

**G3N1U5<sup>®</sup>**



Data is the new water in the  
digital age

Anthony G Nolan OAM JP

[tony@g3n1u5.com](mailto:tony@g3n1u5.com)

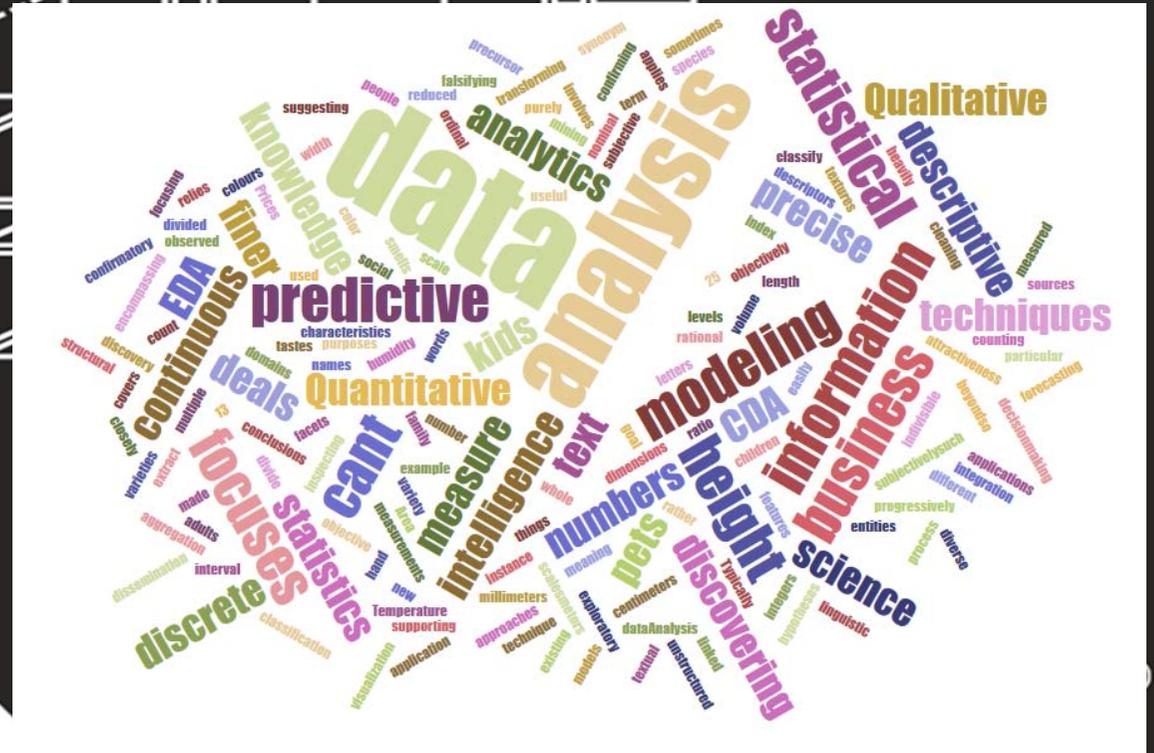
0417270664

**G3N1U5<sup>®</sup>**



# What is Data?

Data is everything, and it can be found everywhere. It exists in two states at the same time, data and metadata.



# G3NTUS

Data ->

Colour

Sex

Shape

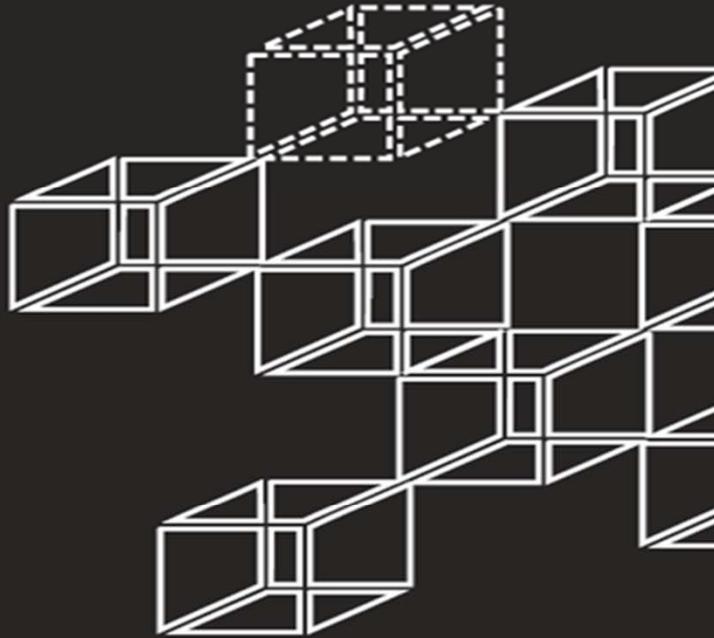
Height

Weight

Material

Jewellery

etc



# G3N105<sup>®</sup>

Jewellery  
Metadata ->

Purpose

Distance

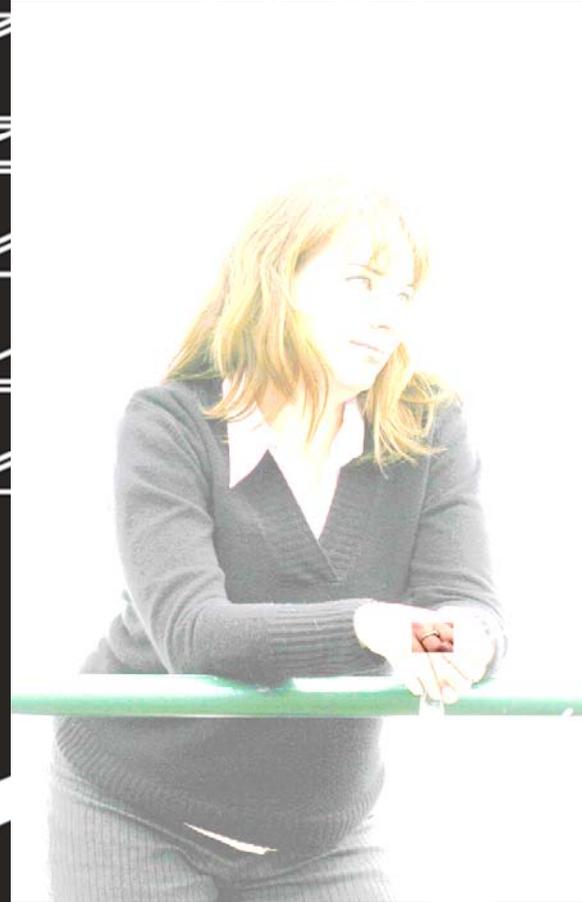
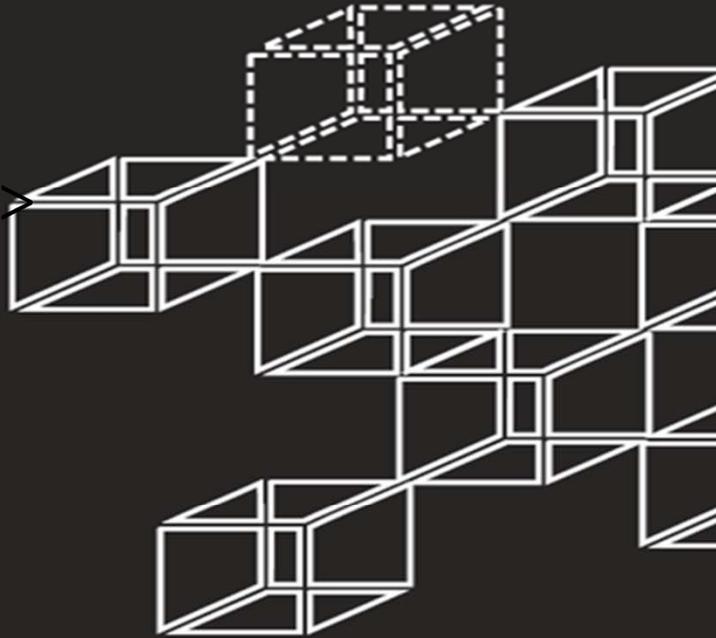
Sex

Colour

Size

Weight

Pose



**G3N105<sup>®</sup>**

# Open Source

The world is full of open source data, and every part of reality can be found some where on the internet, in either the world wide web, or the dark web.



# G3N1U5

## The Data Pillars

These are my five different facets of data



- Data that is directly observed.
- Data that describes the primary data.
- Data that is transformed, transferred, transposed
- Space/ Time about primary data
- Activity about the primary data
- Casualty to and from the primary data
- Data is always subjected to the scenario

# G3N1U5®

## Golden Rules

Everything to most basic

Rescale data to true scale

Back to reality

Distance makes the difference

- $\$34.45 = 3445$

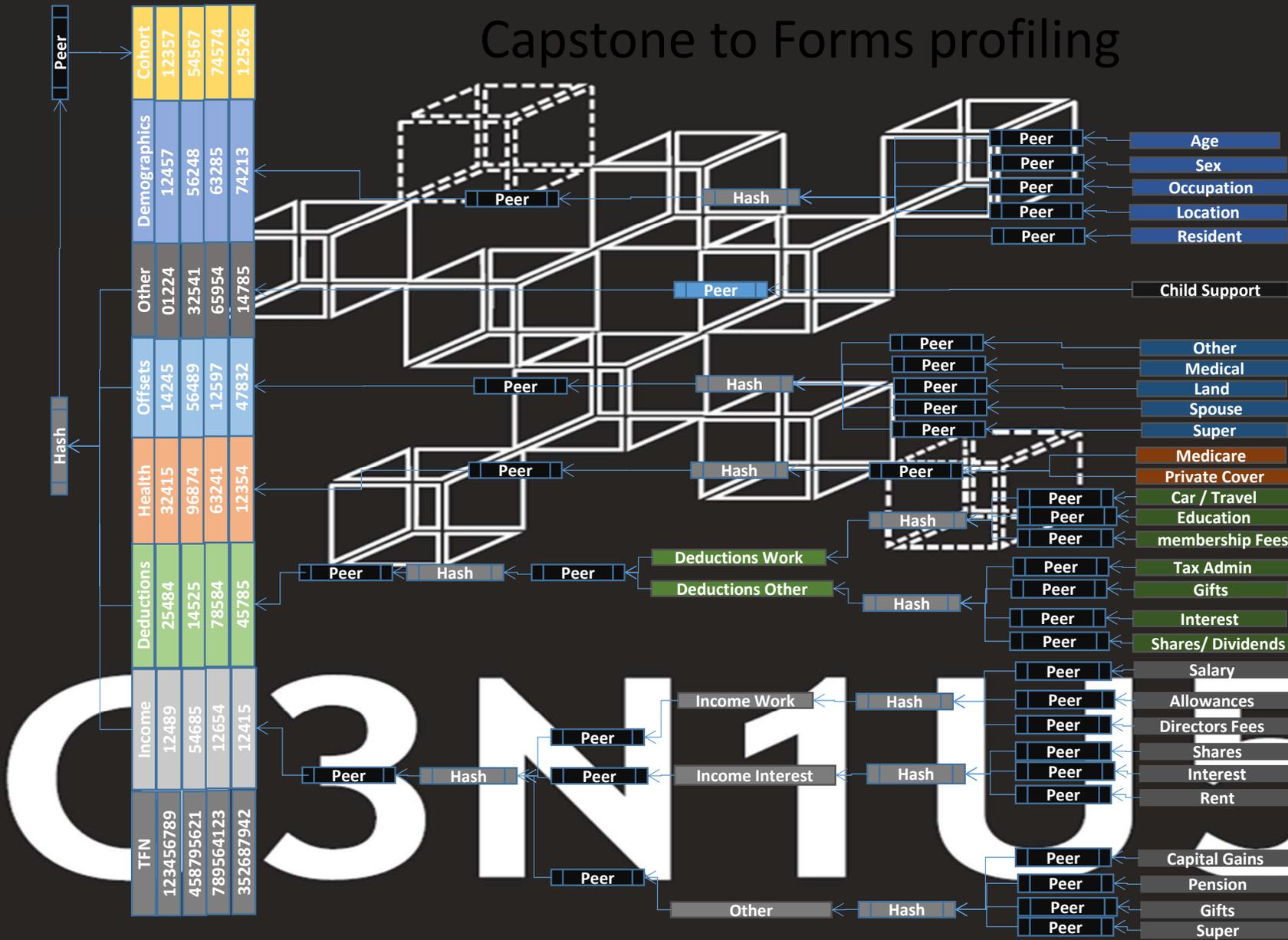
- $16C = 290k$

- $2\% = 51,526$

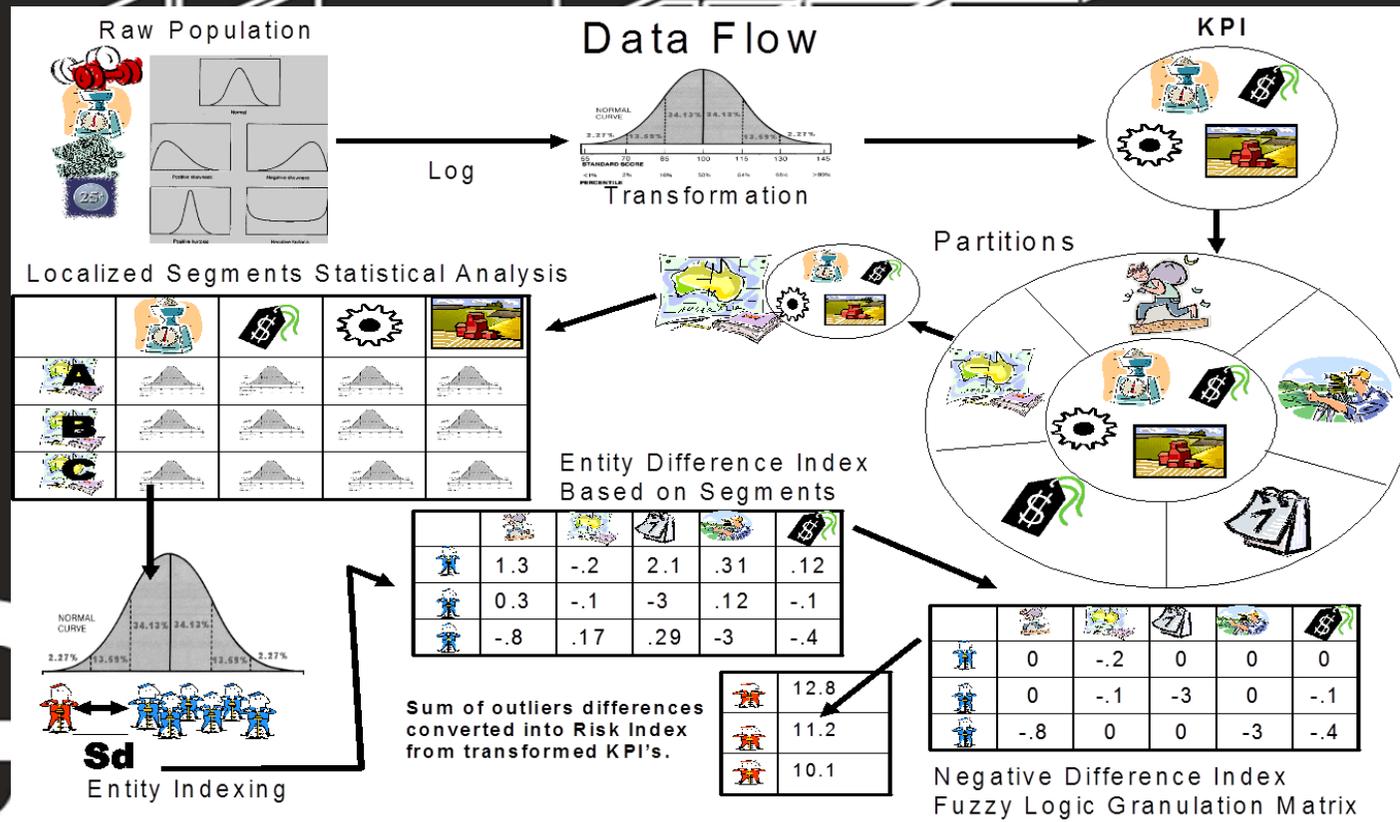


# G3NTU5<sup>®</sup>

# Capstone to Forms profiling

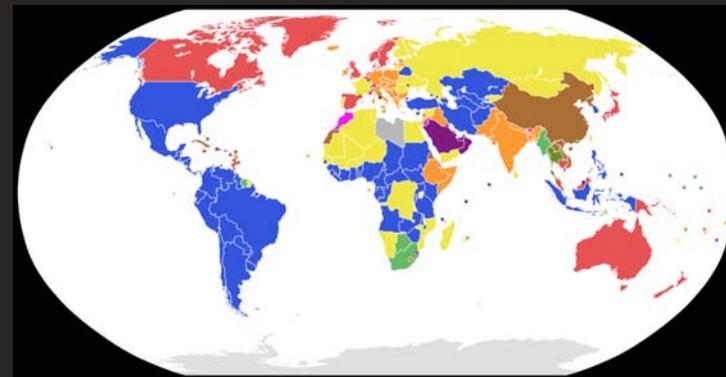
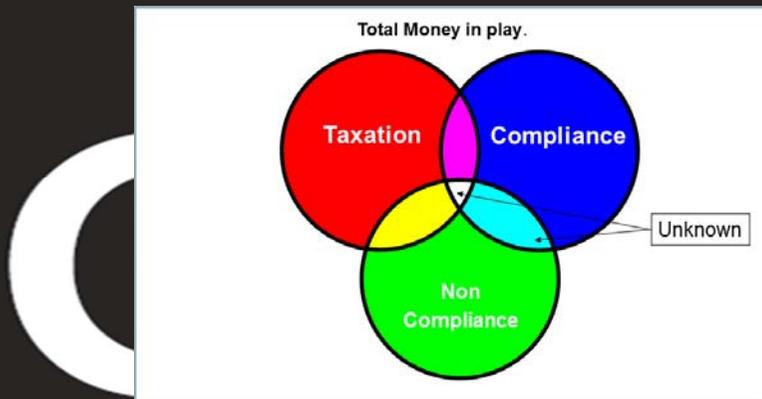
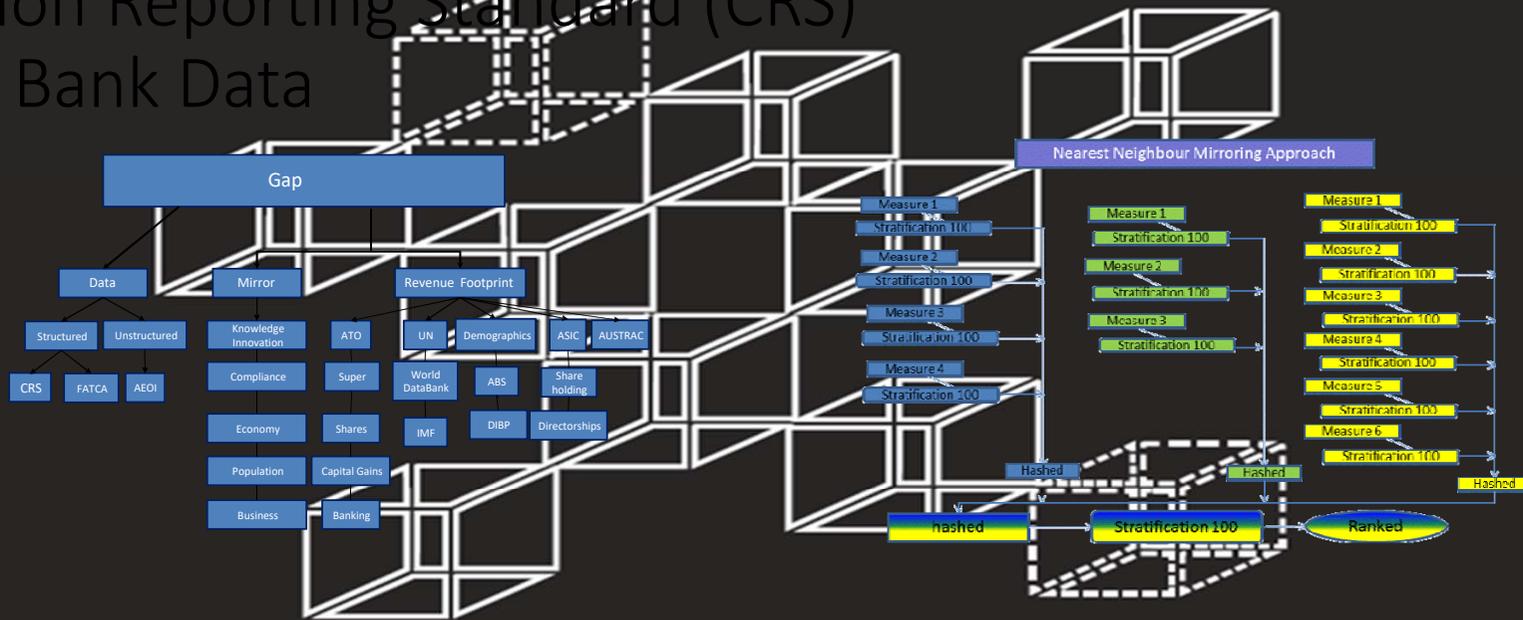


# A system of systems approach to Data Modelling.



**G** **5** <sup>®</sup>

# Common Reporting Standard (CRS) World Bank Data



## GIS / Demographics

- Target CCD
- Similar CCD
- Used Combinations
- Match Profiles
- Near Match Profiles

- Like geographic zones, must be matched with Demographic characteristics.
- Other degrees of freedom can be associated with near matches if desired.

16

## Describing the target

Male	Sex
22	Age
\$1,857	Income Week
Engineer	Occupation
Transport	Employment
Single	Married Status
	Other

- This is an example of one of our target population.

15

## Locate Other CCD's

- Similar CCD
- Dissimilar CCD
- Target CCD

- This is an example of other CCD's that are close by that have a similar Index value.

8

## SEFIA Area Example

- This is an example which shows the different Geographic Zones within Australia, and how these break down to State level and then to CCD zones, combined with demographic profiles, derived from the ABS CDATA tables.

8

# Exclusive vs Inclusive & Extended Dataset

Inclusive analytics – Where you use as many variables as you can.

An enhanced dataset, is designed to be used in a variety of different software applications, and used in a number of different techniques.

Exclusive analytics – Where you reduce your dataset

An exclusive dataset, is reduced to as little variables as possible, and is mainly designed for a single software application or methodology.

Raw	NA	Val Type	- NA	Base Adj	B Adj Log	Peer Strat
NA	Y	1	0	145	2.16	11
0	N	2	0	145	2.16	11
1245	N	4	1245	1390	3.143	99
-145.23	N	3	-145	1.23	0.089	1

Raw	<u>DayN</u>	<u>DayT</u>	<u>MntN</u>	<u>MntT</u>	Year	<u>WeekN</u>	<u>FinYWN</u>
01/01/16	01	Fri	01	Jan	2016	01	27
30/08/16	30	Tue	08	Aug	2016	35	10
15/06/15	15	Mon	06	Jun	2015	24	49

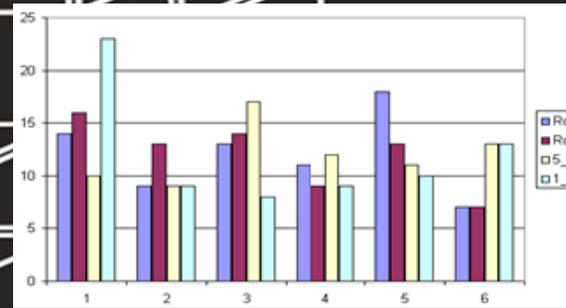
# G3NTU5

- Show numbers 1 to 45

# Quasi Random

There is no such thing as randomise in the real world, every thing is cause and effect

Pseudo Randomness, is events that happen below the threshold we can observe.

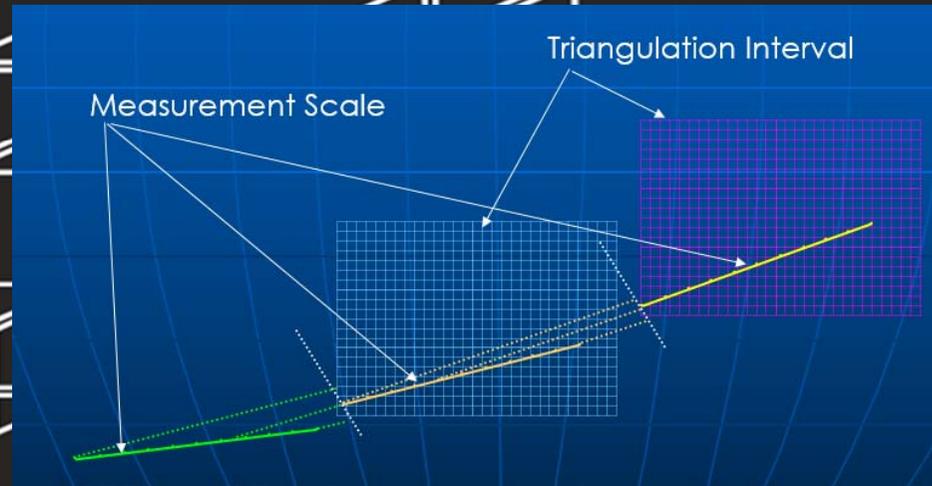
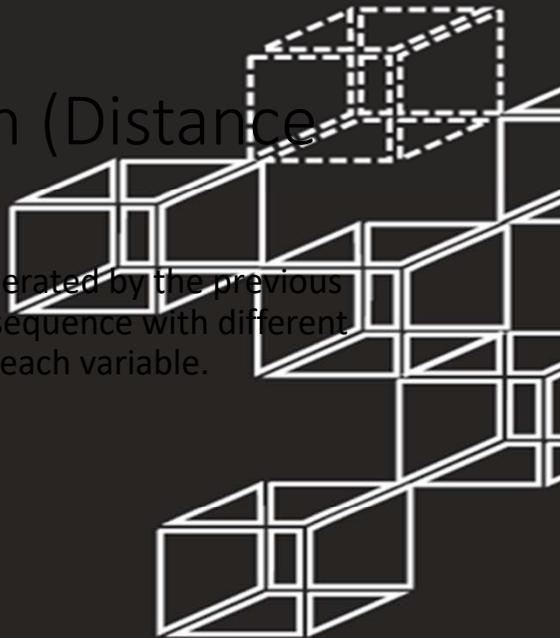


39		9		12		30		33		45		5		7		19
42		7		36		23		39		29		16		19		43
24		34		17		3		11		1		43		39		4
22		33		18		12		45		43		28		23		13
13		17		30		7		32		41		10		20		14
45	2	23	2	5		4	3	33	2	1	3	29	4	39	3	40
39	1	35		31		4	1	26		23	1	41	2	11	4	13
41	1	26	1	28	4	33	2	3	5	6		32	3	40	2	1
15		41	1	17	4	10	4	4	2	26	1	30	4	9	3	29
14	5	29	1	9	1	23	3	18	6	19	8	5	4	8		37

# G3N1U5®

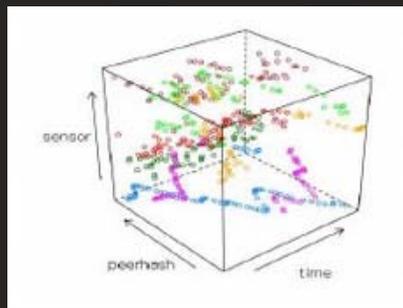
# Digital Hash (Distance Preserving)

A unique number, generated by the previous values in a cascading sequence with different weightings applied to each variable.



$$\text{Score} = \sum_{i=1}^n X_i * \text{Cos}(Y*i)$$

# G3N1



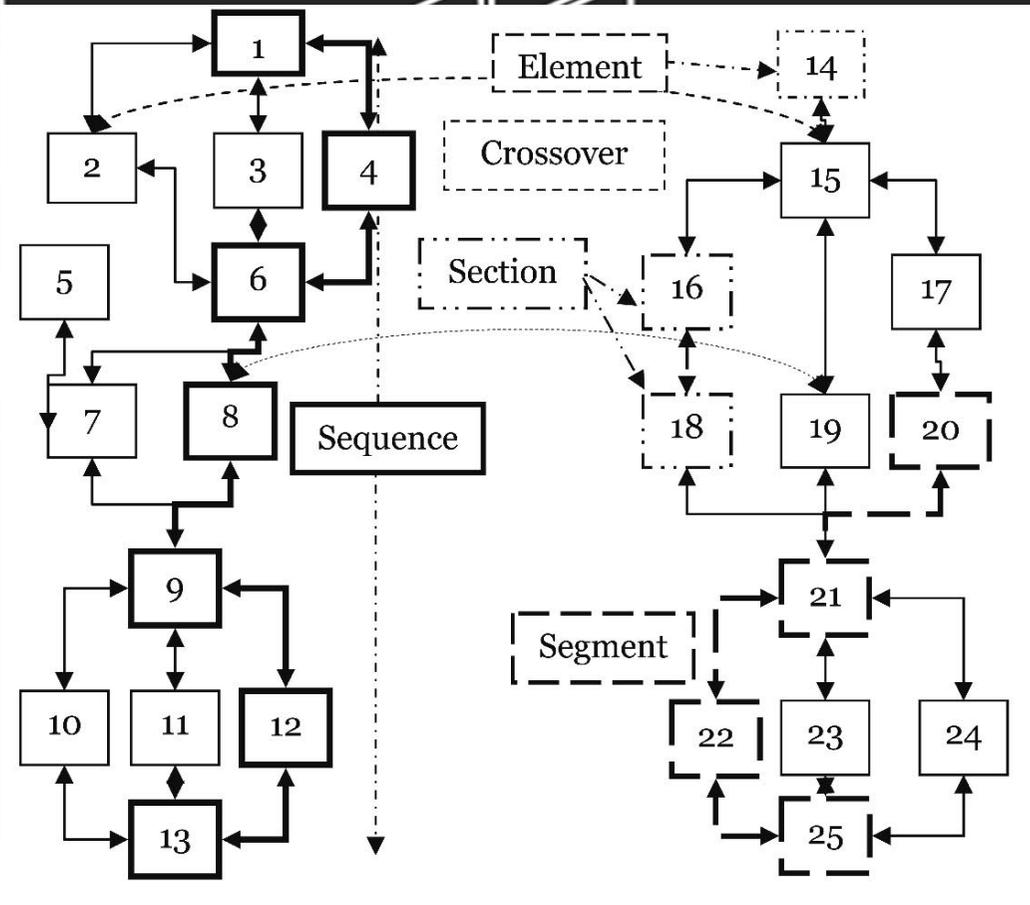
# 5<sup>®</sup>

# Hyperpanofiction

Stories within stories

Hyperpanofiction is a series of short stories linked together to make bigger stories. The difference is that these links connect the stories, not only in a forward direction as in a usual story timelines, but also backwards as well. Hyperpanofiction also has multiple lead characters. Using the links you can move between main characters, the same way you can move within time lines. But don't think that it has to stop there.

With Hyperpanofiction, you can also move between locations, genres, themes, and storylines. Its really up to the author to provide the links. Hyperpanofiction is really about alternatives, and your opportunity to move between them in an interesting and investigative manner.



# G3NTO3

# Peer Transformation

This is a reshape transformation, that rescales into the desired number of stratifications or sup-populations. It is not mean dependant, so it not affected by skewness

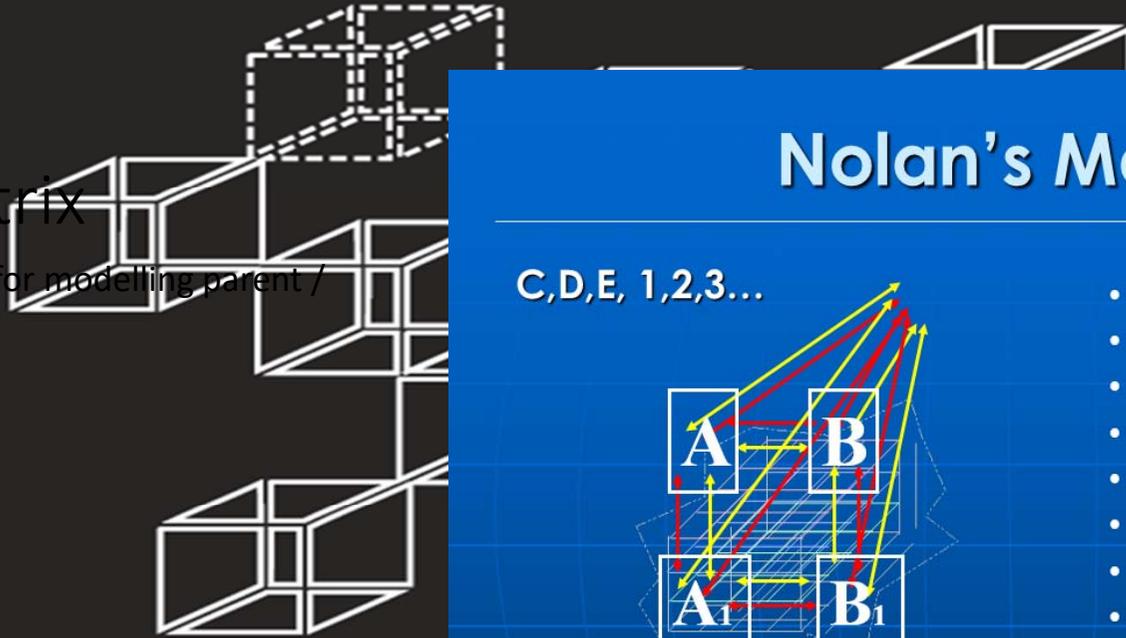
A	12	54	35	67	B	43	46	35	77	C	73	63	25	44	D	75	23	45	54
A	41	21	36	14	B	35	15	43	62	C	85	64	31	45	D	12	65	32	56
A	15	32	64	14	B	52	18	38	65	C	85	47	45	78	D	65	96	21	45
A	16	41	32	41	B	14	36	89	15	C	47	56	98	32	D	89	24	69	74
A	65	17	34	31	B	18	63	42	79	C	71	54	27	61	D	74	24	64	36

$$X = \frac{(x - x_{\min})}{(x_{\max} - x_{\min})} * n$$

# G3N1U5<sup>®</sup>

# Nolans Matrix

An interactive matrix for modelling parent / child relationships.



## Nolan's Matrix

C,D,E, 1,2,3...

- Finite Mathematics
- Fuzzy Logic
- Granulation
- Clustering
- Quasi Fractals Patterns
- Measurement Scales
- Between & Within Group
- Hyper Dimensional Plotting
- Hyper-Intervals
- Dynamic Modeling

Hyperdimensional Relationships

Hierarchical Analysis

Bounded Rationality

Pattern Recognition

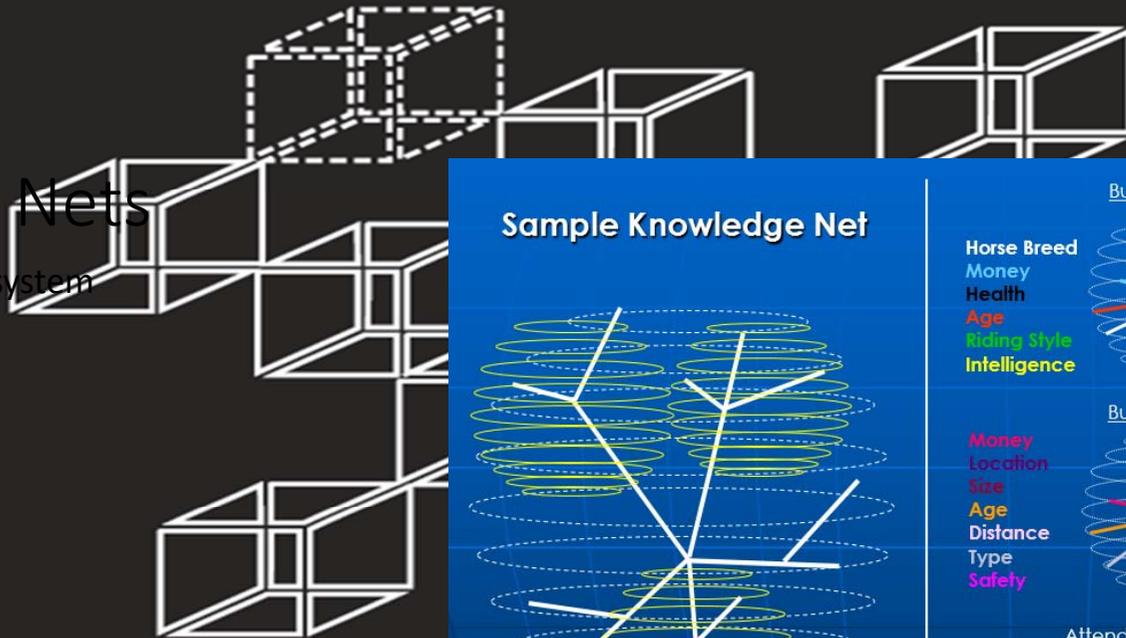
Modeling

# G3N105

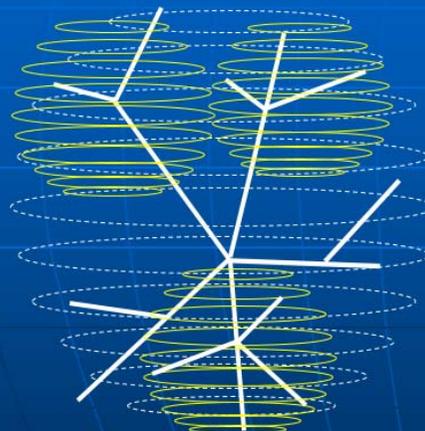
# Knowledge Nets

Subject classification system

3d fuzzy logic globe



## Sample Knowledge Net



Category	Value	Subject
Horse Breed	000	Generalities
Money	100	Philosophy
Health	200	Religion
Age	300	Social Sciences
Riding Style	400	Language
Intelligence	500	Natural Sciences & Mathematics
	600	Technology (Applied Sciences)
	700	The Arts
	800	Literature & Rhetoric
	900	Geography & History

Category	Value	Subject
Money	000	Generalities
Location	100	Philosophy
Size	200	Religion
Age	300	Social Sciences
Distance	400	Language
Type	500	Natural Sciences & Mathematics
Safety	600	Technology (Applied Sciences)
	700	The Arts
	800	Literature & Rhetoric
	900	Geography & History

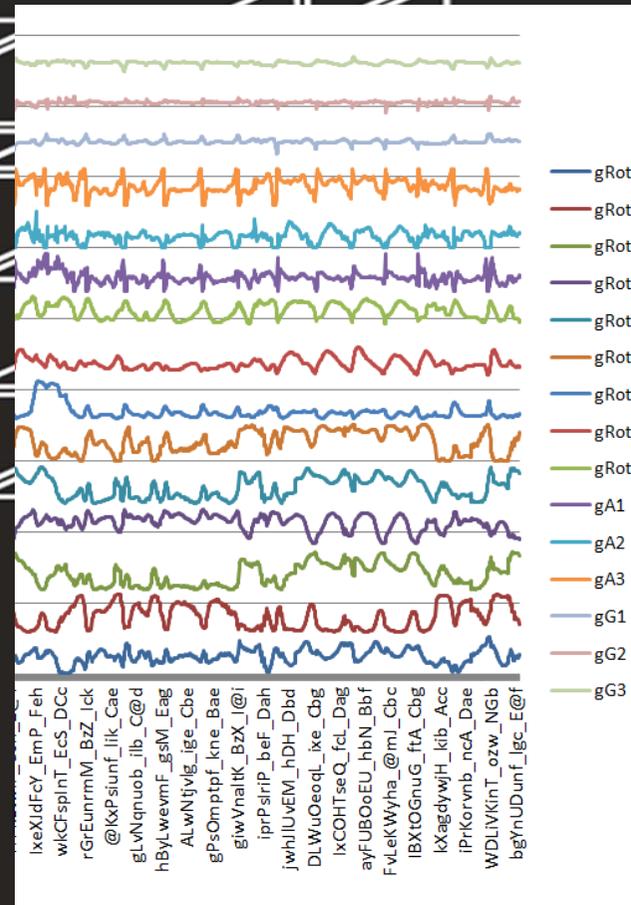
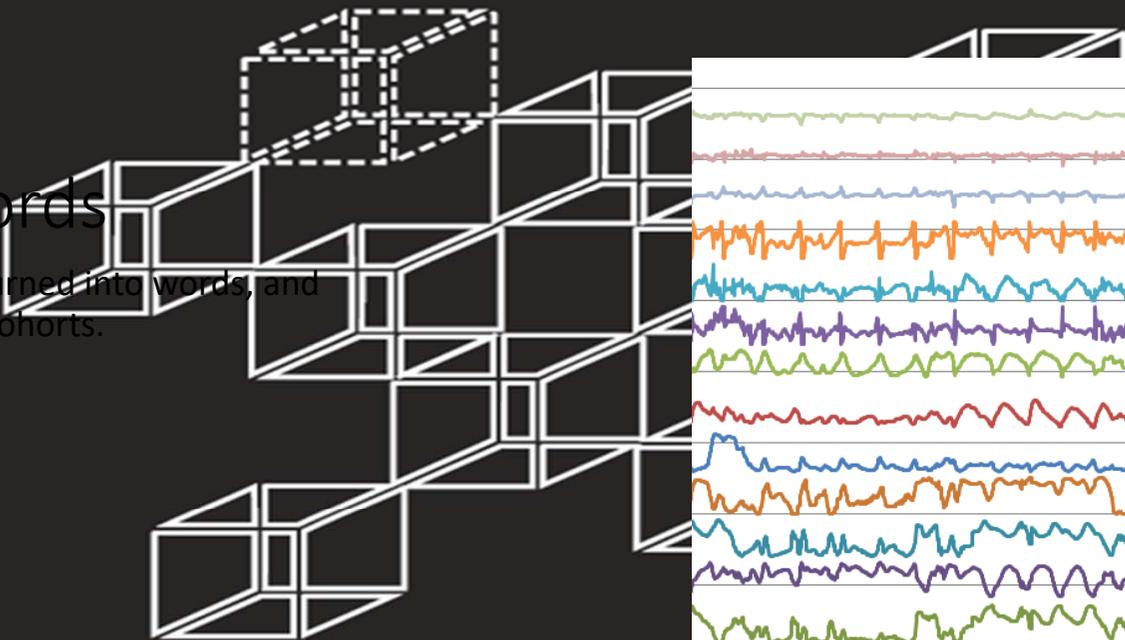
  

Category	Value	Subject
Money	000	Generalities
Travel	100	Philosophy
Academic	200	Religion
Celebration	300	Social Sciences
Insurance	400	Language
	500	Natural Sciences & Mathematics
	600	Technology (Applied Sciences)
	700	The Arts
	800	Literature & Rhetoric
	900	Geography & History

# G3NTUS

# Number words

Where numbers are turned into words, and can be used to make cohorts.



# G3N TO5



39	9	12	30	33	45	5	7	19	399123033455719	Q	h	k	H	K	W	r	p	d	QnkHKWrp
42	7	36	23	39	29	16	19	43	42736233929161943	T	p	N	A	Q	G	g	d	U	TpNAQGgd
24	34	17	3	11	1	43	39	4	243417311143394	B	l	f	t	l	v	U	Q	s	BLftlvUQ
22	33	18	12	45	43	28	23	13	223318124543282313	a	K	e	k	W	U	F	A	j	aKekWUFA
13	17	30	7	32	41	10	20	14	13173073241102014	j	f	H	p	J	S	m	c	i	jfHpJSmc
45	23	5	4	33	1	29	39	40	452354331293940	W	A	r	s	K	v	G	Q	R	WArsvKvGQR
39	35	31	4	26	23	41	11	13	39353142623411113	Q	M	l	s	D	A	S	L	i	QMIsDASI
41	26	28	33	3	6	32	40	1	412628333632401	S	D	F	K	t	a	J	R	v	SDFKtqJR
15	41	17	10	4	26	30	9	29	1541171042630929	h	S	f	m	s	D	H	n	G	hSfmsDHn
14	29	9	23	18	19	5	8	37	142992318195837	i	G	n	A	e	d	r	0	0	iGnAedr00
35	12	17	2	40	31	39	33	13	35121724031393313	M	k	f	u	R	l	Q	K	j	MkfuRIQK
12	44	1	40	29	6	22	34	4	124414029622344	k	V	v	R	G	q	a	L	s	kVvRGqaL
41	26	36	4	2	38	11	19	20	4126364238111920	S	D	N	s	u	P	l	d	c	SDNsuPldc
27	16	28	44	25	37	23	19	22	271628442537231922	E	g	F	V	C	O	A	d	a	EgFVCOAd
22	23	3	20	37	8	41	21	28	2223320378412128	a	A	t	c	O	O	S	b	F	aAtcOOSbF
20	14	17	8	3	21	2	38	7	20141783212387	c	i	f	0	t	b	u	P	p	cif0tbuPp
5	32	26	4	30	23	28	11	21	5322643023281121	r	J	D	s	H	A	F	l	b	rJDsHAFIb
13	5	1	21	11	3	35	37	44	135121113353744	j	r	v	b	l	t	M	O	V	jrvbltMO
17	33	6	41	27	45	25	1	9	173364127452519	f	K	q	S	E	W	C	v	n	fKqSEWCVn



# Big Data Files Mapping

## - Indexing your files in their names.

So if your data warehouse or electronic document filing system is so big that you do not know every file name by heart, then you have a problem. Lets say that your storage system is inside the movie TRON, then it would probably look like the picture above. Especially if you had to physically port in, find the file and port out.

As I have said in my previous LinkedIn Pulse posts, handling big volumes of knowledge, is nothing new, libraries have been doing it for over 2000 plus years. They have been doing battle with the curse of Precision verse Recall factor, since collections got that big, they they had to be indexed. Of course now, every library uses an index system to be able to retrieve stored material, so if it has worked for them for over 2000 years, why cant you take advantage it for your system now.

So lets jump forward to now. The way I approach it, is that dedicate a set number of characters at the front of the file name. Especially as file names are now no longer limited to 8 characters, etc.

### • FMNI = Md4H551AUS

Type	No of Var	No of Rows	Subject	Country
M	d	4H	551	AUS
Mixed	< 30	<400	Weather	Australia

- -Type of Data (single Character)  
T for Text, N for Number, M for mixture
- -Number of Variables (single Characters)  
a 5, b 10, c20, d30, e40, f50, g100, h250, J500, k1000, l2500, m5000, N= 5000+
- -Number of rows (two characters)  
1, 2, 3, 4, 5, 6, 7, 8, 9, 0 = category representations of numbers  
D10, H100, K1000, L 100000, M1000000, G1000000000, T1000000000000
- -Subject of dataset  
Using a Library Classification subject number
- -3 letter country code where the dataset was created  
International standing for prefix for countries by UN.

# G3N1U5<sup>®</sup>

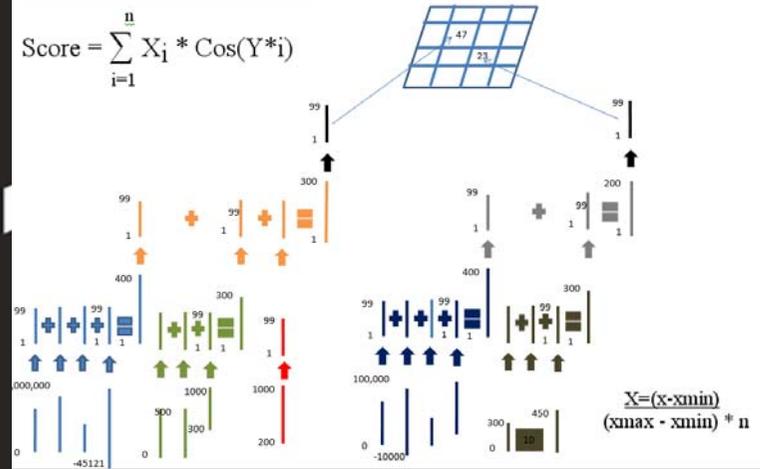
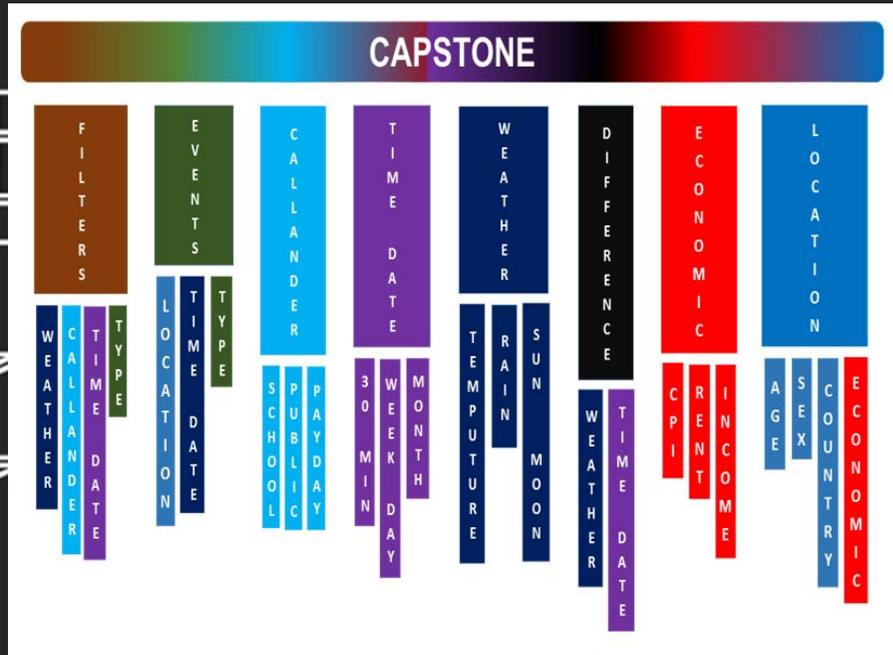
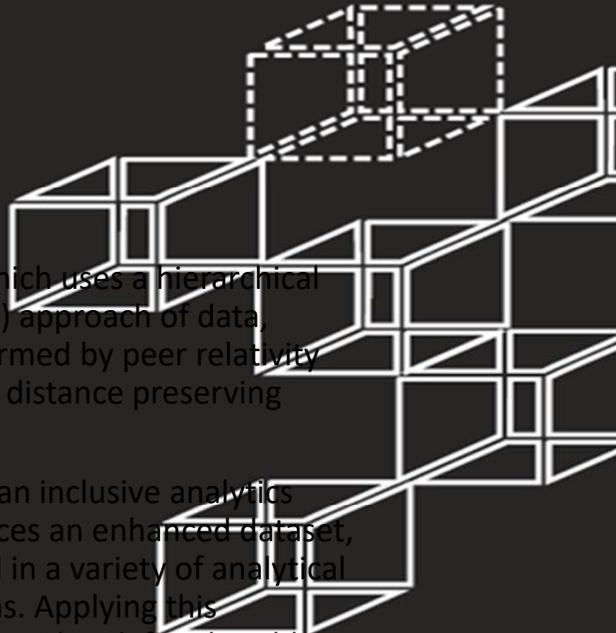
# Capstone

Capstone Modelling, which uses a hierarchical System of System (SoSE) approach of data, which has been transformed by peer relativity and is represented by a distance preserving digital hash.

The Capstone model is an inclusive analytics approach, which produces an enhanced dataset, which can then be used in a variety of analytical or reporting applications. Applying this technique usually starts with a defined problem, with a basic or limited dataset.

The dataset is then turned into an enhanced dataset, where linkage variables are identified, and a number of other datasets are used to generate metadata, as well as other data observations.

# G3N



# NLIS

George Kingsley Zipf suggested in the mid 1940's, that human beings were either wanting to be more efficient or they were just plain lazy. Hence they would shorten words over time, as they became more popular. So longer words would become shorter words. For instance TELEPHONE became PHONE, THOU became THE, and AUTOMOBILE became CAR, etc.

There have been studies, which have concluded that different Fields of Study, Employment Streams, Document Types, and Age of writers, all have different frequency usage of letters in the Alphabet. That there is a type of ranking, where the higher education a person needs to be able to participate in an activity, the less frequent letters are used in greater quantities. So a Doctor or a Lawyer will use different language with longer words, than a kindergarten teacher, who will use more common and shorter words. In Literature the first 7 letters are ( E A I T N O R ), where as in religion the order is ( E T I A N O S ), however in Chemistry the order is ( E I A O T N S ).

In the late 1980's, I theorised that you could develop a language complexity measure, that would not only use the letters of the English alphabet, but you could use the sounds of words also. So, by using the International Phonetic Alphabet, and applying it to a number of different language samples. I was able to develop a complexity measure which could be applied to a body of text, to index its complexity.

Word	Pos	LC	LC rS	PC	PCrS	Type	Emot	Arch	Sub	Keyw	Cat
on	1	2	11	2	29	preposition	0	0	0	N	Word
12/12/2012	4	10	0	0	0	noun	0	0	0	N	Date
the	15	3	13	2	54	determiner	0	0	0	N	Word
stupid	19	6	11	7	94	adjective	-15	5	153	Y	Word
car	26	3	14	2	31	noun	0	0	620	Y	Word
engine	30	6	6	5	51	noun	0	0	621	Y	Word
blew	37	4	27	3	69	verb	-20	28	662	Y	Word
up	42	2	30	2	27	adverb	0	0	0	N	Word
and	45	3	7	3	29	conjunction	0	0	0	N	Word
we	49	2	21	2	24	pronoun	0	0	0	N	Word
were	52	4	21	2	50	verb	0	0	0	N	Word
stuck	57	5	11	4	57	adjective	-5	15	994	Y	Word
there	63	5	13	2	65	adverb	0	0	0	N	Word

**Word** – The word that is being used.

**Position** – The starting position of the word.

**Letter Count** – The number of letters used in the word.

**Letter Rank Score** – The sum of the frequency ranking of the letters contained in that word.

**Phonetic Count** – The number of phonetics used in the word.

**Phonetic Rank Score** – The sum of the frequency ranking of the phonetics contained in that word.

**Language Type Index** – The type of word usage for that word in the context of the text.

**Emotional Index** – The category and emotional intensity of the word in the context of the text.

**Archetype Index** – The type of person most likely to use that word in the context of the text.

**Subject Index** – The category of subject of the word in context of the text.

**Keyword** – Is the word considered to be a keyword in the piece of text being analysed.

**Category** – Is the word a number, word, date, other, etc.

# G S N

# Mapping Events by Forecastable Factors

Income

Education

Dwellings

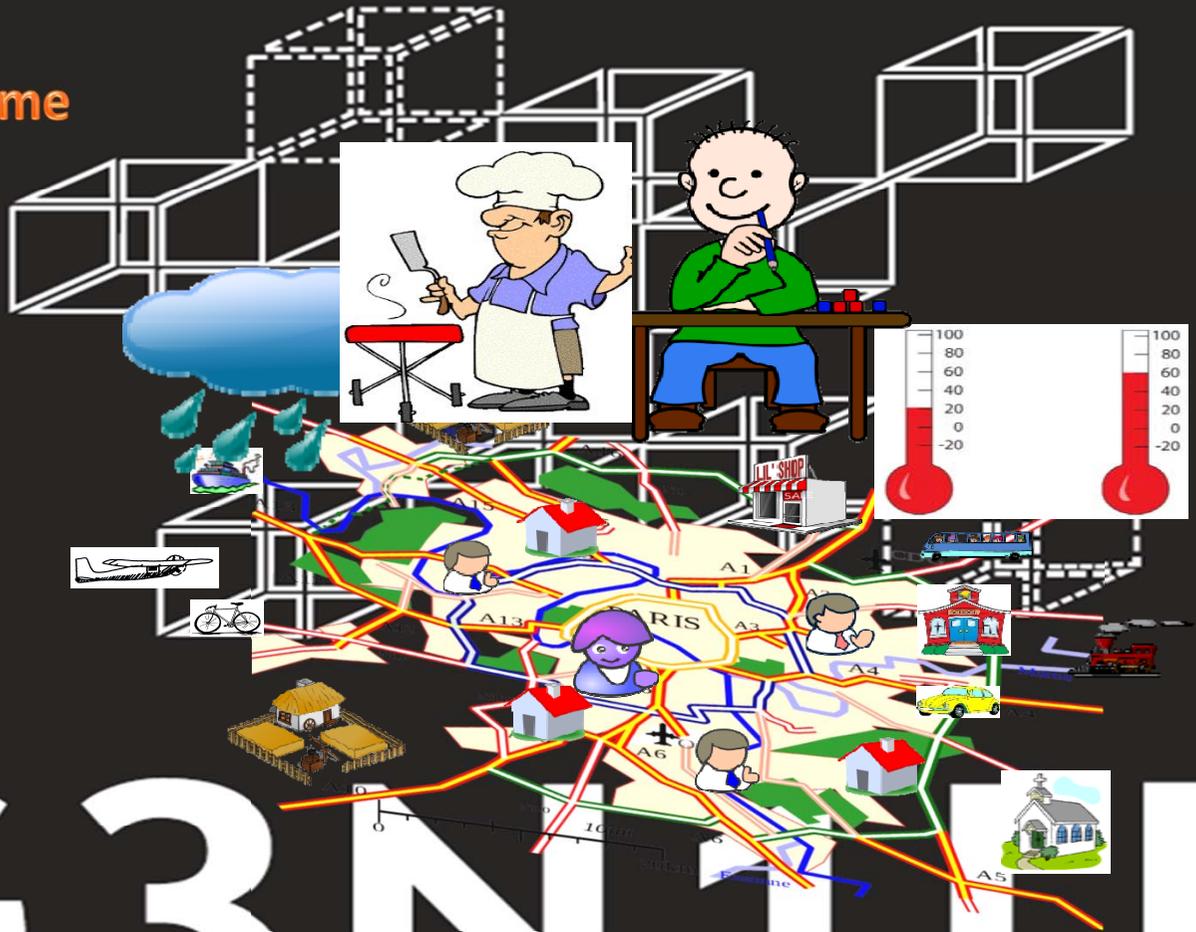
Age

Nationality

Occupation

Business

Sex



# G3N115<sup>®</sup>

Economic

Transport

Emergency

Domestic

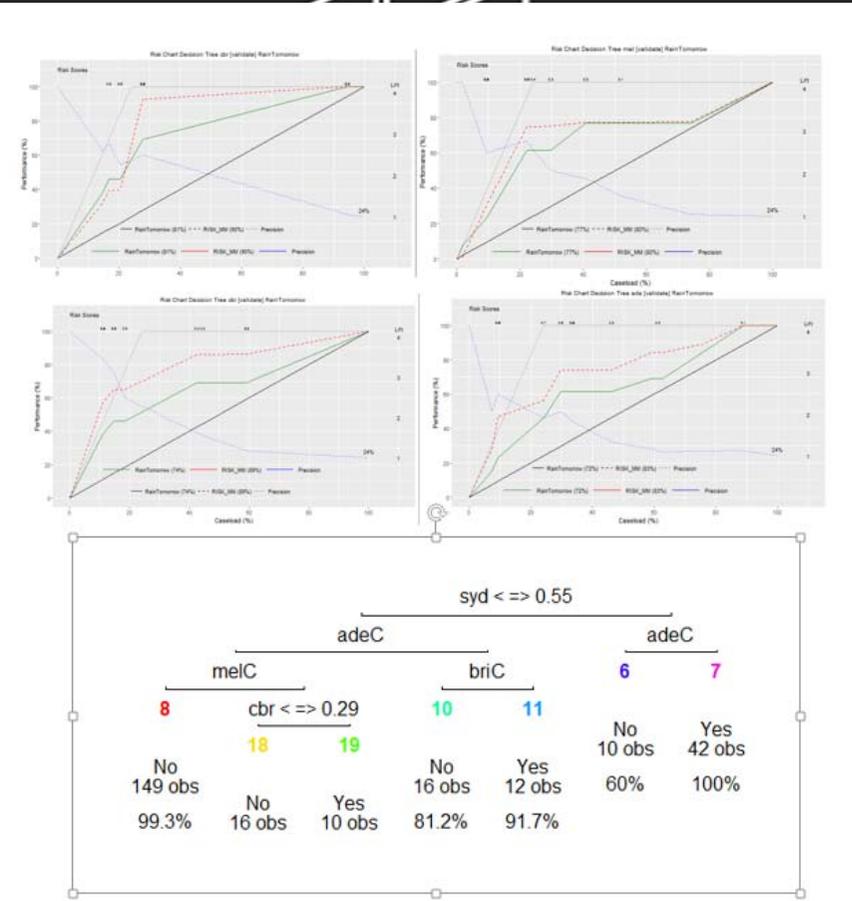
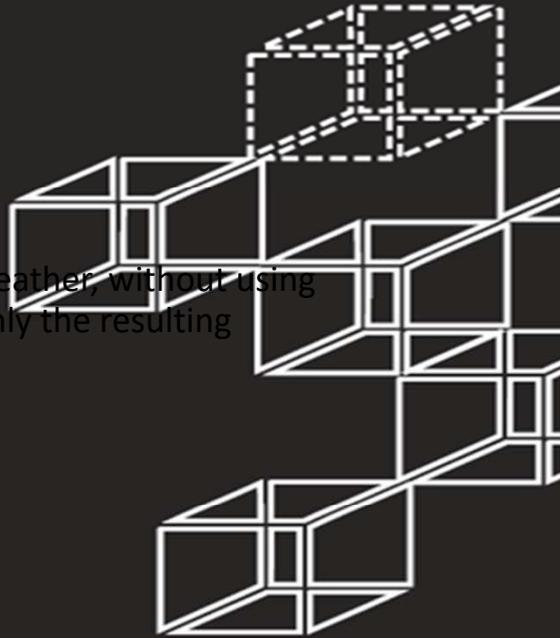
Recreation

Employment

Behavioral

# Weather

Predicting Sydney's weather, without using Sydney's variables, only the resulting observations.



# G3NTRUBS



# Dendromatrix

General purpose problem solving process which includes Total Quality Management and Brainstorming to map and solve problems through the integration of qualitative and quantitative data.

**Targeting the Problem** – Understanding the problem from all sides, history, etc

**Describe the Process** – Mapping out the problem and examine cross over effects.

**Brainstorming the Problem** – As much lateral thinking as possible focused on the problem.

**Cause and Effect** – CEDAC – Grouping common aspects of the problem.

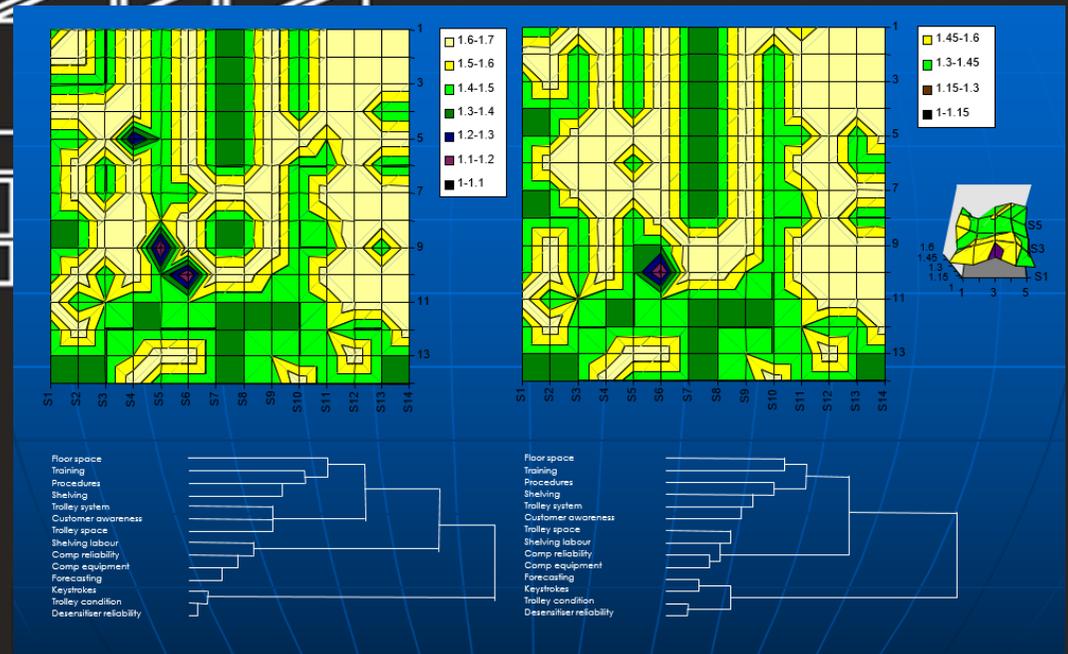
**Action Statement** – Turning negatives to positives and depersonalise comments.

**The Interaction Matrix** – Every element is measured against all others in both + and – impact.

**The Dendrogram** – A statistical driven decision tree with branches to trunk approach.

**Brainstorming the Solutions** – Lateral think a series of solutions / reflections on answers.

**Prioritising the Options** – Making an action working plan.



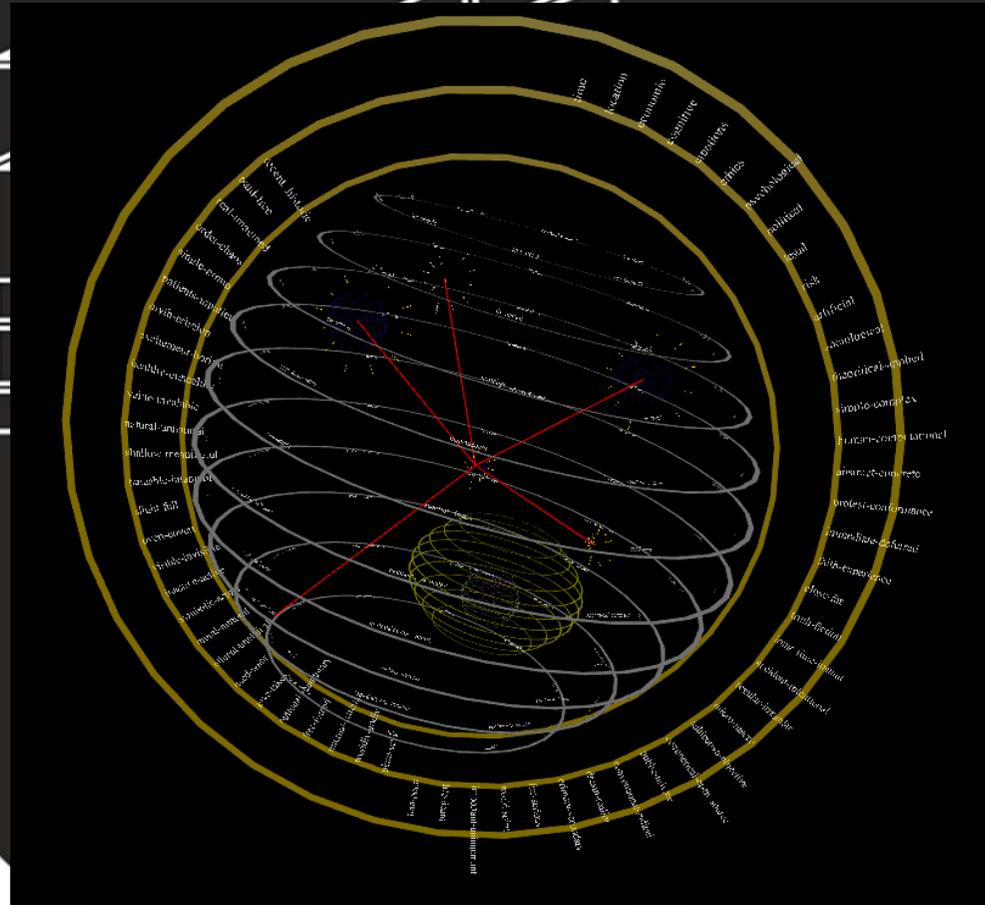
# G3N1U5

# Panocauseology

## D3 x3dom

Panocauseology is about modelling the cause and effect of everything. We live in an objective world, but we experience it through subjective sense making. When you think about everything in the universe, be it physical, behavioural, cognitive or imaginative. They all have five things in common. The first is that we can observe it. The second is that we have a name for label for them. The third is that we can classification them in a knowledge classification system. The forth is that we can observe and measure the degree of cause and effect between it and anything else in the universe. The fifth is that everything is part of a system, and nothing exists in isolation.

While there has been much research in trying to find a theory of everything or in finding physics based universal constant, none have been found. But in my view, the only interconnecting links, are those of if there is a significant and observable cause and effect. Plus when dealing in the knowledge and cognitive world through perception, observation, and imagination, many limitations fall away. The secret to this type of modelling is that it is based on categorical frameworks for the labelling and classification of everything, and an ordinal scale for the cause and effect.



## RPG

I am interested in choice analysis, and why people make the decisions they do. To undertake this research, I am using gamification back onto the gaming environment, to retrieve data from the people who play the role play game (RPG).

However, in bedded in the RPG are both a number of IQ style puzzles to solve, but within these puzzles, I bury survey questions.

- A player responds to each situation with a number of set responses from a drop down box, or radio button text box. However, based on the characters profile and their choices and pathways through the gaming environment, each survey becomes a unique experience which is tailored to that person.
- When each person starts the game, they are asked a series of demographic questions. For instance their Age, Sex, Gender, Occupation, Education Level, etc. They are also asked a series of basic profiling questions. These questions help to establish a baseline to start the game, and to also setup their choice profile, and their character profile.
- The questions are designed to serve multiple functions within the game. The first function is to move forward within the game. The second function is to look at their range of choices, to analyse which is the best next section for game play. The third function is to look at the semantic analysis of the text in the choice selection, as well as any chose dialogue options, to examine and generate a personality profile.
- The specific word usage within the game is tied to an index, which has a list of words in the rows, which is also tied to context. The columns use a number of psychographic measurement scales, taken from Dungeons and Dragons gaming platform, criminal profiling, market analysis, and behavioural economics. Each cell within the choice index table, has an individual value using a cause & effect measure. The choice index table is then integrated with the game progression table, to adjust the players profile table, etc. Each segment of the game is encoded with metadata for different types of profiles, and the most suitable game segment based on the desired outcome of the game is then presented to the player. This game is a type of advanced choose your own adventure game written in Hyperpanofiction style, which allows for game play across multiple time lines, character swapping, and direction is story time.
- The Analytics Engine that sits behind this uses transformations, profiling indexes, digital fingerprinting, and character matrixes, which I have invented to undertake my concept of this type of decision analysis. I have already used this type of analytical activity in my employment activities in law enforcement and noncompliance modelling activities. <sup>®</sup>

# G3N1U5

# Some of my linkedin

<https://www.linkedin.com/in/tony-nolan-oam-99908111/>

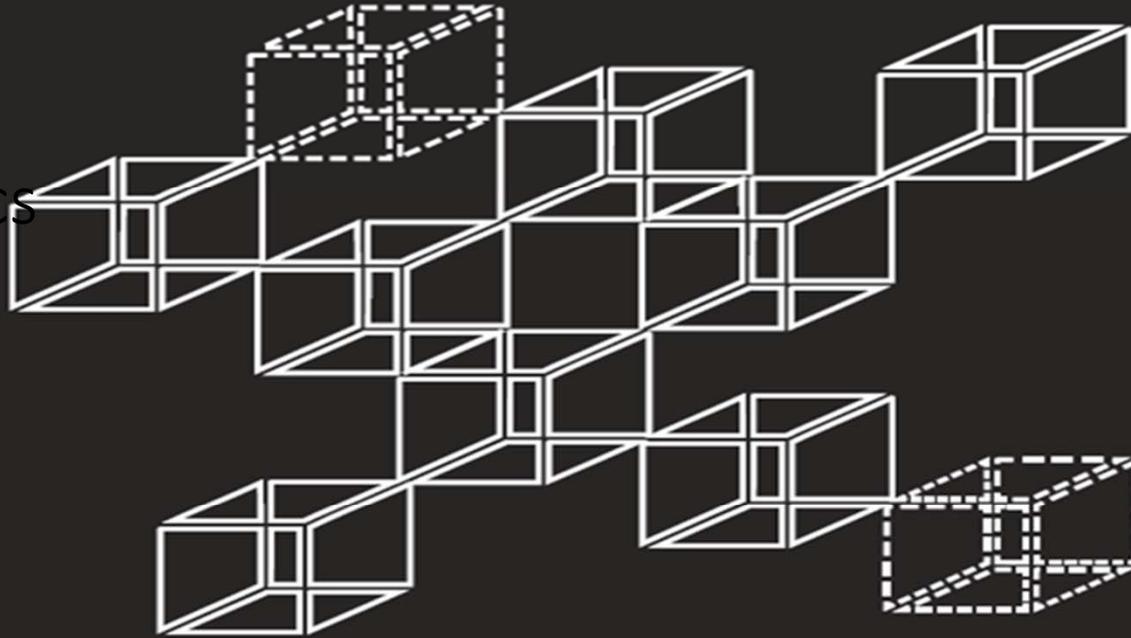
- 50 shades of data. (Confessions of a NymphoDATAct).
- Would you date your data?
- Transhumanism
- The blended Family
- Who will advocate for Artificial Life Forms
- An interconnected universe: Everything or Nothing?
- A PolyBlendedMenage?
- What are we doing tonight?
- Minding the Gap. Dealing with missing data using mirrors and open source.
- Too much Gender, not enough Sex!
- Are you a big data fool?
- Panocauseology
- Mapping reality through gaming practices that copy reality.

**G3N1U5<sup>®</sup>**

Other topics

Safety Audits

Gap analysis surveys



# G3N1U5<sup>®</sup>