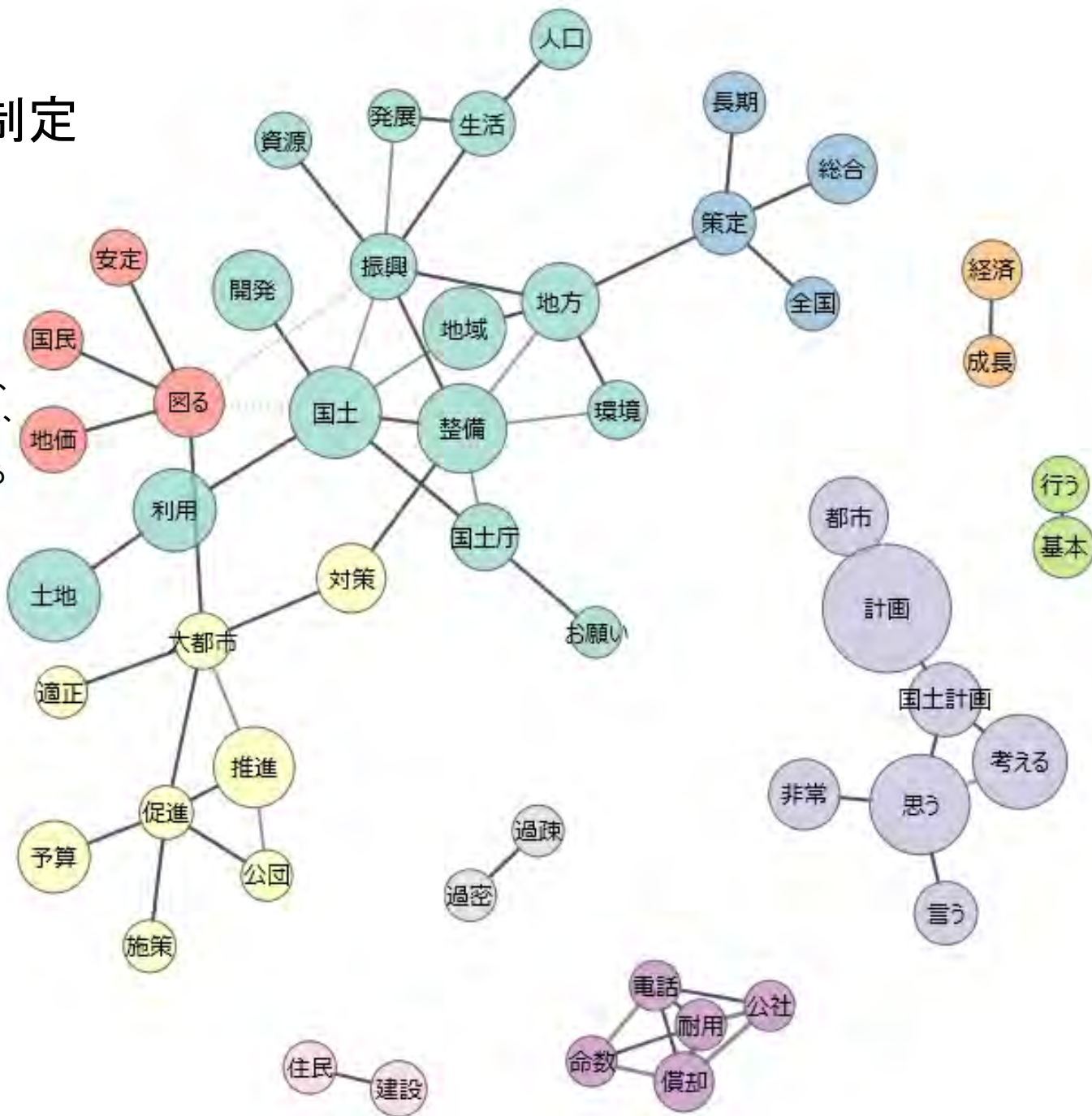
三全総

1977年11月4日制定

1. 安定成長経済
2. 人口、産業の地方分散の兆し
3. 国土資源、エネルギー等の有限性の顕在化

人間居住の総合的環境の整備

- 限られた国土資源を前提として、地域特性を生かしつつ、歴史的、伝統的文化に根ざし、人間と自然との調和のとれた安定感のある健康で文化的な人間居住の総合的環境を計画的に整備する。



1977

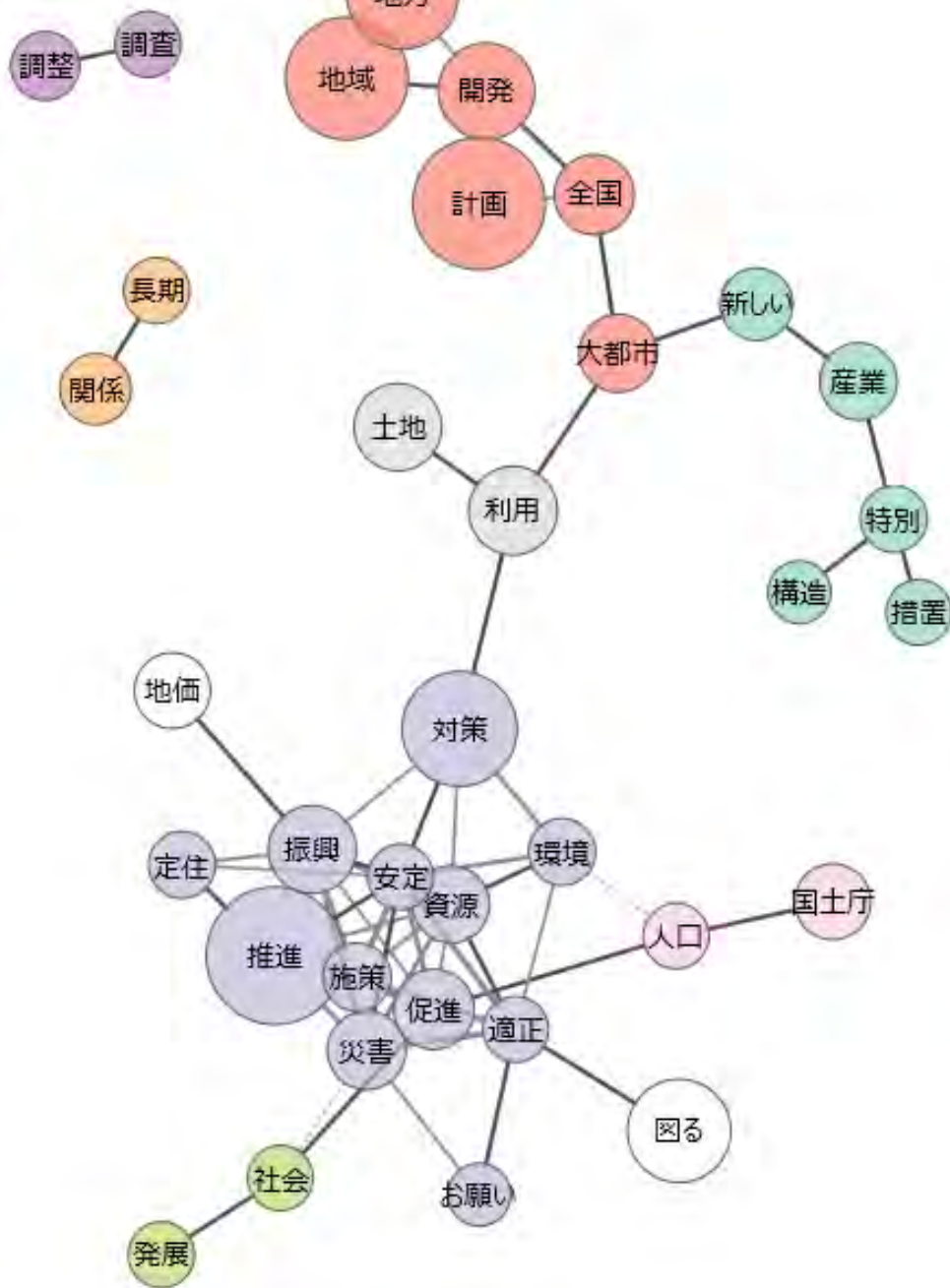
四全総

1987年6月30日制定

1. 人口、諸機能の東京一極集中
2. 産業構造の急速な変化等により、地方圏での雇用問題の深刻化
3. 本格的国際化の進展

多極分散型国土の構築

- 安全でうるおいのある国土の上に、特色ある機能を有する多くの極が成立し、特定の地域への人口や経済機能、行政機能等諸機能の過度の集中がなく地域間、国際間で相互に補完、触発しあいながら交流している国土を形成する。



1987

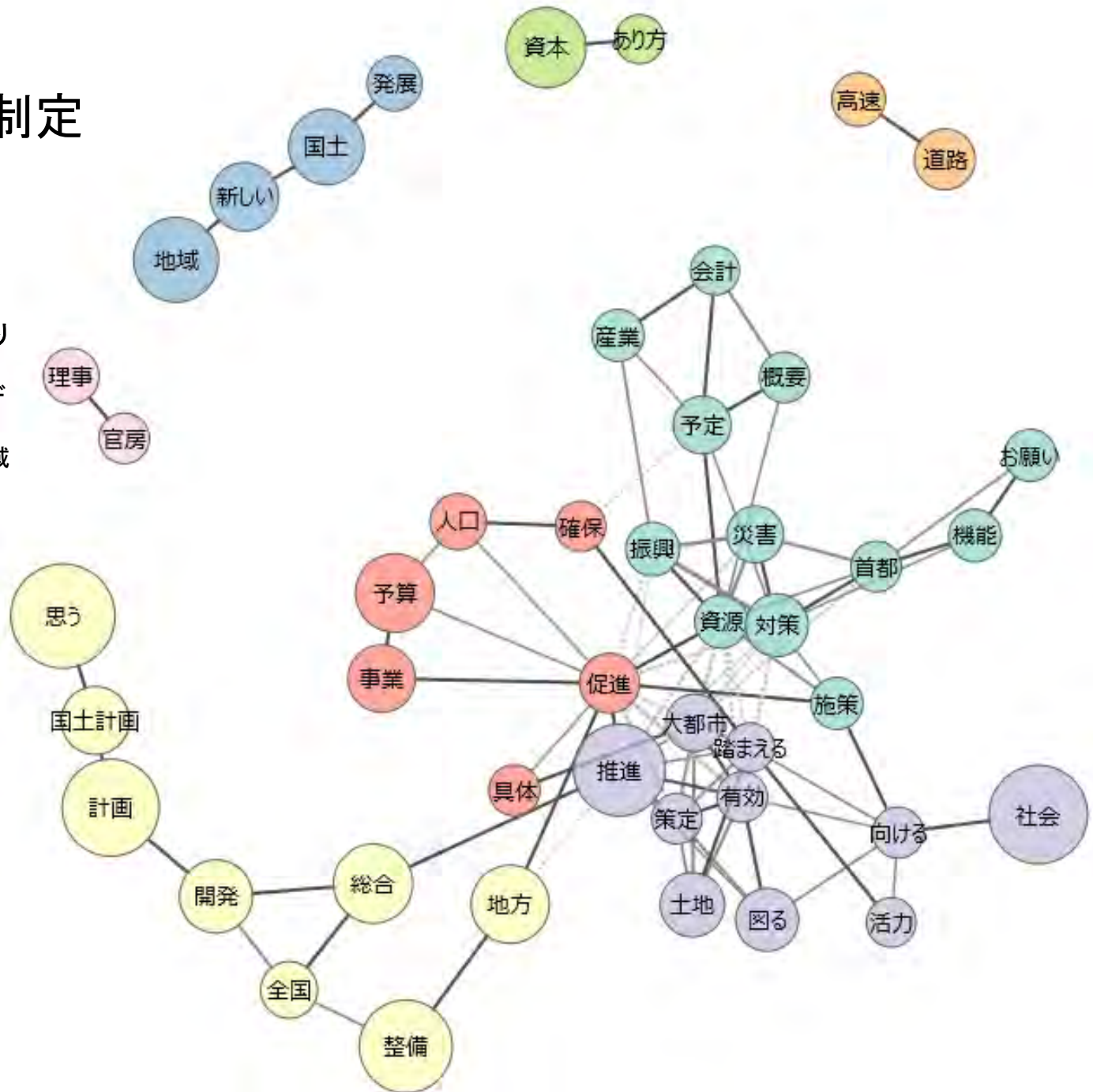
五全総

1998年3月31日制定

1. 地球時代(地球環境問題、大競争、アジア諸国との交流)
2. 人口減少・高齢化時代
3. 高度情報化時代

多軸型国土構造形成の基礎づくり

- 多軸型国土構造の形成を目指す「21世紀の国土のグランドデザイン」実現の基礎を築く。
- 地域の選択と責任に基づく地域づくりの重視



1998

ひとまずの「まとめ」・・・

- ツールを駆使することで「コードを1行も書かず」にこのような分析を短時間(正味4時間未満)で行うことができたが・・・
- 手早く分析するために発言に「国土計画」を含むレコードのみ対象にしてしまったが、本来はその発言の前後も含めて(=発言のあった会議全体を再参照して)分析する必要があった
- もっと分析対象に即した「適切な検索キーワード」や「抽出期間」を設計した方が有意な結果が出せそうな気配が・・・
- でも、特定テーマの時系列変遷を国会議事録で追うのは興味深いので、モクモクと習作を続けてみたい