Title: Can science be wikified

Date: Mar 26, 2018 02:45 PM

URL: http://pirsa.org/18030102

Abstract:

TBD

Ivar Martin

Can science be wikified? (and should it)

Ivar Martin

Knowen.org

Outline

- Does wikification makes sense?

Outline

- Does wikification makes sense?
- How do we get there?

Wikipedia

Everyone can read and edit

- 18B page views and 500M unique visitors/ month (Alexa rank 5)
- Number of articles: 5.5M (English only), 40K "high quality"
- 33M registered editors, 140K active editors
- 50% edits are done by 500 people, 0.7% (2009)

Source: Wikipedia

The Future of Science
by Michael Nielsen on July 17, 2008

A failure of science online: Wikipedia

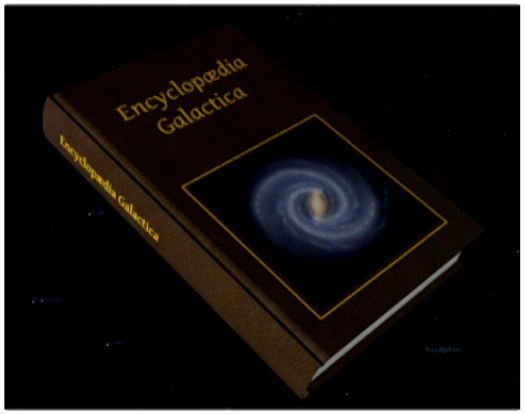
Wikipedia is a second example where scientists have missed an opportunity to innovate online. Wikipedia has a vision statement to warm a scientist's heart: "Imagine a world in which every single human being can freely share in the sum of all knowledge. That's our commitment."

...

Nowadays, Wikipedia's success has to some extent legitimized contribution within the scientific community. But how strange that the modern day Library of Alexandria had to come from outside academia.

Wikipedia appeal

- Universal − central − repository
- Generally good high-level articles
- Active participation (billions of users, thousands of contributors)
- As close as can get to



If Wikipedia ***were used*** for science

- Potential benefits
  - **Universal** resource, instead of zillions of journals
  - Existing **context** for new results
  - Fairer **attribution**
- Potential problems
  - Public quarrels
  - Fractionalized content (“hard forks”)
  - …

Why scientist **don't** use Wikipedia

Why scientist **don't** use Wikipedia

- Mixed **quality** of material

Why scientist **don't** use Wikipedia

- Mixed **quality** of material
- No tangible **credit** to contributor

Why scientist **don't** use Wikipedia

- Mixed **quality** of material
- No tangible **credit** to contributor
- Not a place for **original** material

Why scientist **don't** use Wikipedia

- Mixed **quality** of material
- No tangible **credit** to contributor
- Not a place for **original** material
- **Transience** (editing over)

Why scientist **don't** use Wikipedia

- Mixed **quality** of material
- No tangible **credit** to contributor
- Not a place for **original** material
- **Transience** (editing over)
- **Consensus**: not (always) expertise based

Why scientist **don't** use Wikipedia

- Mixed **quality** of material
- No tangible **credit** to contributor
- Not a place for **original** material
- **Transience** (editing over)
- **Consensus**: not (always) expertise based
- Content intended for **non-experts**

Wikipedia-for-science requirements

Wikipedia-for-science requirements

- **Incentives**
  - Improve **visibility** of work
  - Enhance personal **productivity** and **quality** of work
    - Note taking, long-term preservation
  - Improve **collaborations**
  - Build **reputation**
    - Track/value all contributions

Wikipedia-for-science requirements

- **Incentives**
  - Improve **visibility** of work
  - Enhance personal **productivity** and **quality** of work
    - Note taking, long-term preservation
  - Improve **collaborations**
  - Build **reputation**
    - Track/value all contributions
- Access control
  - Public or Limited access

Wikipedia-for-science requirements

- **Incentives**
  - Improve **visibility** of work
  - Enhance personal **productivity** and **quality** of work
    - Note taking, long-term preservation
  - Improve **collaborations**
  - Build **reputation**
    - Track/value all contributions
- Access control
  - Public or Limited access
- **Structure**
  - Levels of **refinement**

Wikipedia-for-science requirements

- **Incentives**
  - Improve **visibility** of work
  - Enhance personal **productivity** and **quality** of work
    - Note taking, long-term preservation
  - Improve **collaborations**
  - Build **reputation**
    - Track/value all contributions
- Access control
  - Public or Limited access
- **Structure**
  - Levels of **refinement**
- AI
  - Suggest connections to existing content
  - Analytics tools
    - performance,
    - trends

Incentives

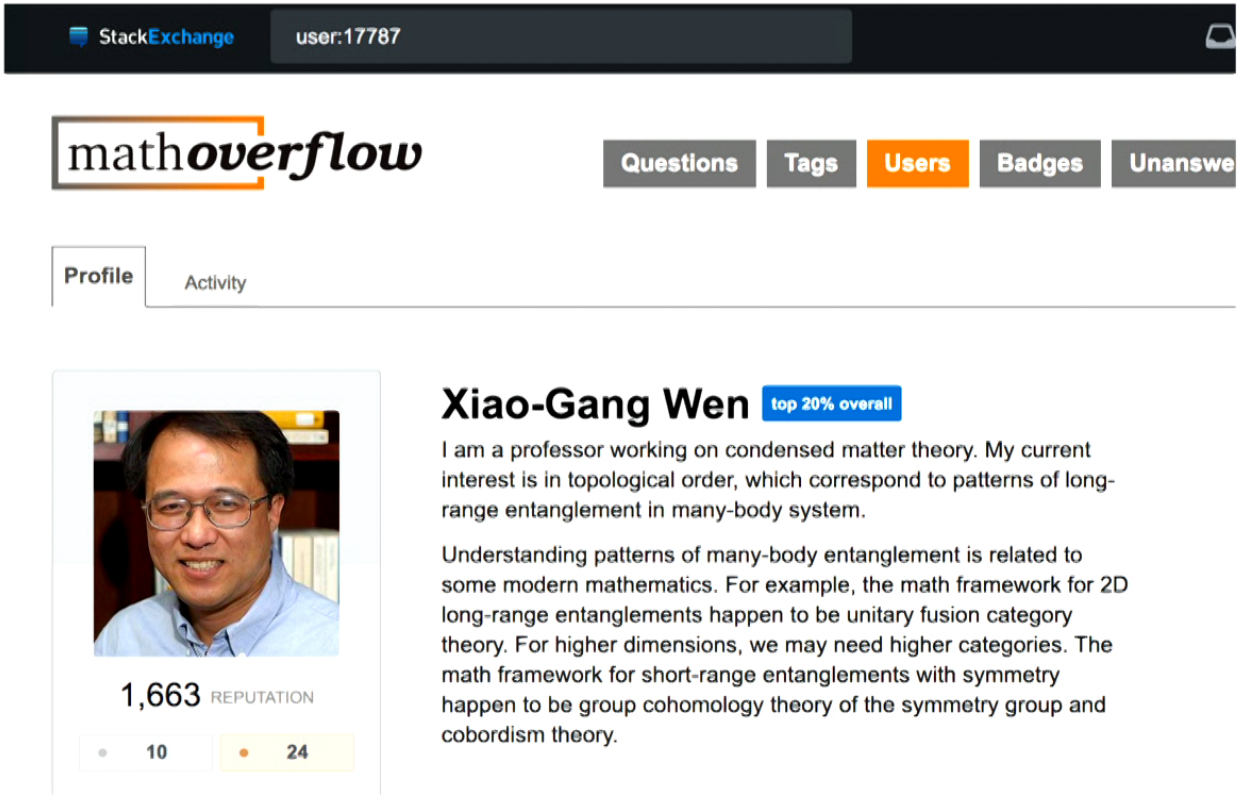
- StackExchange model
  - Reputation, badges, karma

Incentives

- StackExchange model
  - Reputation, badges, karma
- Blockchain model
  - Tokens (utility, work), convertible into $

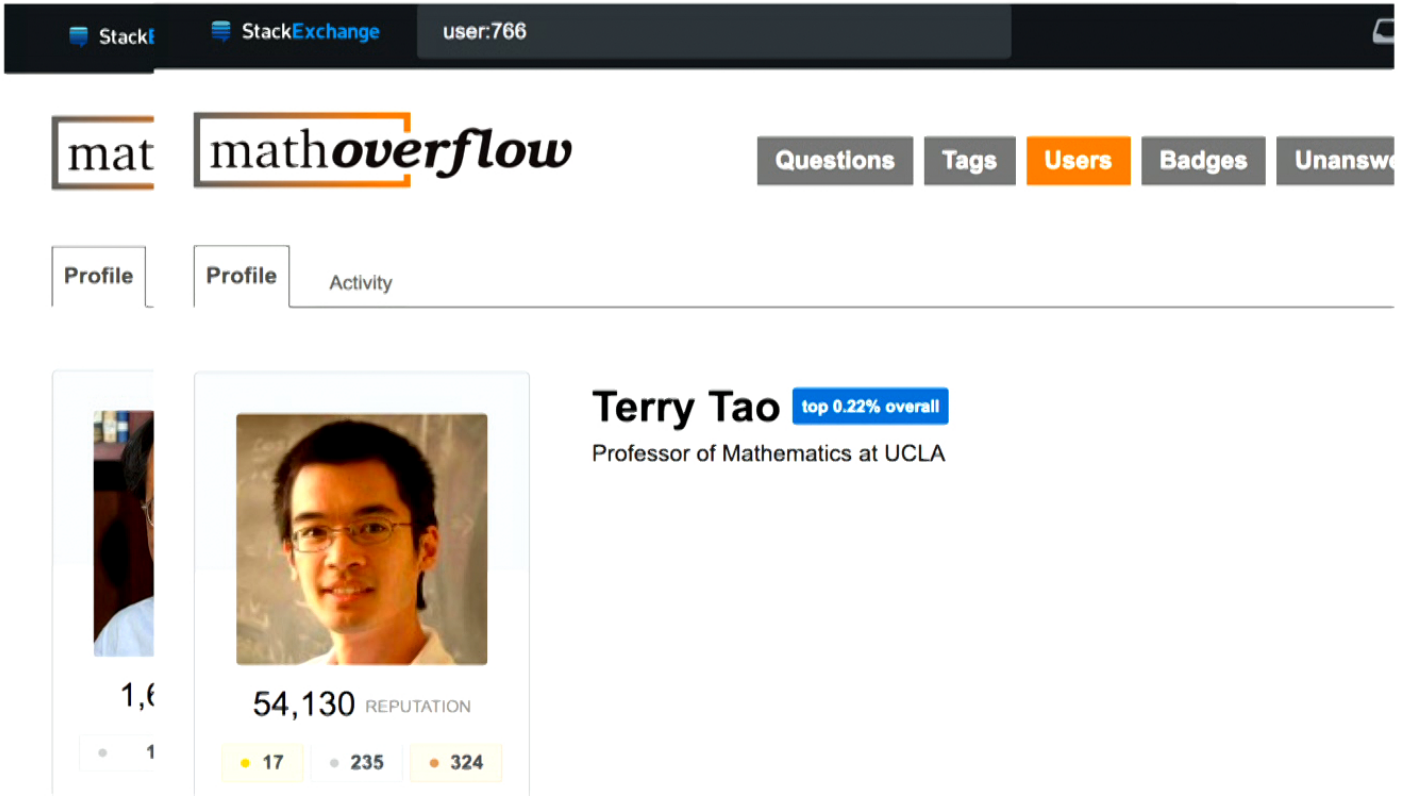
Incentives

- StackExchange model



Incentives

- StackExchange model



Structure

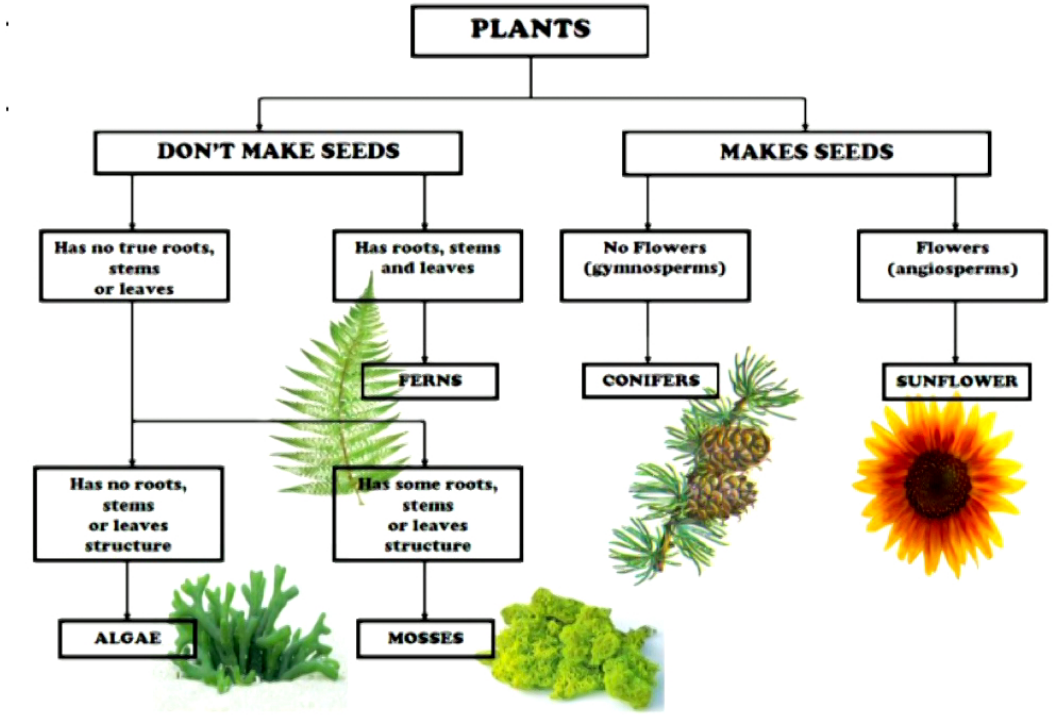
What is the structure of knowledge?

What is the structure of knowledge?

- Collections of facts -> connection of facts
  - Generalization/refinement
  - Efficient Compression (sparse recovery/ compressed sensing)

What is the structure of knowledge?

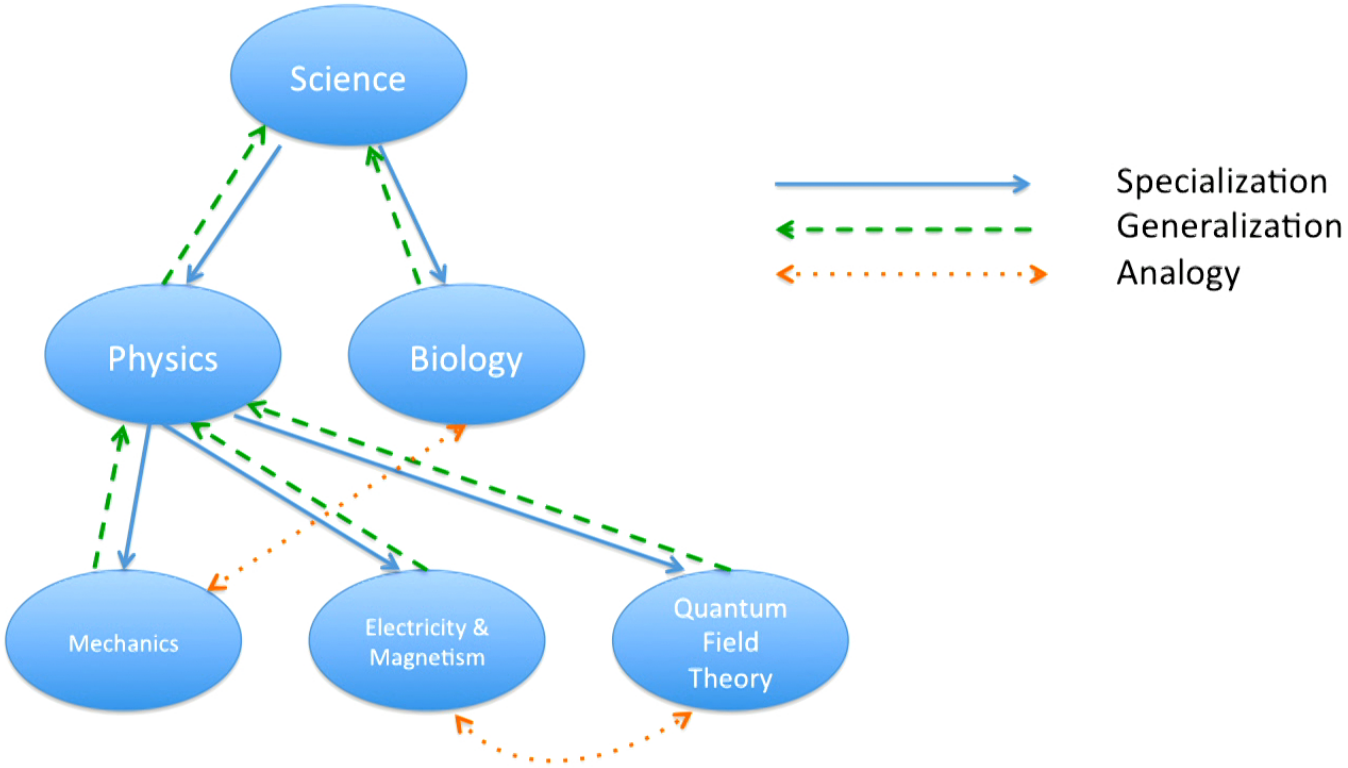
- Collections of facts -> connection of facts



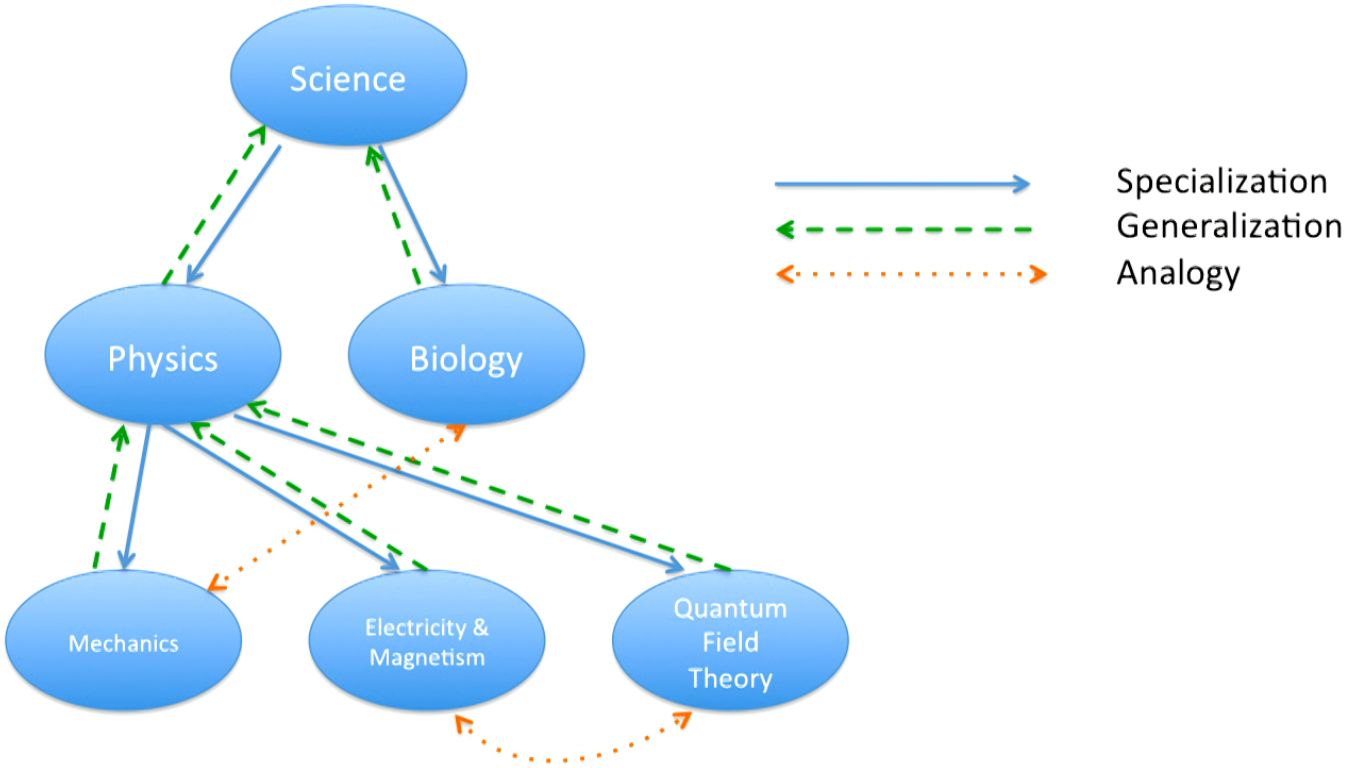
Knowledge hierarchies

- **Math**
  - Mathematical Subject Classification – MSC2010: https://zbmath.org/classification/
    Wolfram: http://mathworld.wolfram.com/
    NLab: https://ncatlab.org/nlab/show/mathematics
    Enc of Math: https://www.encyclopediaofmath.org/index.php/Talk:EoM:This_project#Categories
- **Physics:**
  - https://physh.aps.org/about
    PhySH: https://physh.aps.org/browse
    PACS: https://www.aip.org/publishing/pacs/pacs-2010-regular-edition
    NLab: https://ncatlab.org/nlab/show/physics
- **Medicine:**
  - https://www.nlm.nih.gov/pubs/factsheets/mesh.html
- **Computer science**
  - Computing classification system: http://dl.acm.org/ccs/ccs.cfm
    Computing research repository: http://arxiv.org/corr/subjectclasses
- **Economics**
  - Journal of Econ Lit: https://www.aeaweb.org/jel/guide/jel.php
- **General**
  - Dewey: https://en.wikipedia.org/wiki/List_of_Dewey_Decimal_classes
    Lib of Congress: https://www.loc.gov/catdir/cpso/lcco/

Is knowledge a tree?

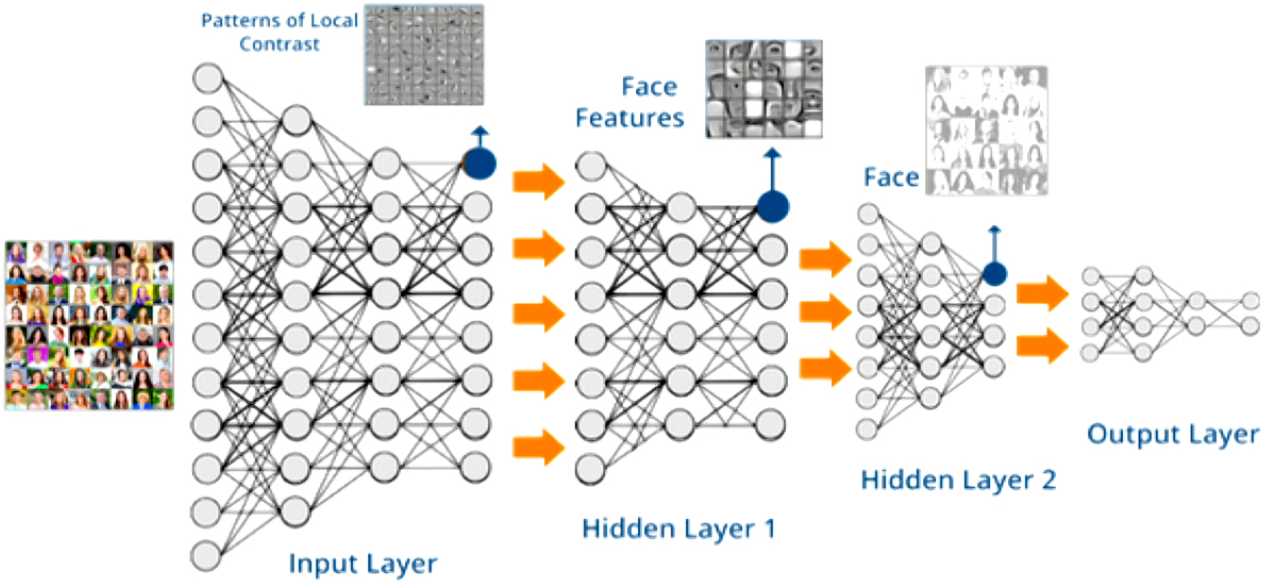


Is knowledge a tree?

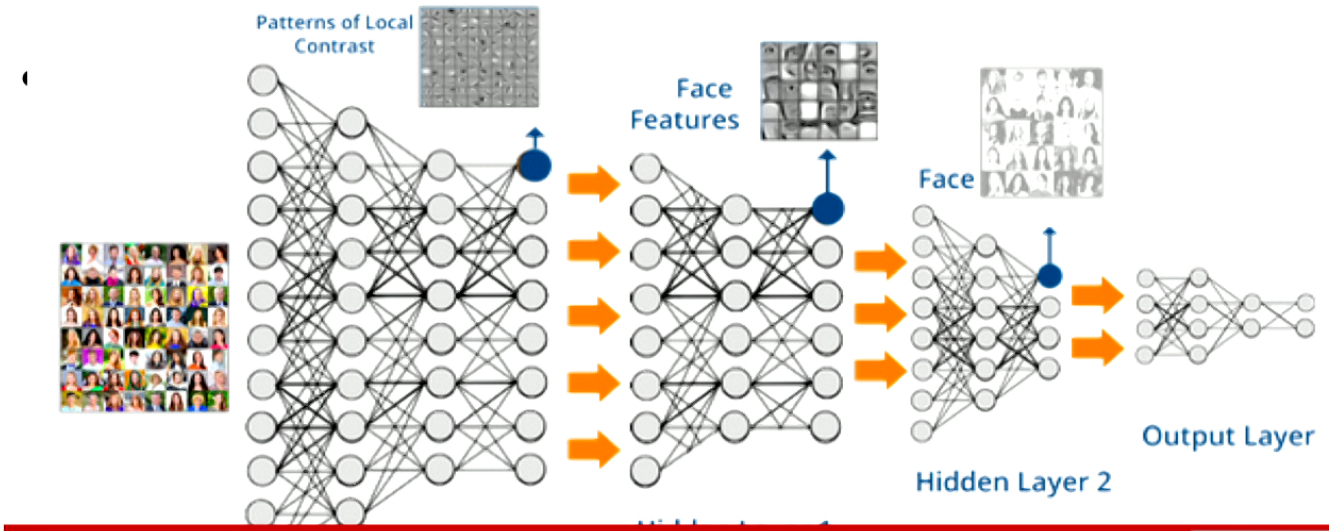


Can it be shaped into a tree?

Inspiration from RG/Machine learning



Inspiration from RG/Machine learning



arXiv.org > stat > arXiv:1410.3831

Search or Article II

(Help | Advanced searc

Statistics > Machine Learning

An exact mapping between the Variational Renormalization Group and Deep Learning

Pankaj Mehta, David J. Schwab

(Submitted on 14 Oct 2014)

Inspiration

THE ART OF COMPUTER PROGRAMMING

THIRD EDITION

CONTENTS

Chapter 1 — Basic Concepts . . . 1

1.1. Algorithms . . . 1
1.2. Mathematical Preliminaries . . . 10
1.2.1. Mathematical Induction . . . 11
1.2.2. Numbers, Powers, and Logarithms . . . 21
1.2.3. Sums and Products . . . 27
1.2.4. Integer Functions and Elementary Number Theory . . . 39
1.2.5. Permutations and Factorials . . . 45
1.2.6. Binomial Coefficients . . . 52
1.2.7. Harmonic Numbers . . . 75
1.2.8 Fibonacci Numbers . . . 79
1.2.9 Generating Functions . . . 87
1.2.10 Analysis of an Algorithm . . . 96
*1.2.11 Asymptotic Representations . . . 107
*1.2.11.1 The O-notation . . . 107
*1.2.11.2 Euler's summation formula . . . 111
*1.2.11.3 Some asymptotic calculations . . . 116
1.3 MIX . . . 124
1.3.1. Description of MIX . . . 124
1.3.2. The MIX Assembly Language . . . 144
1.3.3. Applications to Permutations . . . 164
1.4. Some Fundamental Programming Techniques . . . 186
1.4.1. Subroutines . . . 186
1.4.2. Coroutines . . . 193
1.4.3. Interpretive Routines . . . 200
1.4.3.1. A MIX simulator . . . 202
*1.4.3.2. Trace routines . . . 212
1.4.4. Input and Output . . . 215
1.4.5. History and Bibliography . . . 229

Chapter 2 — Information Structures . . . 232

2.1. Introduction . . . 232
2.2. Linear Lists . . . 238
2.2.1. Stacks, Queues, and Deques . . . 238
2.2.2. Sequential Allocation . . . 244
2.2.3. Linked Allocation . . . 254
2.2.4. Circular Lists . . . 273
2.2.5. Doubly Linked Lists . . . 280
2.2.6. Arrays and Orthogonal Lists . . . 298
2.3. Trees . . . 308
2.3.1. Traversing Binary Trees . . . 318
2.3.2. Binary Tree Representation of Trees . . . 334
2.3.3. Other Representations of Trees . . . 348
2.3.4. Basic Mathematical Properties of Trees . . . 362
2.3.4.1. Free trees . . . 363
2.3.4.2. Oriented trees . . . 372
*2.3.4.3. The "infinity lemma" . . . 382
*2.3.4.4. Enumeration of trees . . . 386
2.3.4.5. Path length . . . 399
*2.3.4.6. History and bibliography . . . 406
2.3.5. Lists and Garbage Collection . . . 408
2.4. Multilinked Structures . . . 424
2.5. Dynamic Storage Allocation . . . 435
2.6. History and Bibliography . . . 457

Answers to Exercises . . . 466

Appendix A — Tables of Numerical Quantities . . . 619

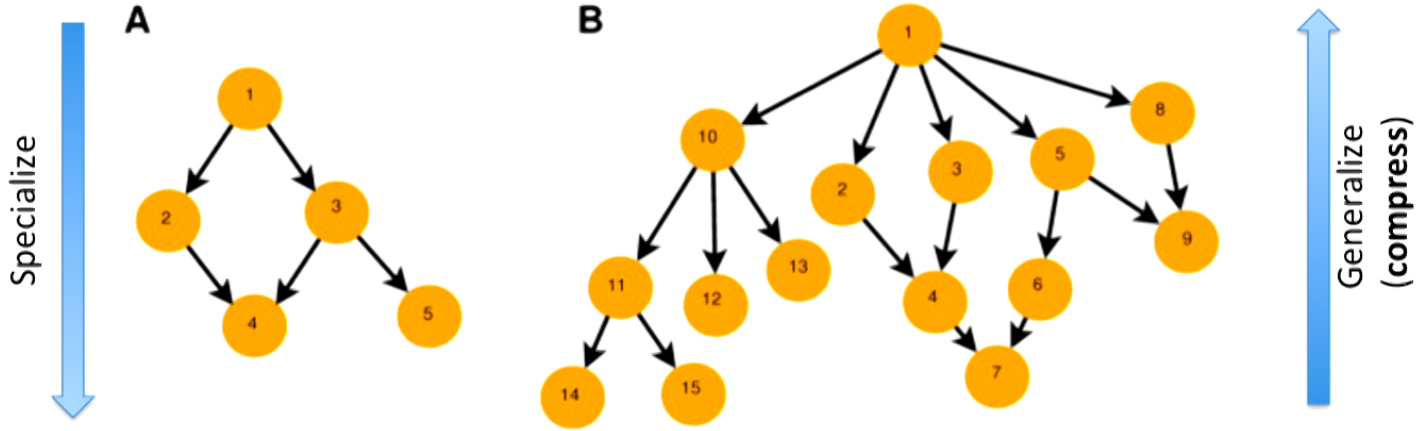
1. Fundamental Constants (decimal) . . . 619
2. Fundamental Constants (octal) . . . 620
3. Harmonic Numbers, Bernoulli Numbers, Fibonacci Numbers . . . 621

Appendix B — Index to Notations . . . 623

Index and Glossary . . . 628

Why Tree or DAG structure is good

- depth is ~log N
- Easy to navigate
- Easy to learn
- “Plug-and-Play”



How to get there?

How to get there?

Chicken or Egg?

How to get there?

- Seed from existing content
  - Wikipedia
  - Published articles
  - Taxonomies/classification schemes

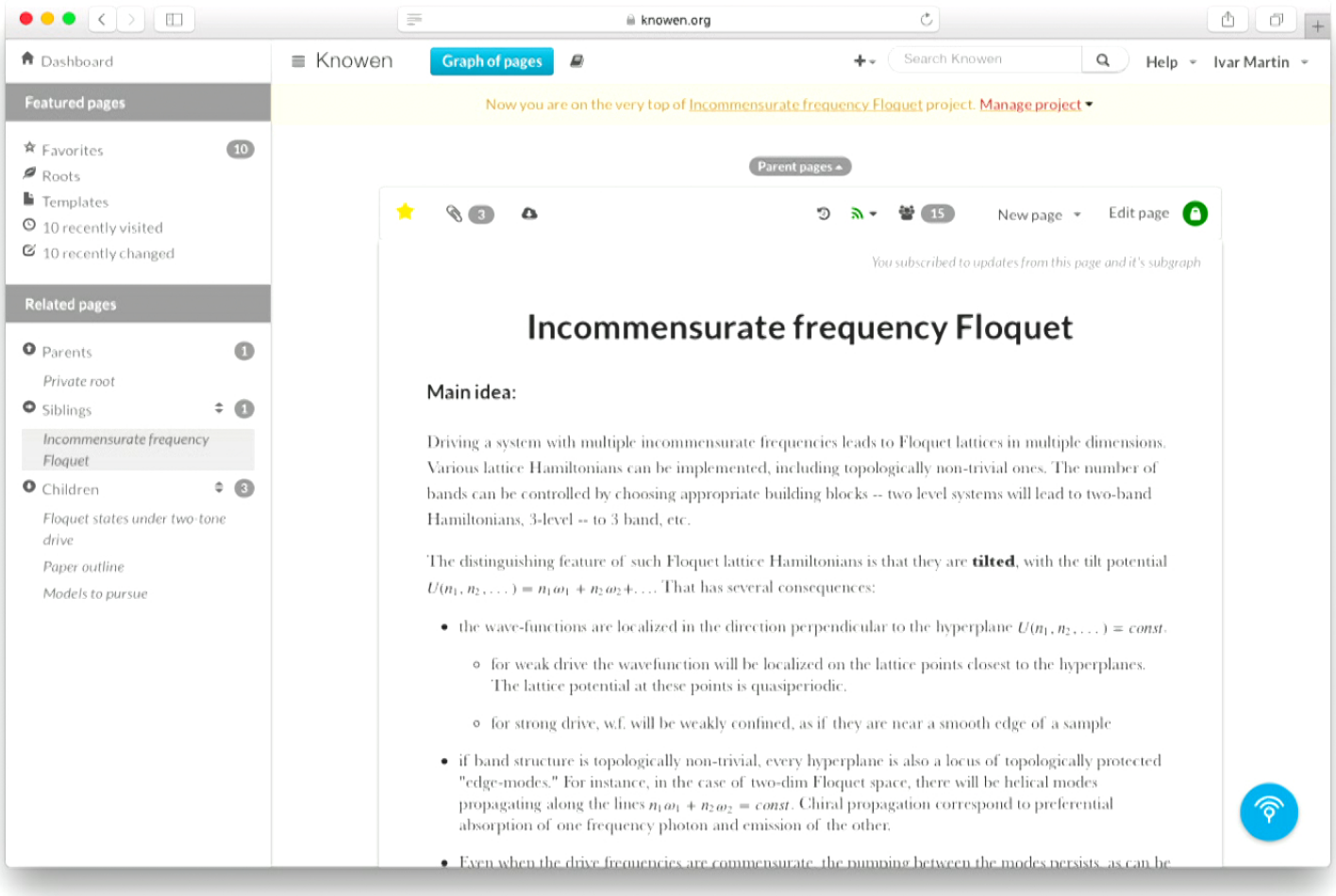
Chicken or Egg?

How to get there?

- Seed from existing content
  - Wikipedia
  - Published articles
  - Taxonomies/classification schemes
- Organically, new collaborative projects
  - Bottom up
  - Knowen

Chicken or Egg?

Experiment: knowen.org

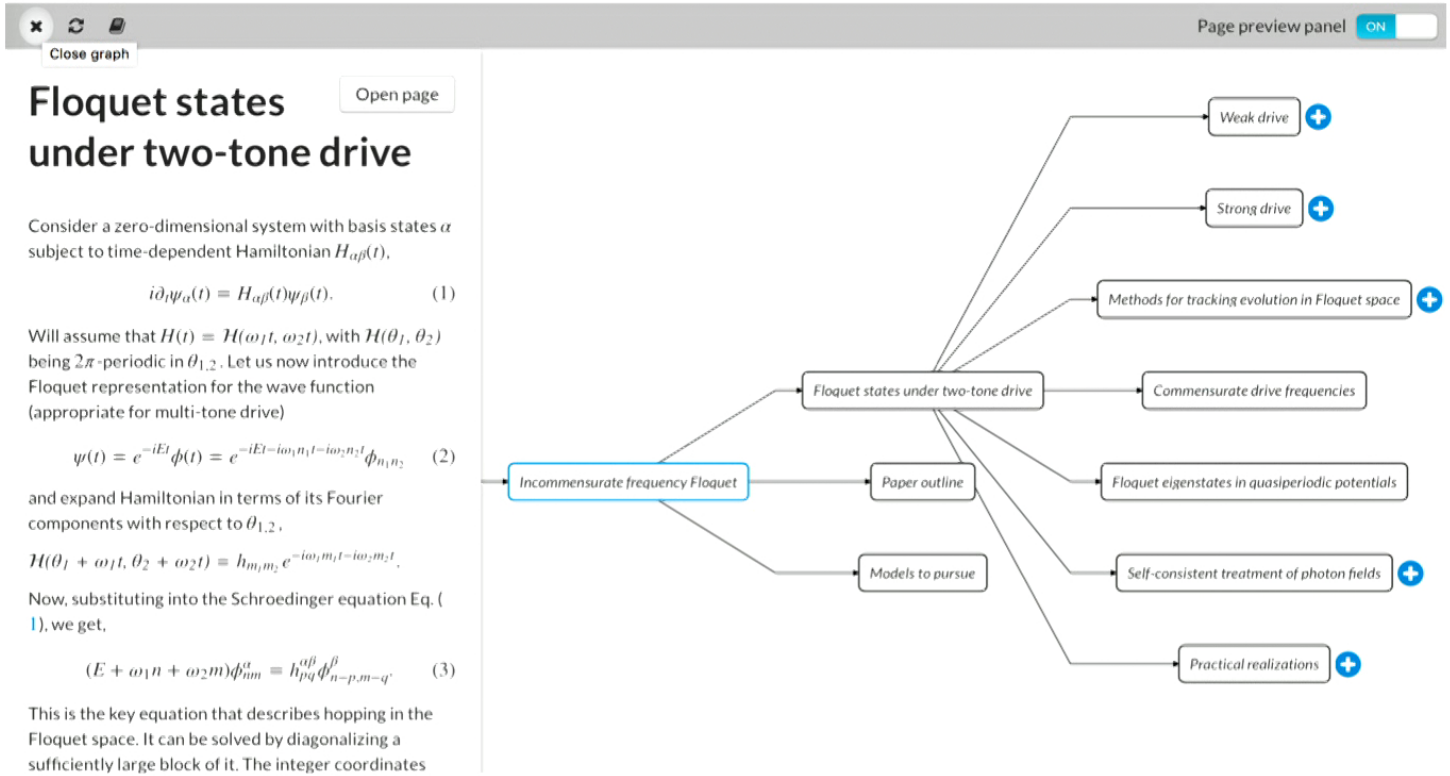


Elements

- Nodes: (Text, data, scripts)
- Commits and commenting (cf. github)
- Structure: Tree/Directed acyclic graph (DAG)
- Access: public and private projects
- Reputation tracking

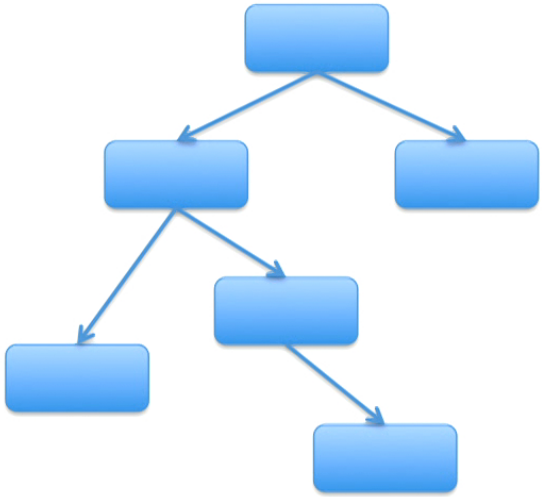
Also: collaborative editing, subscriptions, feedback mechanisms

Knowen.org

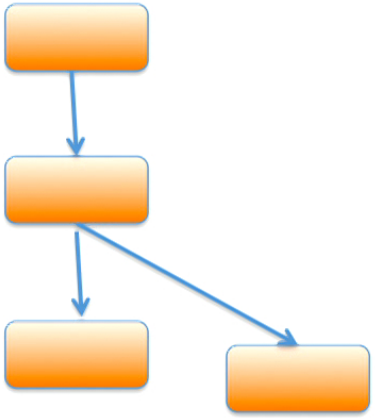

Close graph
Page preview panel
ON
Open page
Floquet states under two-tone drive
Consider a zero-dimensional system with basis states α subject to time-dependent Hamiltonian $H_{\alpha\beta}(t)$,
$$i\partial_t\psi_\alpha(t) = H_{\alpha\beta}(t)\psi_\beta(t). \quad (1)$$
Will assume that $H(t) = \mathcal{H}(\omega_1 t, \omega_2 t)$, with $\mathcal{H}(\theta_1, \theta_2)$ being 2π-periodic in $\theta_{1,2}$. Let us now introduce the Floquet representation for the wave function (appropriate for multi-tone drive)
$$\psi(t) = e^{-iEt}\phi(t) = e^{-iEt-i\omega_1 n_1 t-i\omega_2 n_2 t}\phi_{n_1 n_2} \quad (2)$$
and expand Hamiltonian in terms of its Fourier components with respect to $\theta_{1,2}$,
$$\mathcal{H}(\theta_1 + \omega_1 t, \theta_2 + \omega_2 t) = h_{m_1 m_2} e^{-i\omega_1 m_1 t-i\omega_2 m_2 t}.$$
Now, substituting into the Schroedinger equation Eq. (1), we get,
$$(E + \omega_1 n + \omega_2 m)\phi^\alpha_{nm} = h^{\alpha\beta}_{pq}\phi^\beta_{n-p,m-q}. \quad (3)$$
This is the key equation that describes hopping in the Floquet space. It can be solved by diagonalizing a sufficiently large block of it. The integer coordinates
Weak drive
Strong drive
Methods for tracking evolution in Floquet space
Floquet states under two-tone drive
Commensurate drive frequencies
Incommensurate frequency Floquet
Paper outline
Floquet eigenstates in quasiperiodic potentials
Models to pursue
Self-consistent treatment of photon fields
Practical realizations

Public-Private Division

- Public



- Private



Collaboration tool – most projects are

private

Workflow:

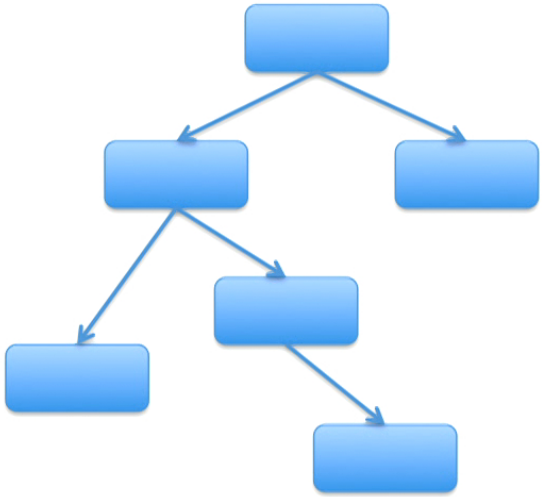
- Formulate Idea – *create top level node*
- Explore sub-branches of idea/ sub-ideas – *create subnodes*
- get results/refine idea – *go back and forth between nodes*

All along:

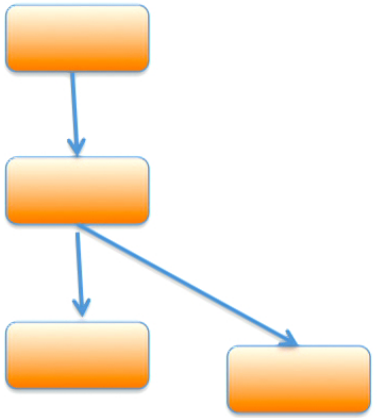
- Invite collaborators
- Track and comment on any changes
- Build reputation

Public-Private Fusion ("plug-and-play")

- Public

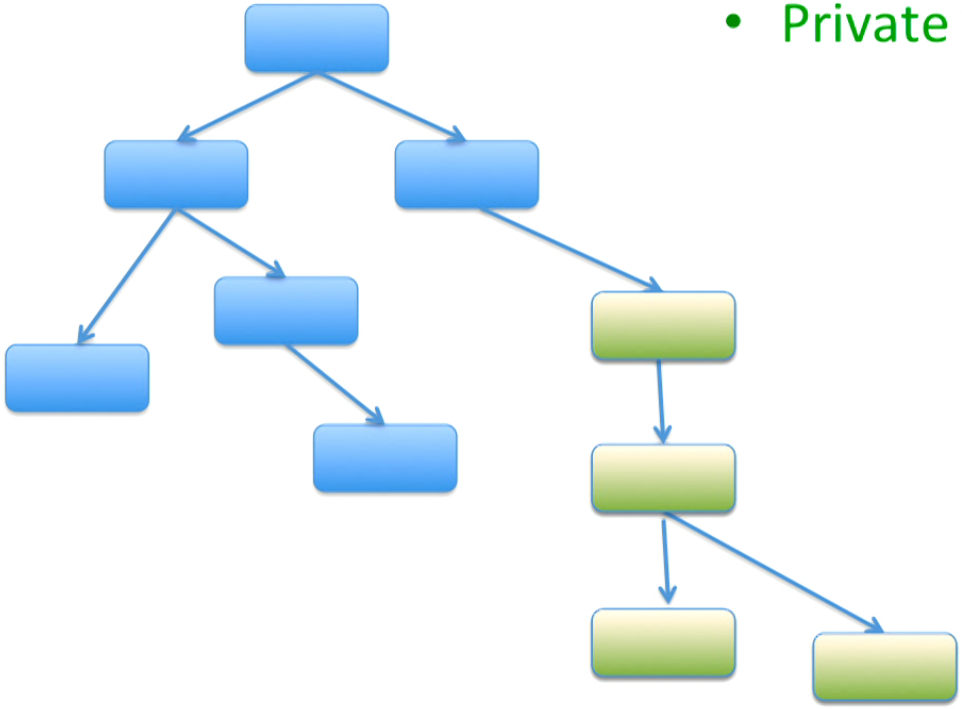


- Private



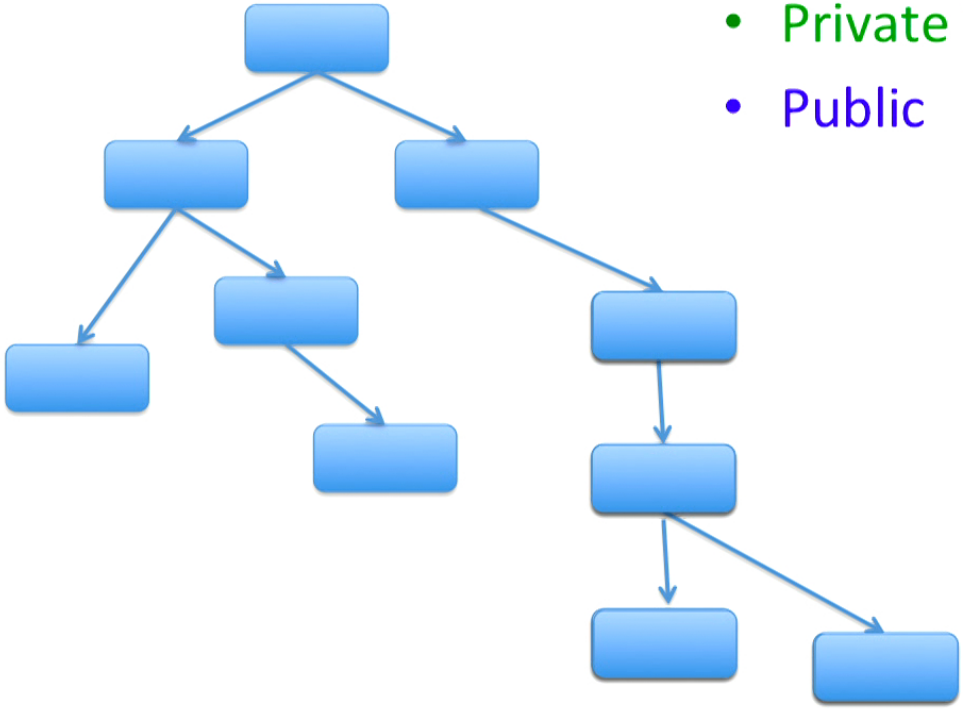
Public-Private Fusion ("plug-and-play")

- Public
- Private
- Private Visible



Public-Private Fusion ("plug-and-play")

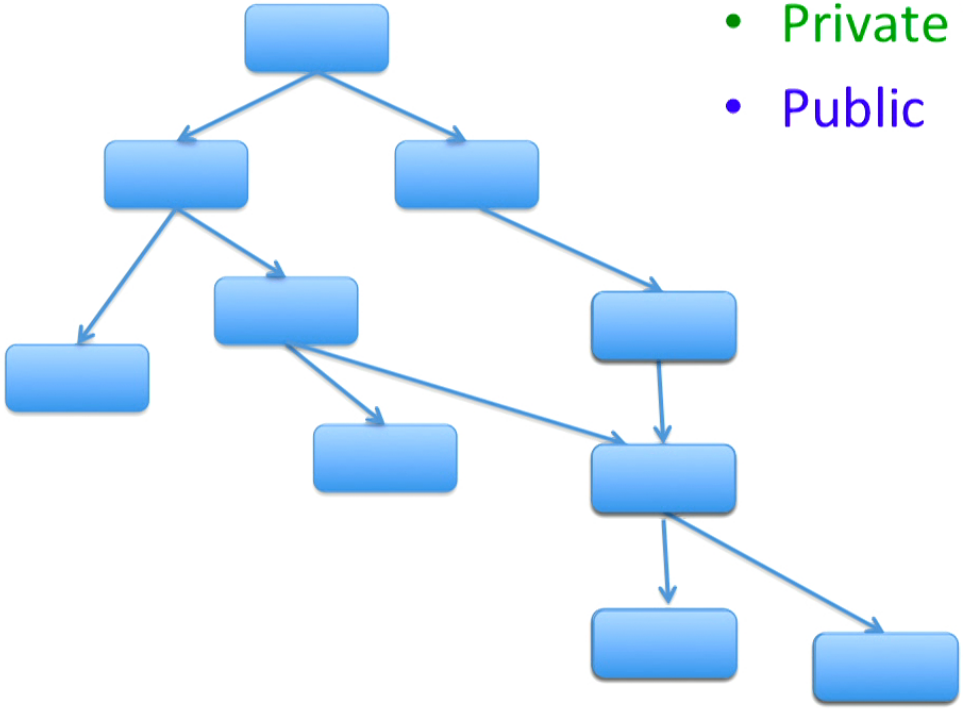
- Public



- Private
- Private Visible
- Public

Public-Private Fusion ("plug-and-play")

- Public

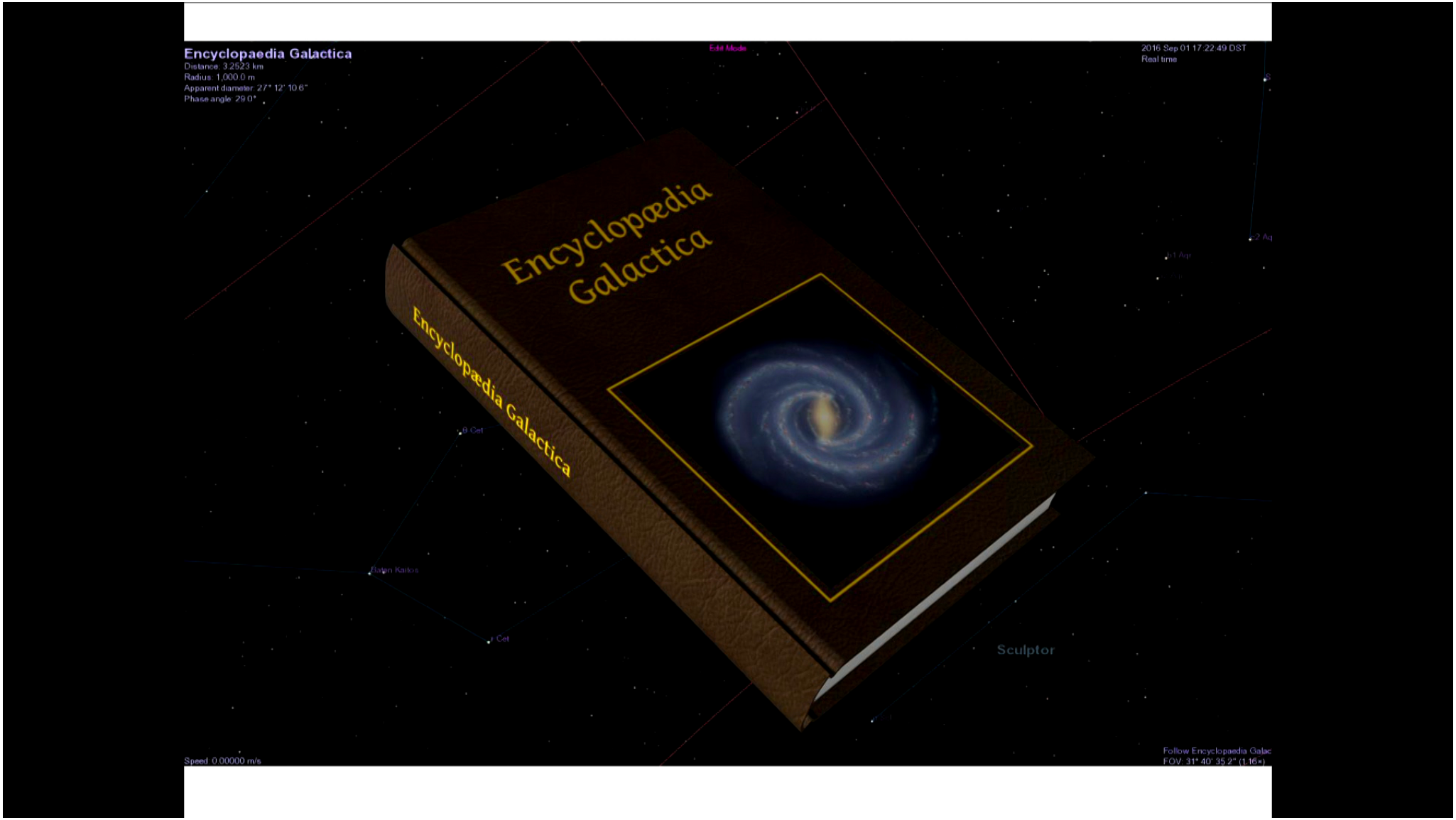


- Private
- Private Visible
- Public

Vision

- Global (decentralized) structured wiki-like science resource (knowen.org is an experimental prototype)
- Uses:
  - Outreach
  - Learning
  - Sharing results/methods/data
  - Raw material for packaging into books and reviews
- Reputation/priority recording
- Intelligent Automatic suggestions for content placement and search

Can exist in parallel with current article/journal infrastructure
Sustained by community (cf blockchain), or nonprofits (cf arxiv)


Encyclopaedia Galactica
Distance: 3.2523 km
Radius: 1,000.0 m
Apparent diameter: 27° 12' 10.6"
Phase angle: 29.0°
Edit Mode
2016 Sep 01 17:22:49 DST
Real time
Encyclopædia Galactica
Encyclopædia Galactica
9 Cet
Baten Kaitos
τ Cet
Sculptor
Speed: 0.00000 m/s
Follow Encyclopaedia Galac
FOV: 31° 40' 35.2" (1.16×)