Knowledge Crystallization

Card, Mackinlay, Shneiderman (1999), "Information Visualization",
Introduction to "Readings in Information Visualization: Using Vision to Think"

Knowledge crystallization

The process of:

- wanting to find out some information (your task)
- looking at all the data you can find relating to the task
- creating a schema/framework to describe the data
- packaging it so that it can be communicated to others
- … and using it to solve your task

Sue is asked by her colleagues for a laptop recommendation

She gathers a whole load of information from vendors' websites, brochures, friends and colleagues, trade magazines…

… and, in doing so, identifies several attributes (processor speed, weight, screen size etc.)

The Knowledge Crystallization process, CMS p10, edited

Sue is asked by her colleagues for a laptop recommendation

She gathers a whole load of information from vendors' websites, brochures, friends and colleagues, trade magazines…

… and, in doing so, identifies several attributes (processor speed, weight, screen size etc.)

She creates a table: columns are different products; rows are attributes (the 'schema'). In doing so, she realises

- some attributes are missing, some attributes need to be separated into two, some models have been discontinued etc., some attributes are irrelevant … so she changes the schema
- some information is missing… so she seeks it out, and adds it

The Knowledge Crystallization process, CMS p10, edited

Sue is asked by her colleagues for a laptop recommendation

She gathers a whole load of information from vendors' websites, brochures, friends and colleagues, trade magazines…

… and, in doing so, identifies several attributes (processor speed, weight, screen size etc.)

She creates a table: columns are different products; rows are attributes (the 'schema'). In doing so, she realises

- some attributes are missing, some attributes need to be separated into two, some models have been discontinued etc., some attributes are irrelevant … so she changes the schema
- some information is missing… so she seeks it out, and adds it

In doing so, she realises that some attributes are related to each other (weight, screen size); some are trade-offs (cost, processing speed)… so she changes the schema row order to highlight these relationships

The Knowledge Crystallization process, CMS p10, edited

Sue is asked by her colleagues for a laptop recommendation

She gathers a whole load of information from vendors' websites, brochures, friends and colleagues, trade magazines…

… and, in doing so, identifies several attributes (processor speed, weight, screen size etc.)

She creates a table: columns are different products; rows are attributes (the 'schema'). In doing so, she realises

- some attributes are missing, some attributes need to be separated into two, some models have been discontinued etc., some attributes are irrelevant … so she changes the schema
- some information is missing… so she seeks it out, and adds it

In doing so, she realises that some attributes are related to each other (weight, screen size); some are trade-offs (cost, processing speed)… so she changes the schema row order to highlight these relationships

She can then present this information to her colleagues in a way that **shows the insight that she herself has gained during this process**

The Knowledge Crystallization process, CMS p10, edited

	New Dell Inspiron Black	HUAWEI MateBook D	HP 15-dw1004na	Lenovo V15 ADA 15.6" Full HD Laptop	Lenovo IdeaPad	Lenovo V15 ADA
Price	£524.00	£599.99	£435.23	£399.00	£529.93	£524.40
Sold By	COMPSOLUK	Amazon.co.uk	Amazon Warehouse	MESH Computers	TEKshop	Techno world Plc
Computer Memory Size	8 GB	8 GB	8 GB	8 GB	8 GB	8 GB
Connectivity Technology	Bluetooth; USB; HDMI; Wi-Fi	Bluetooth; Wi-Fi; USB	Bluetooth; Ethernet; HDMI; USB; Wi-Fi	—	—	Wi-Fi
CPU Model	Ryzen 5 3500U	Ryzen 5 3500U	Core i3	Athlon	Core i5 Family	R Series
CPU Model Manufacturer	AMD	AMD	Intel	AMD	Intel	AMD
CPU Speed	3.7 GHz	3.4	2.1 GHz	2.4 GHz	1 GHz	2.6 GHz
Item Weight	2 kg	1.53 kg	2 kg	0 grammes	1.7 kg	2.1 kg
Screen Size	15.6 in	15.6 in	15.6 in	15.6 in	15.6 in	15.6 in
Display Technology	LED	—	—	—	—	—
Resolution	1080p	1920 x 1080	1080p	1920 x 1080	1920 x 1080	1920 x 1080
Hard Disk Description	Flash Memory Solid State	Flash Memory Solid State	Flash Memory Solid State	SSD	SSD	SSD
Hard Disk Size	256 GB	256 GB	—	256 GB	256 GB	256 GB
Operating System	Windows 10	Windows 10 Home	Windows 10 Home	Windows 10 Home	Windows 10 S	Windows 10 Home
Processor Count	1	1	2	2	4	2
Processor Description	AMD Ryzen	AMD R5, Win 10 Home	Intel Core	AMD Athlon	Intel Core i5 10th Gen	AMD Ryzen
RAM Type	DDR4 SDRAM	DDR4 SDRAM	DDR4 SDRAM	DDR4 SDRAM	DDR4 SDRAM	DDR SDRAM

https://www.amazon.co.uk/dp/B08NQ1VX6Z/ (08/07/21)

	New Dell Inspiron Black	HUAWEI MateBook D	HP 15-dw1004na	Lenovo V15 ADA 15.6" Full HD Laptop	Lenovo V15 ADA
Price	£524.00	£599.99	£435.23	£399.00	£524.40
Sold By	COMPSOLUK	Amazon.co.uk	Amazon Warehouse	MESH Computers	Techno world Plc
Computer Memory Size	8 GB	8 GB	8 GB	8 GB	8 GB
RAM Type	DDR4 SDRAM	DDR4 SDRAM	DDR4 SDRAM	DDR4 SDRAM	DDR SDRAM
Operating System	Windows 10	Windows 10 Home	Windows 10 Home	Windows 10 Home	Windows 10 Home
Connectivity Technology	Bluetooth; USB; HDMI; Wi-Fi	Bluetooth; Wi-Fi; USB	Bluetooth; Ethernet; HDMI; USB; Wi-Fi	—	Wi-Fi
CPU Model	Ryzen 5 3500U	Ryzen 5 3500U	Core i3	Athlon	R Series
CPU Model Manufacturer	AMD	AMD	Intel	AMD	AMD
CPU Speed	3.7 GHz	3.4	2.1 GHz	2.4 GHz	2.6 GHz
Item Weight	2 kg	1.53 kg	2 kg	0 grammes	2.1 kg
Screen Size	15.6 in	15.6 in	15.6 in	15.6 in	15.6 in
Display Technology	LED	—	—	—	—
Resolution	1080p	1920 x 1080	1080p	1920 x 1080	1920 x 1080
Hard Disk Description	Flash Memory Solid State	Flash Memory Solid State	Flash Memory Solid State	SSD	SSD
Hard Disk Size	256 GB	256 GB	—	256 GB	256 GB
Processor Count	1	1	2	2	2
Processor Description	AMD Ryzen	AMD R5, Win 10 Home	Intel Core	AMD Athlon	AMD Ryzen

https://www.amazon.co.uk/dp/B08NQ1VX6Z/ (08/07/21)

John needs to decide which hill races to enter this summer

He gathers all the information he can find about all the summer races, with their attributes (length, climb, record time, cost)

…. and creates a schema of rows (races), and columns (attributes)

John needs to decide which hill races to enter this summer

He gathers all the information he can find about all the summer races, with their attributes (length, climb, record time, cost)

…. and creates a schema of rows (races) and columns (attributes)

In doing so, he discovers new attributes (e.g. point-to-point *vs* loop, date, registration process), some races include more than one distance, different costs for club members etc. He changes his schema accordingly.

John needs to decide which hill races to enter this summer

He gathers all the information he can find about all the summer races, with their attributes (length, climb, record time, cost)

…. and creates a schema of rows (races) and columns (attributes)

In doing so, he discovers new attributes (e.g. point-to-point vs loop, date, registration process), some races include more than one distance, different costs for club members etc. He changes his schema accordingly.

He also identifies races that are too long/short for his interest, and removes them, and he seeks out information that he does not have.

He sees some races are held the same day, or are so close together there is insufficient time to recover from one race before the next

John needs to decide which hill races to enter this summer

He gathers all the information he can find about all the summer races, with their attributes (length, climb, record time, cost)

…. and creates a schema of rows (races) columns (attributes)

In doing so, he discovers new attributes (e.g. point-to-point vs loop, date, registration process), some races include more than one distance, different costs for club members etc. He changes his schema accordingly.

He also identifies races that are too long/short for his interest, and removes them, and he seeks out information that he does not have.

He sees some races are held the same day, or are so close together there is insufficient time to recover from one race before the next

He realises that the race dates are an important attribute, and so changes his schema to a timetable representation, where the number of days between each race is clearly seen

This makes it easier to find a sensible schedule through the summer that identifies the races he will take part in.

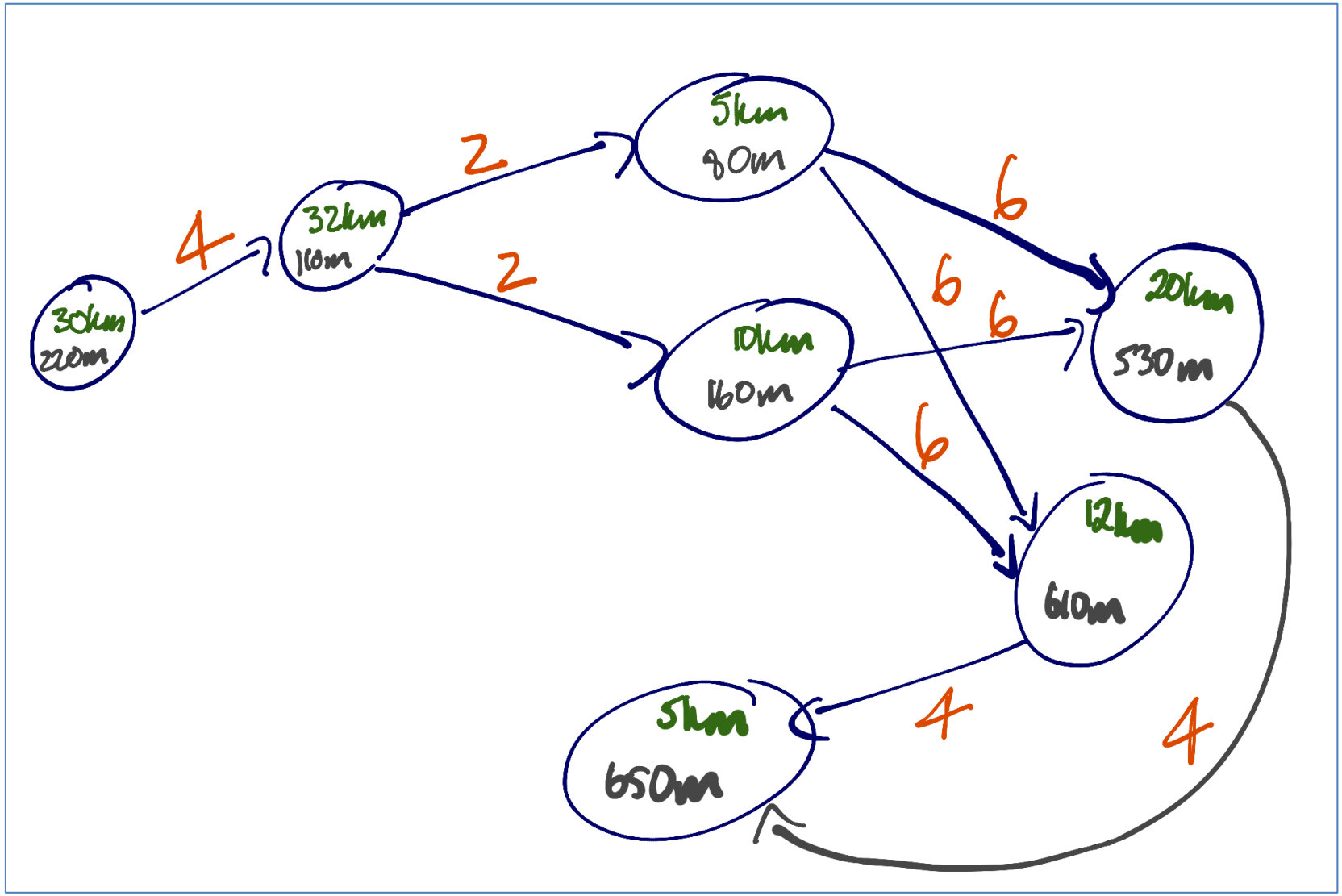
John's first schema

	distance	climb	cost	route	date
A	30km	220m	£10	P2P	4 June
B	12km	610m	£15	loop	16 June
C	5km	650m	£5	P2P	20 June
D	32km	110m	£7	loop	8 June
Ea	5km	80m	£8	loop	10 June
Eb	10km	160m	£10	loop	10 June
F	20km	530m	£12	loop	16 June

John's second schema

4	5	6	7	8	9	10	11	12	13
A (30km,220m)				D (32km,110m)		Ea (5km,80m) Eb (10km,160m)			
14	15	16	17	18	19	20	21	22	23
		F (20km,530m) B (12km,610m)				C (5km, 650m)			

John's third schema



The steps of Knowledge Crystallization

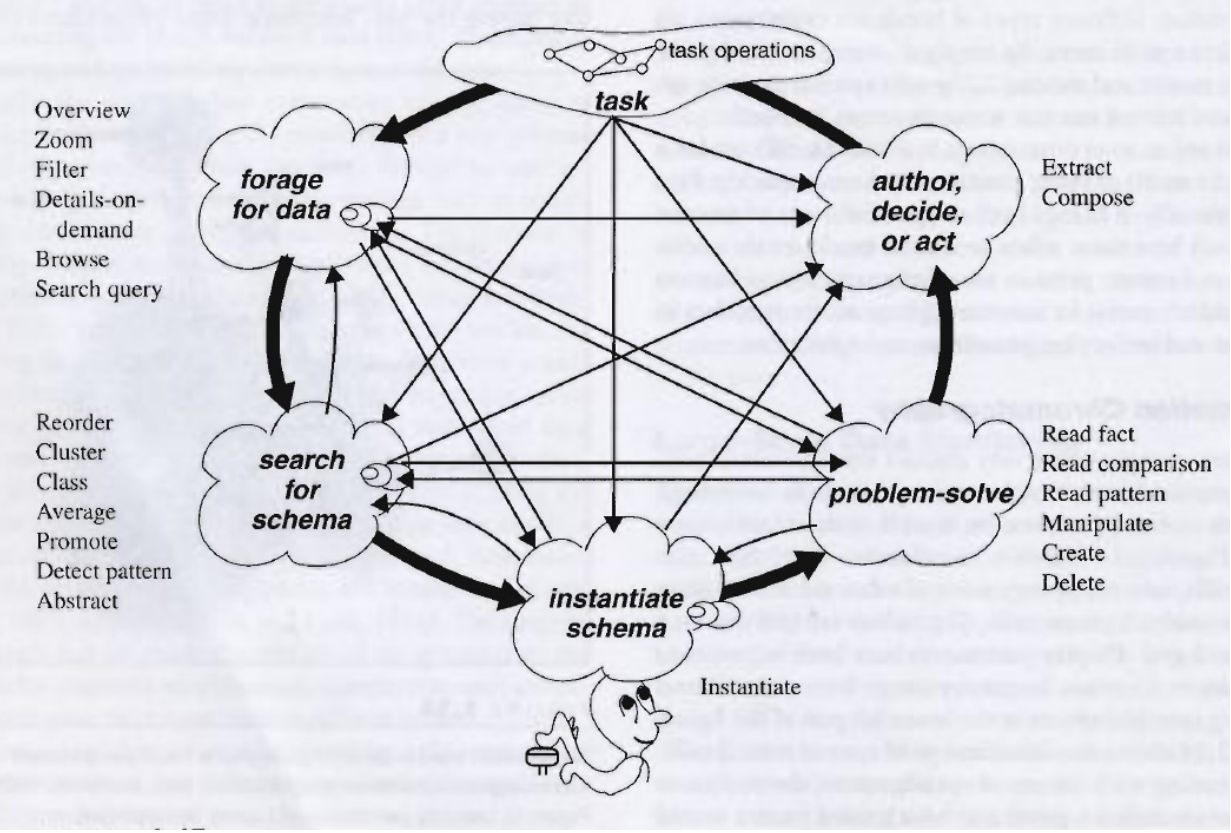
1. Forage for information/data
2. Create schema – the framework for holding the information
3. Instantiate the schema by putting the information in it
    a. solve the problems that emerge: change or extend schema, add or remove information
    b. manipulate the schema to reveal insight about the information gathered along the way
    c. repeat until the most effective and efficient schema is found
4. Present the information to others, using the best schema
5. Use it to solve your task

Requirements for Knowledge Crystallization:

data, task, existing or potential schema

CMS p10, edited

Knowledge crystallisation



Information Visualisation can support all stages

CMS p10, edited

Knowledge Crystallization