

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*

*Known.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

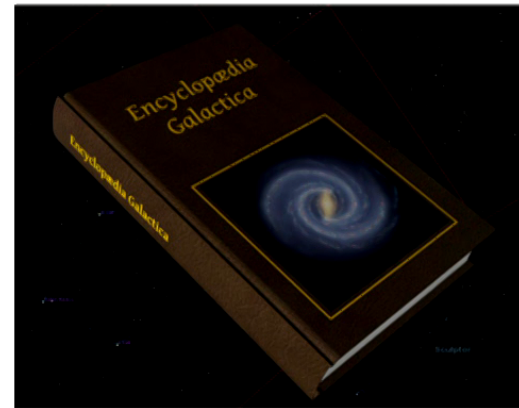
Wikipedia is a second example where scientists have missed an opportunity to innovate online. Wikipedia has a vision statement to warm a scientist's heart: "Imagine a world in which every single human being can freely share in the sum of all knowledge. That's our commitment."

...

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

 StackExchange


user:17787



**mathoverflow**

QuestionsTagsUsersBadgesUnanswered

ProfileActivity



1,663 REPUTATION

1024

**Xiao-Gang Wen** top 20% overall

I am a professor working on condensed matter theory. My current interest is in topological order, which correspond to patterns of long-range entanglement in many-body system.

Understanding patterns of many-body entanglement is related to some modern mathematics. For example, the math framework for 2D long-range entanglements happen to be unitary fusion category theory. For higher dimensions, we may need higher categories. The math framework for short-range entanglements with symmetry happen to be group cohomology theory of the symmetry group and cobordism theory.

# Incentives

- StackExchange model

The screenshot shows the StackExchange website interface. At the top, there's a navigation bar with 'Stack' and 'StackExchange' logos, and a user profile 'user:766'. Below this, the 'mathoverflow' logo is prominent, along with navigation tabs for 'Questions', 'Tags', 'Users' (highlighted), 'Badges', and 'Unanswered'. Under the 'Users' tab, there are two profile cards. The first card shows a partial profile with a reputation of 1,600 and 1 badge. The second card is for 'Terry Tao', a 'top 0.22% overall' user, who is a 'Professor of Mathematics at UCLA'. His profile shows a reputation of 54,130 and three badges: 17 (gold), 235 (silver), and 324 (bronze).

Stack StackExchange user:766

mat mathoverflow Questions Tags Users Badges Unanswered

Profile Profile Activity

1,600 1

54,130 REPUTATION

17 235 324

**Terry Tao** top 0.22% overall  
Professor of Mathematics at UCLA

# Structure

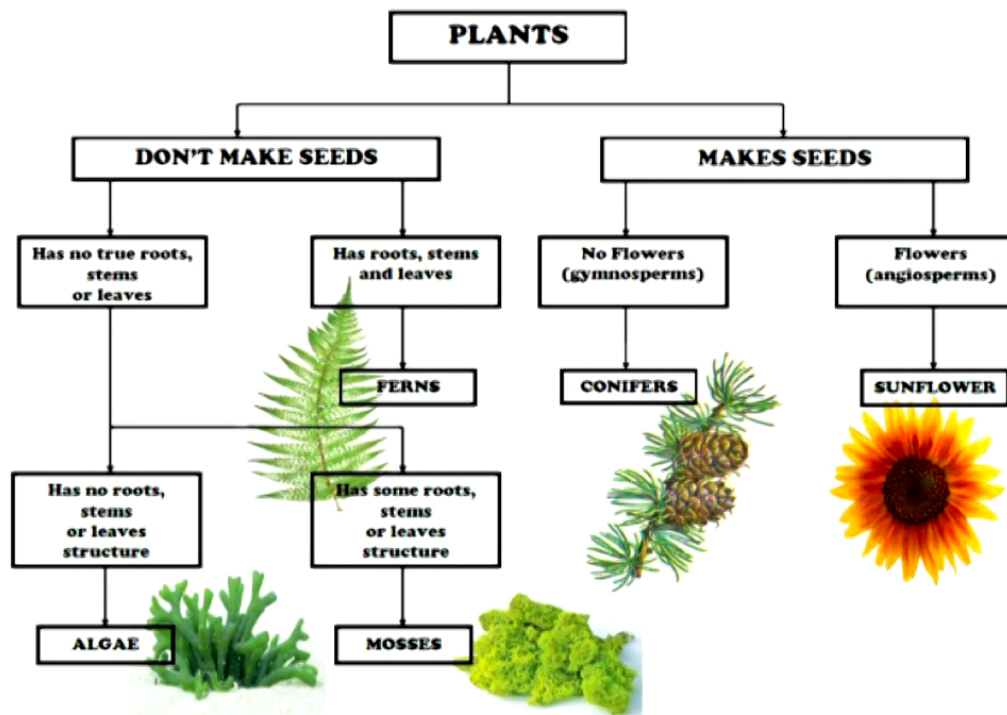
# What is the structure of knowledge?

# 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?

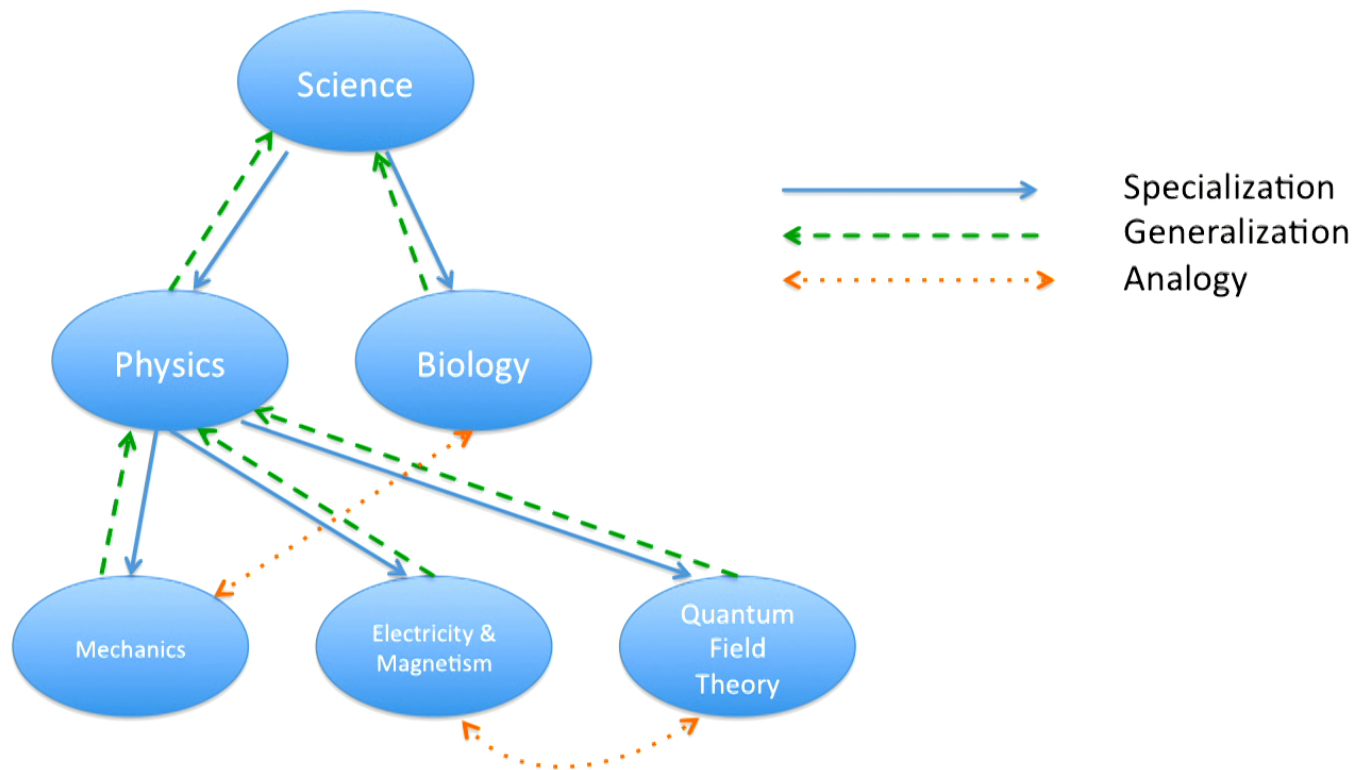
- Collections of facts -> connection of facts



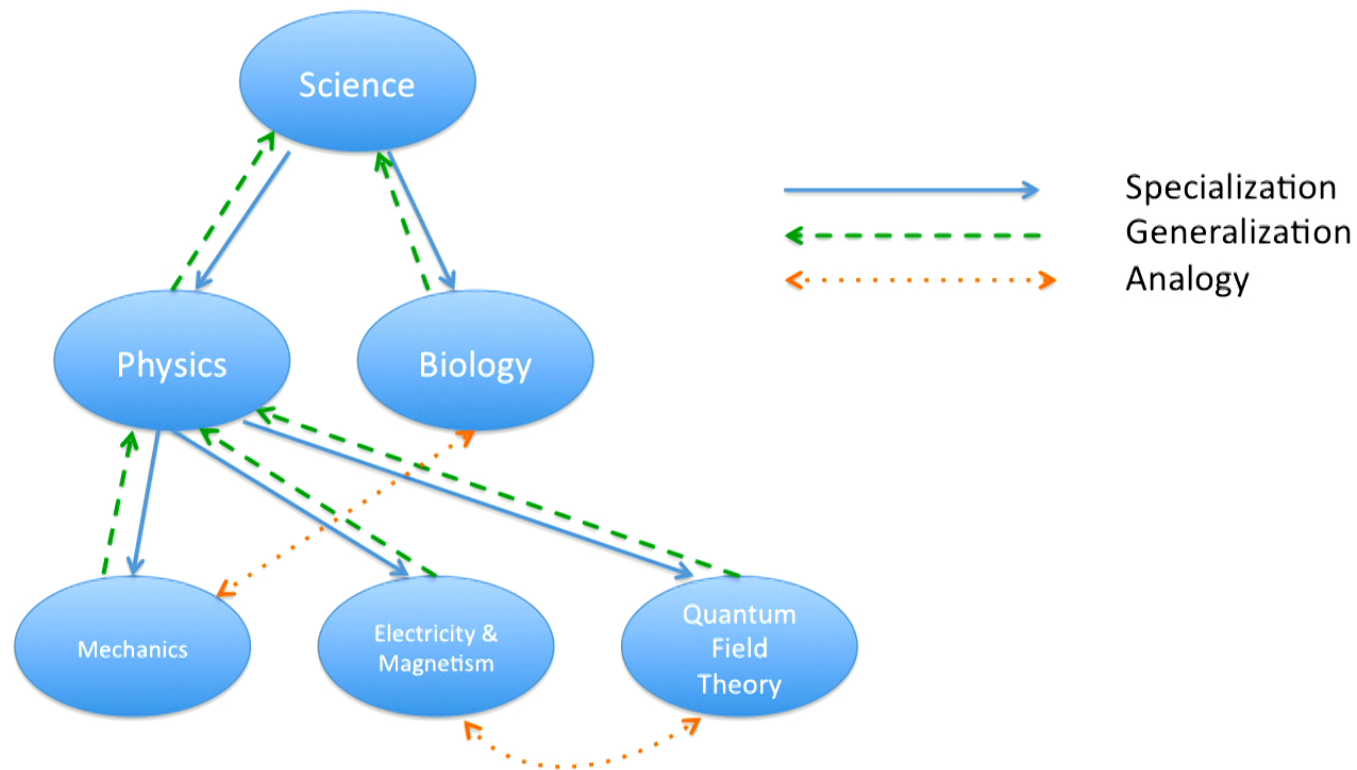
# Knowledge hierarchies

- **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](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](https://en.wikipedia.org/wiki/List_of_Dewey_Decimal_classes)
  - Lib of Congress: <https://www.loc.gov/catdir/cpsol/lcco/>

# Is knowledge a tree?

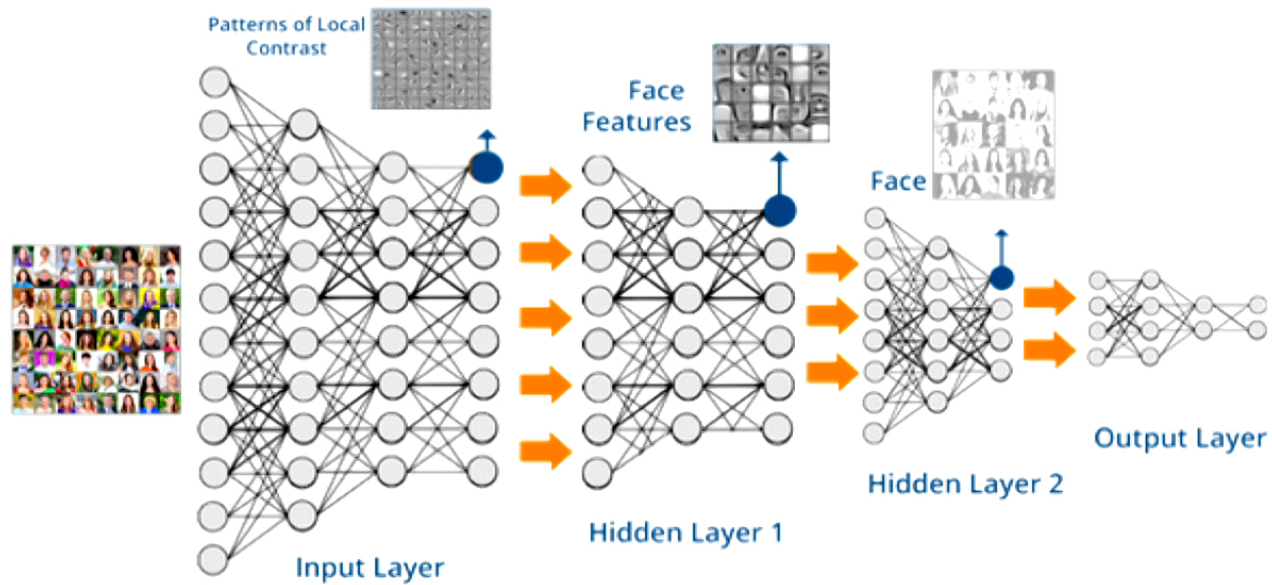


# Is knowledge a tree?

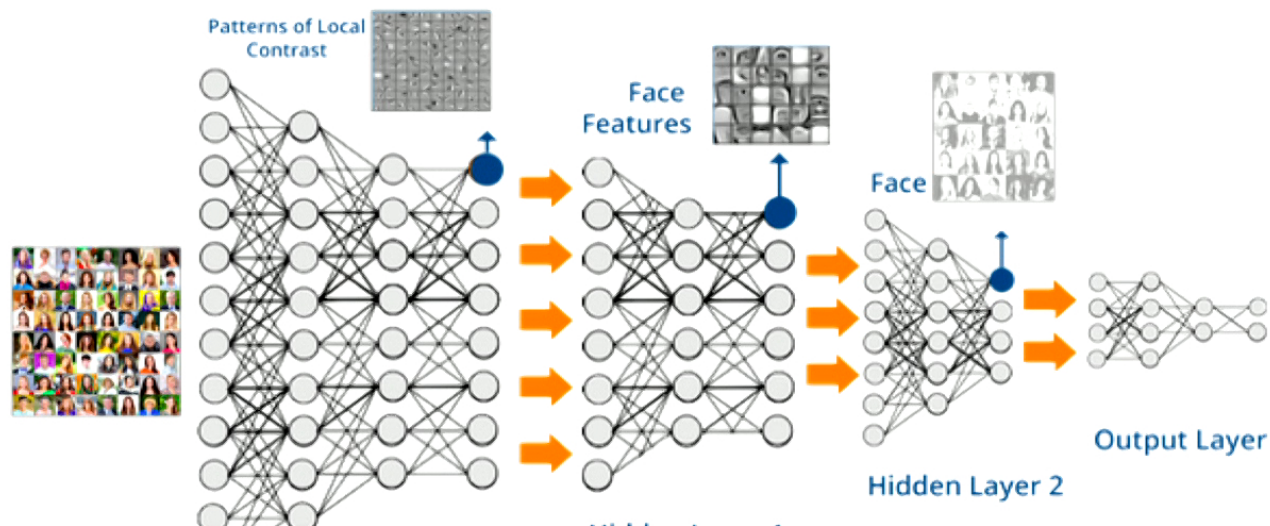


## Can it be shaped into a tree?

# Inspiration from RG/Machine learning



# Inspiration from RG/Machine learning



arXiv.org > stat > arXiv:1410.3831

Search or Article II

(Help | Advanced search)

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

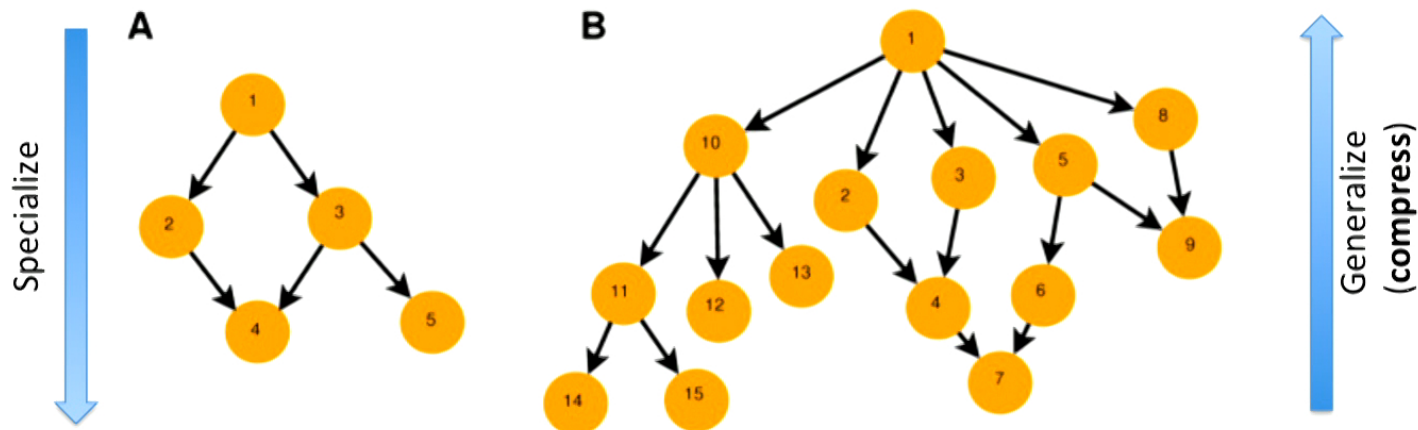
<b>Chapter 1—Basic Concepts</b>	1
1.1. Algorithms	1
1.2. Mathematical Preliminaries	10
1.2.1. Mathematical Induction	11
1.2.2. Numbers, Powers, and Logarithms	21
1.2.3. Sums and Products	27
1.2.4. Integer Functions and Elementary Number Theory	39
1.2.5. Permutations and Factorials	45
1.2.6. Binomial 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
<b>Chapter 2—Information Structures</b>	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

xviii

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
<b>Answers to Exercises</b>	466
<b>Appendix A—Tables of Numerical Quantities</b>	619
1. Fundamental Constants (decimal)	619
2. Fundamental Constants (octal)	620
3. Harmonic Numbers, Bernoulli Numbers, Fibonacci Numbers	621
<b>Appendix B—Index to Notations</b>	623
<b>Index and Glossary</b>	628

# Why Tree or DAG structure is good

- depth is  $\sim \log N$
- 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
  - Known

Chicken or Egg?



# Experiment: knownen.org

The screenshot shows a web browser window with the URL `knownen.org`. The page layout includes a left sidebar with navigation options: Dashboard, Knownen, Graph of pages, Search Knownen, Help, and Ivar Martin. The main content area displays a page titled "Incommensurate frequency Floquet" with a subtitle "Main idea:". The page content discusses Floquet lattices and Hamiltonians, mentioning concepts like "tilted" and "edge-modes". A sidebar on the left lists "Featured pages" (Favorites, Roots, Templates, 10 recently visited, 10 recently changed) and "Related pages" (Parents, Siblings, Children). A blue circular icon with a white signal symbol is visible in the bottom right corner of the page.

Dashboard Knownen Graph of pages Search Knownen Help Ivar Martin

Now you are on the very top of [Incommensurate frequency Floquet project](#). [Manage project](#)

Parent pages

You subscribed to updates from this page and it's subgraph

## Incommensurate frequency Floquet

**Main idea:**

Driving a system with multiple incommensurate frequencies leads to Floquet lattices in multiple dimensions. Various lattice Hamiltonians can be implemented, including topologically non-trivial ones. The number of bands can be controlled by choosing appropriate building blocks -- two level systems will lead to two-band Hamiltonians, 3-level -- to 3 band, etc.

The distinguishing feature of such Floquet lattice Hamiltonians is that they are **tilted**, with the tilt potential  $U(n_1, n_2, \dots) = n_1 \omega_1 + n_2 \omega_2 + \dots$ . That has several consequences:

- the wave-functions are localized in the direction perpendicular to the hyperplane  $U(n_1, n_2, \dots) = \text{const.}$ 
  - for weak drive the wavefunction will be localized on the lattice points closest to the hyperplanes. The lattice potential at these points is quasiperiodic.
  - for strong drive, w.f. will be weakly confined, as if they are near a smooth edge of a sample
- if band structure is topologically non-trivial, every hyperplane is also a locus of topologically protected "edge-modes." For instance, in the case of two-dim Floquet space, there will be helical modes propagating along the lines  $n_1 \omega_1 + n_2 \omega_2 = \text{const.}$  Chiral propagation correspond to preferential absorption of one frequency photon and emission of the other.
- Even when the drive frequencies are commensurate, the pumping between the modes persists, as can be

# 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

Close graph
Page preview panel ON

## Floquet states under two-tone drive

Consider a zero-dimensional system with basis states  $\alpha$  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  $2\pi$ -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  $\theta_{1,2}$ ,

$$\mathcal{H}(\theta_1 + \omega_1 t, \theta_2 + \omega_2 t) = \sum_{m_1, m_2} h_{m_1 m_2} e^{-i\omega_1 m_1 t - i\omega_2 m_2 t},$$

Now, substituting into the Schrodinger equation Eq. (1), we get,

$$(E + \omega_1 n + \omega_2 m) \phi_{nm}^\alpha = \sum_{p, q} h_{pq}^{\alpha\beta} \phi_{n-p, m-q}^\beta. \quad (3)$$

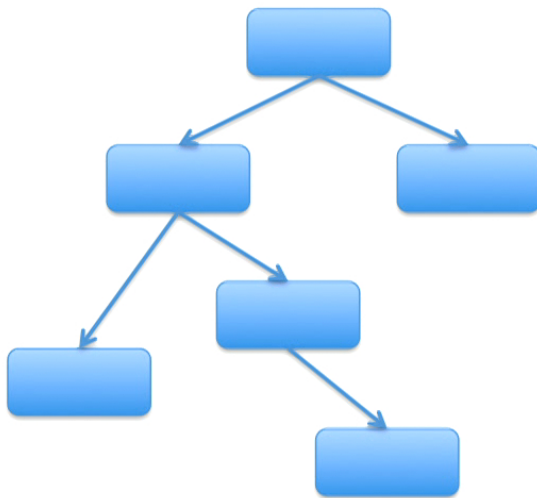
This is the key equation that describes hopping in the Floquet space. It can be solved by diagonalizing a sufficiently large block of it. The integer coordinates

```

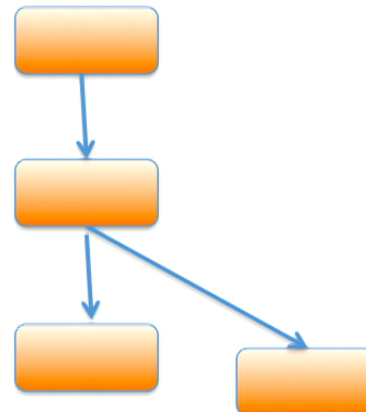
graph TD
    A[Floquet states under two-tone drive] --> B[Weak drive]
    A --> C[Strong drive]
    A --> D[Methods for tracking evolution in Floquet space]
    A --> E[Commensurate drive frequencies]
    A --> F[Floquet eigenstates in quasiperiodic potentials]
    A --> G[Self-consistent treatment of photon fields]
    A --> H[Practical realizations]
    A --> I[Models to pursue]
    A --> J[Paper outline]
    A --> K[Incommensurate frequency Floquet]
    
```

# Public-Private Division

- Public



- Private



# Collaboration tool – most projects are **private**

## Workflow:

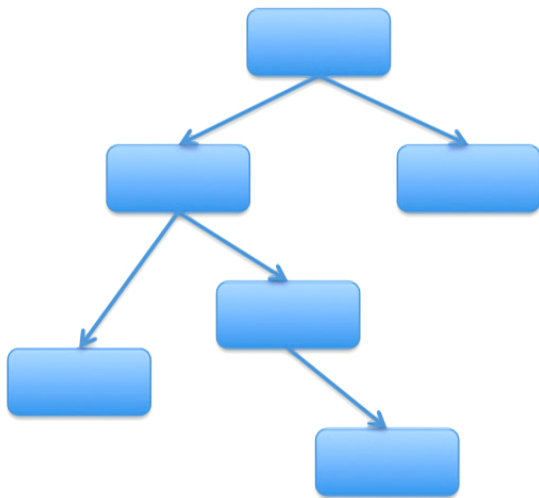
- Formulate Idea – *create top level node*
- Explore sub-branches of idea/ sub-ideas – *create subnodes*
- get results/refine idea – *go back and forth between nodes*

## *All along:*

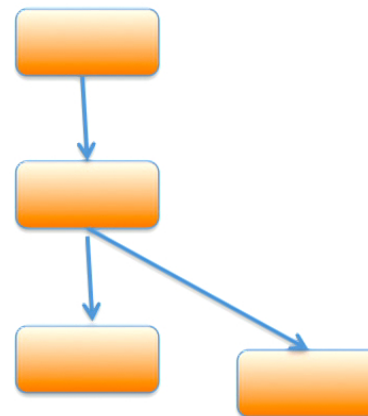
- Invite collaborators
- Track and comment on any changes
- Build reputation

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

- Public



- Private

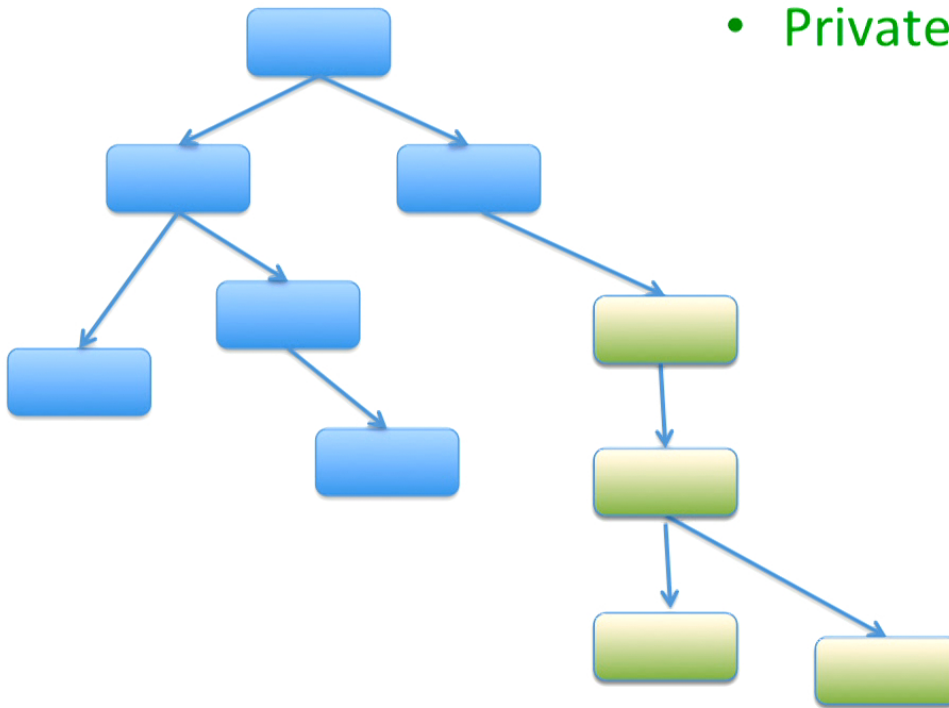


# Public-Private Fusion ("plug-and-play")

- Public

- Private

- Private Visible



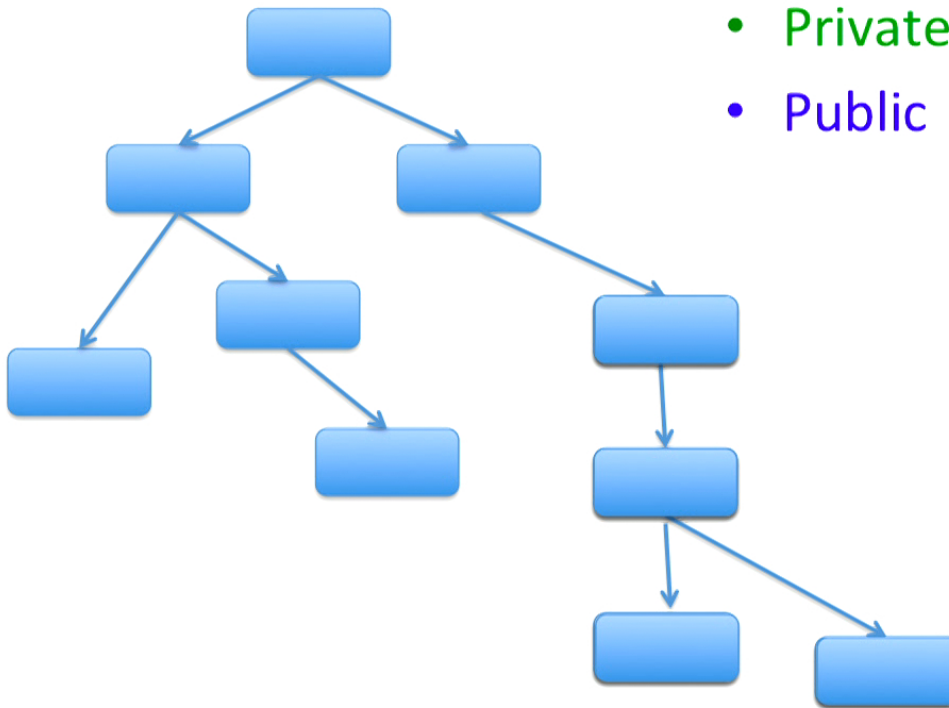
# Public-Private Fusion ("plug-and-play")

- Public

- Private

- Private Visible

- Public



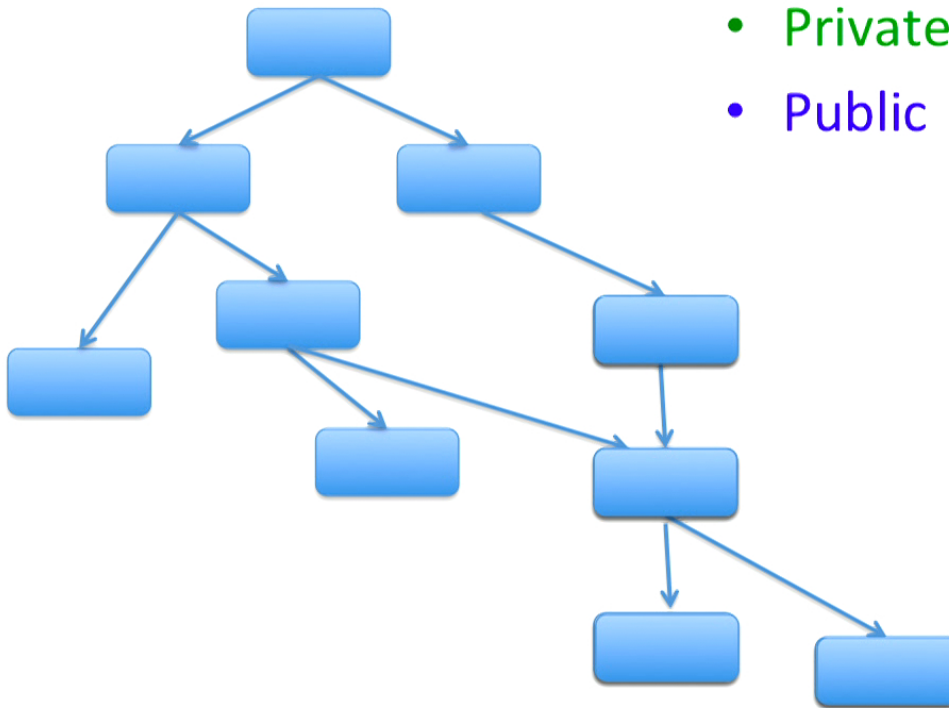
# Public-Private Fusion ("plug-and-play")

- Public

- Private

- Private Visible

- Public



# Vision

- Global (decentralized) structured wiki-like science resource (knownen.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)

## Encyclopædia Galactica

Distance: 3,252,3 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

