

# Metamodels and Feature Models: Complementary Approaches to Formalize Product Comparison Matrices

---

Guillaume Bécan

Ph.D defense

September 23, 2016

Jury

Laurence Duchien - University of Lille 1

Andrzej Wąsowski - IT University of Copenhagen

Felienne Hermans - Delft University of Technology

Olivier Ridoux - University of Rennes 1

Benoit Baudry - Inria Rennes

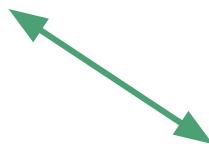
Mathieu Acher - University of Rennes 1

# Context



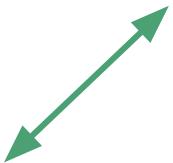
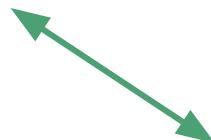
# Context

**SONY**  
**Canon**



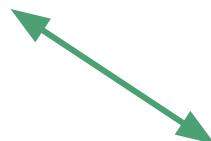
# Context

**SONY**  
**Canon**

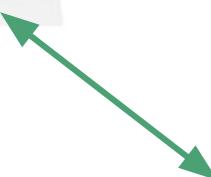


# Context

**SONY**  
**Canon**

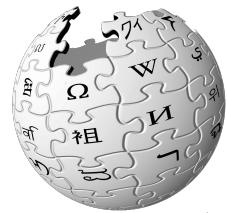


which one ?



# Context

**SONY**  
**Canon**



which one ?



# Product Comparison Matrix (PCM)

Comparison of digital SLRs



Features

Brand	Model name	Sensor size	Effective megapixels	Movie mode	Memory card type	Weight (g) <sup>[1]</sup>	In production
Canon	1D X Mark II	Full frame	20.2	yes	CF+CFast	1530	yes
Canon	1D X	Full frame	18.1	yes	CF (2x)	1530	no
Canon	1Ds Mark III	Full frame	21.1	no	CF+SD	1205	no
Canon	1D Mark IV	APS-H	16.1	yes	CF+SD	1180	no
Canon	5D Mark III	Full frame	22.3	yes	CF+SD	950 (860 without battery)	yes
Canon	5D Mark II	Full frame	21.1	yes	CF	810	no
Canon	6D	Full frame	20.2	yes	SD	755 (680 without battery)	no
Canon	7D Mark II	APS-C	20.2	yes	CF+SD	910 (820 without battery)	yes

Products

Cell

The **product** “5D Mark III” has the **feature** “movie mode”

# Product Comparison Matrix (PCM)

Can you find the **lightest** camera with

- > 20 megapixels
- Movie mode
- Full frame sensor



which one ?



Brand	Model name	Sensor size	Effective megapixels	Movie mode	Memory card type	Weight (g) <sup>[1]</sup>	In production
Canon	<a href="#">1D X Mark II</a>	Full frame	20.2	yes	CF+CFast	1530	yes
Canon	<a href="#">1D X</a>	Full frame	18.1	yes	CF (2x)	1530	no
Canon	<a href="#">1Ds Mark III</a>	Full frame	21.1	no	CF+SD	1205	no
Canon	<a href="#">1D Mark IV</a>	APS-H	16.1	yes	CF+SD	1180	no
Canon	<a href="#">5D Mark III</a>	Full frame	22.3	yes	CF+SD	950 (860 without battery)	yes
Canon	<a href="#">5D Mark II</a>	Full frame	21.1	yes	CF	810	no
Canon	<a href="#">6D</a>	Full frame	20.2	yes	SD	755 (680 without battery)	no
Canon	<a href="#">7D Mark II</a>	APS-C	20.2	yes	CF+SD	910 (820 without battery)	yes

# Product Comparison Matrix (PCM)

Can you find the **lightest** camera with

- > 20 megapixels
- Movie mode
- Full frame sensor



which one ?



Brand	Model name	Sensor size	Effective megapixels	Movie mode	Memory card type	Weight (g) <sup>[1]</sup>	In production
Canon	<a href="#">1D X Mark II</a>	Full frame	20.2	yes	CF+CFast	1530	yes
Canon	<a href="#">1D X</a>	Full frame	18.1	yes	CF (2x)	1530	no
Canon	<a href="#">1Ds Mark III</a>	Full frame	21.1	no	CF+SD	1205	no
Canon	<a href="#">1D Mark IV</a>	APS-H	16.1	yes	CF+SD	1180	no
Canon	<a href="#">5D Mark III</a>	Full frame	22.3	yes	CF+SD	950 (860 without battery)	yes
Canon	<a href="#">5D Mark II</a>	Full frame	21.1	yes	CF	810	no
Canon	<a href="#">6D</a>	Full frame	20.2	yes	SD	755 (680 without battery)	no
Canon	<a href="#">7D Mark II</a>	APS-C	20.2	yes	CF+SD	910 (820 without battery)	yes

# Product Comparison Matrix (PCM)

Brand	Model name	Sensor size	Effective megapixels	Lens mount	Viewfinder type	Viewfinder coverage (% of the frame)	Histogram zones	Focus points	Lowest ISO	Highest ISO	DxO Mark sensor score	DxO ISO performance	Cont shr (B5)	LCD size (in)	Live view	Movie mode	Memory card type	Dimensions (mm)	Weight (g)	Announced (date)	In production	Reference
Canon	1D X Mark II	Full frame	20.2	EF	Pentaprism	100	216	61	60	409600	TBA	TBA	14	3.2	yes	CF+EFest	159x168x83	1530	Feb 2016	yes	B101 B102	
Canon	1D X	Full frame	18.1	EF	Pentaprism	100	252	61	60	204800	82	2766	14	3.2	yes	CF (2D)	159x164x83	1530	Oct 2011	no	B101 B102	
Canon	1Dx Mark III	Full frame	21.1	EF	Pentaprism	100	63	46	60	3200	80	1663	5	3	yes	CF+SD	159x160x80	1205	Aug 2007	no	B101 B102	
Canon	1Dx Mark IV	APS-C	16.1	EF	Pentaprism	100	63	46	60	102400	74	1320	10	3	yes	CF+SD	156x157x80	1180	Oct 2009	no	B101 B102	
Canon	5D Mark III	Full frame	22.3	EF	Pentaprism	100	63	61	60	102400	81	2293	6	3.2	yes	CF+SD	152x117x77	950 (650 without battery)	Mar 2012	yes	B101 B102	
Canon	5D Mark II	Full frame	21.1	EF	Pentaprism	98	36	9	60	25600	79	1815	2.9	3	yes	CF	152x114x75	810	Sep 2008	no	B110 B111	
Canon	6D	Full frame	20.2	EF	Pentaprism	97	63	11	60	102400	82	2340	4.6	3	yes	SD	146x111x71	765 (650 without battery)	Sep 2012	no	B110	
Canon	7D Mark II	APS-C	20.2	EF, EF-S	Pentaprism	100	252	65	100	61200	70	1082	10	3	yes	CF+SD	149x112x78	910 (820 without battery)	Sep 2014	yes	B110 B111	
Canon	7D	APS-C	18.0	EF, EF-S	Pentaprism	100	63	19	100	12800	66	854	8	3	yes	CF	148x111x74	820	Sep 2009	no	B110 B111	
Canon	80D	APS-C	24.2	EF, EF-S	Pentaprism	100	63	45	100	25600	TBA	TBA	7.0	3	yes	SD	139x106x79	730	Feb 2016	yes	B110 B111	
Canon	70D	APS-C	20.2	EF, EF-S	Pentaprism	98	63	19	100	25600	68	926	7.0	3	yes	SD	139x106x79	765	Jul 2013	no	B110 B111	
Canon	60D	APS-C	18.0	EF, EF-S	Pentaprism	96	63	9	100	12800	66	813	5.3	3	yes	SD	146x106x79	765	Aug 2010	yes	B110	
Canon	50D	APS-C	15.1	EF, EF-S	Pentaprism	96	36	9	100	12800	63	696	6.3	3	yes	CF	146x108x74	822	Oct 2008	no	B110 B111	
Canon	40D	APS-C	10.1	EF, EF-S	Pentaprism	96	36	9	100	3200	64	703	6.5	3	yes	CF	146x108x74	740	Aug 2007	no	B110	
Canon	30D	APS-C	8.2	EF, EF-S	Pentaprism	96	36	9	100	3200	69	706	6	2.8	no	no	CF	144x106x74	765	Feb 2006	no	B110 B111
Canon	20D	APS-C	8.2	EF, EF-S	Pentaprism	96	36	9	100	3200	62	721	6	1.8	no	no	CF	144x106x72	770	Aug 2004	no	B110
Canon	10D	APS-C	6.3	EF	Pentaprism	96	36	7	100	3200	67	671	3	1.8	no	no	CF	150x107x75	790 (without battery)	Feb 2003	no	B110
Canon	700D, Rebel T6s, EOS 8000D	APS-C	24	EF, EF-S	Pentamirr	96	63	19	100	25600	70	916	6	3	yes	SD	132x101x78	665g	Feb 2015	yes		
Canon	700D, Rebel T6i, Kiss X8i	APS-C	24	EF, EF-S	Pentamirr	96	63	19	100	25600	71	919	6	3	yes	SD	132x101x78	655g	Feb 2015	yes		
Canon	700D, Rebel T5i, Kiss X7i	APS-C	18.0	EF, EF-S	Pentamirr	96	63	9	100	25600	61	681	6.0	3	yes	SD	134x100x79	580	Mar 2013	yes	B110 B111	
Canon	650D, Rebel T4i, Kiss X6i	APS-C	18.0	EF, EF-S	Pentamirr	95	63	9	100	25600	62	722	6.0	3	yes	SD	134x100x79	575	Jun 2012	no	B110 B111	
Canon	650D, Rebel T3i, Kiss X5	APS-C	18.0	EF, EF-S	Pentamirr	95	63	9	100	12800	65	793	3.7	3	yes	SD	133x100x80	515 (without battery)	Feb 2011	yes	B110 B111	
Canon	650D, Rebel T2i, Kiss X4	APS-C	18.0	EF, EF-S	Pentamirr	95	63	9	100	12800	66	784	3.7	3	yes	SD	129x100x78	530	Feb 2010	no	B110 B111	
Canon	550D, Rebel T1i, Kiss X3	APS-C	15.1	EF, EF-S	Pentamirr	95	36	9	100	12800	63	663	3.4	3	yes	SD	129x98x62	480 (without battery)	Mar 2009	no	B110 B111	
Canon	450D, Rebel Xti, Kiss X2	APS-C	12.2	EF, EF-S	Pentamirr	95	36	9	100	1600	61	692	3.6	3	yes	SD	129x98x62	476 (without battery)	Jan 2008	no	B110	
Canon	400D, Digital Rebel XT, Kiss Digital X	APS-C	10.1	EF, EF-S	Pentamirr	95	36	9	100	1600	644	644	3	2.5	no	no	CF	127x95x65	510 (without battery)	Aug 2006	no	B110
Canon	350D, Digital Rebel XT, Kiss Digital II	APS-C	8.0	EF, EF-S	Pentamirr	95	36	7	100	1600	65	728	2.8	1.8	no	no	CF	127x84x64	540 (405 without battery)	Feb 2005	no	B110
Canon	300D, Digital Rebel Kiss Digital	APS-C	6.3	EF, EF-S	Pentamirr	95	36	7	100	1600	65	726	2.6	1.8	no	no	CF	142x99x72	654 (550 without battery)	Aug 2003	no	B110
Canon	1000D, Rebel SL1, Kiss X7	APS-C	18.0	EF, EF-S	Pentamirr	95	63	9	100	25600	63	741	4	3	yes	SD	117x91x69	407 (370 without battery)	Mar 2013	yes	B110 B111	
Canon	1000D, Rebel TS1, Kiss X70	APS-C	18.0	EF, EF-S	Pentamirr	95	63	9	100	12800	63	724	3.0	3	yes	SD	130x100x78	480 (435 without battery)	Feb 2014	yes	B110 B111	
Canon	1100D, Rebel T3, Kiss X80	APS-C	12.2	EF, EF-S	Pentamirr	95	63	9	100	6400	62	765	3	2.7	yes	SD	139x100x78	495	Feb 2014	yes	B110 B111	
Canon	1000D, Rebel XS, Kiss F	APS-C	10.1	EF, EF-S	Pentamirr	95	36	7	100	1600	62	719	3	2.5	yes	SD	126x98x62	502	Jun 2008	no	B110 B111	
Ilum	D4	Full frame	16.2	F-mount	Pentaprism	100	91000	61	60	294800	89	2965	11	3.2	yes	CF+QD	160x157x81	1180 (without battery)	Jan 2012	yes	B110 B111	
Ilum	D3X	Full frame	24.4	F-mount	Pentaprism	100	2005	61	60	6400	88	1992	6	3	yes	CF (2D)	160x157x88	1220	Dec 2008	yes	B110 B111	
Ilum	D3	Full frame	12.1	F-mount	Pentaprism	100	2005	61	100	102400	82	3263	9	3	yes	CF (2D)	160x157x88	1246	Oct 2009	no	B110 B111	
Ilum	D800	Full frame	36.3	F-mount	Pentaprism	100	91000	61	60	25600	95 (D)	2883 (8000)	4	3.2	yes	CF+SD	146x122x82	900 (without battery)	Feb 2012	yes	B110 B111 B110	
Ilum	D700	Full frame	12.1	F-mount	Pentaprism	95	2005	61	100	25600	80	2303	6	3	yes	CF	147x123x77	995	Jul 2008	no	B110 B111	
Ilum	D600	Full frame	24.3	F-mount	Pentaprism	100	2016	39	100	25600	94	2880	6.5	3.2	yes	SD (2D)	141x119x82	760 (no battery)	Sep 2012	yes	B110 B111	
Ilum	D500	APS-C	12.3	F-mount	Pentaprism	100	2005	61	100	6400	70	787	6	3	yes	CF+SD	147x114x74	918	Jul 2009	yes	B110 B111 B110	
Ilum	D7000	APS-C	16.2	F-mount	Pentaprism	100	2016	39	100	25600	80	1167	6	3	yes	SD (2D)	132x106x77	690 (without battery)	Sep 2010	yes	B110	
Ilum	D90	APS-C	12.3	F-mount	Pentaprism	95	420	11	100	6400	73	977	4.5	3	yes	SD	132x103x77	620 (without battery)	Aug 2008	no	B110 B111	
Ilum	D5100	APS-C	16.2	F-mount	Pentamirr	95	420	11	100	25600	80	1183	4	3	yes	SD	129x107x79	660 (without battery)	Apr 2011	yes	B110 B111	
Ilum	D5000	APS-C	12.3	F-mount	Pentamirr	95	420	11	100	6400	72	868	4	2.7	yes	SD	127x104x80	660 (without battery)	Apr 2009	no	B110 B111	
Ilum	D3200	APS-C	24.2	F-mount	Pentamirr	95	420	11	100	12800	81	1111	4	3	yes	SD	125x96x79	695	Apr 2011	yes	B110 B111	
Ilum	D3100	APS-C	14.2	F-mount	Pentamirr	95	420	11	100	12800	67	919	3	3	yes	SD	124x96x745	505	Aug 2010	yes	B110 B111	
Ilum	D3000	APS-C	10.2	F-mount	Pentamirr	95	420	11	100	3200	62	663	3	3	no	no	SD	126x97x64	636	Jul 2009	no	B110
Olympus	E-6	Four Thirds system	12.3	4/3 system	Pentaprism	100	49	11	100	6400	66	619	5	3.0	yes	CF, SD	142x116x78	810 (without battery)	Sep 2010	yes	B110	
Olympus	E-3	Four Thirds system	10.1	4/3 system	Pentaprism	100	49	11	100	3200	66	575	5	2.5	yes	CF, SD	142x116x78	800 (without battery)	Oct 2007	no	B110 B111	
Olympus	E-30	Four Thirds system	12.3	4/3 system	Pentaprism	98	49	11	100	3200	65	630	5	2.7	yes	CF, SD	142x116x78	790	Feb 2009	yes	B110 B111	
Olympus	E-420	Four Thirds system	12.3	4/3 system	Pentamirr	95	49	7	100	3200	65	636	4	2.7	yes	CF, SD	139x104x69	476 (without battery)	Feb 2009	yes	B110 B111	
Pentax	K-01	APS-C	19.5	PK mount	Pentaprism	98	77	11	100	1600	82	1262	1	3	no	SD (2D)	156x117x119	1400	Mar 2010	yes	B110 B111 - Pentax K-01 54.0MP (PRIME) / K-01 54.0MP (K-Mount)	
Pentax	K-0	APS-C	16.3	K mount	Pentaprism	100	77	11	80	61200	82	1162	7	3	yes	SD (2D)	131x97x73	760	Sep 2010	no	B110 B111	
Pentax	K-7	APS-C	14.6	K mount	Pentaprism	100	77	11	80	6400	61	936	5.2	3	yes	SD	131x97x73	760	May 2009	no	B110 B111	
Pentax	K-0	APS-C	16.3	K mount	Pentaprism	100	77	11	100	61200	79	1120	6	3	yes	SD (2D)	139x107x71	650	June 2013	yes	B110 B111	
Pentax	K-000	APS-C	16.3	K mount	Pentaprism	100	77	11	100	61200	79	1087	6	3	yes	SD (2D)	139x107x71	646	June 2013	yes	B110 B111	
Pentax	K-00	APS-C	16.3	K mount	Pentaprism	100	77	11	100	25600	79	1119	6	3	yes	SD (2D)	131x97x73	640	May 2012	yes	B110 B111	
Pentax	K-0	APS-C	12.4	K mount	Pentamirr	96	46	11	100	25600	72	766	6.0	3.0	yes	SD	122x91x67	516 (without battery)	Sep 2010	no	B110 B111	
Pentax	K-0	APS-C	12.4	K mount	Pentamirr	96	46	11	100	12800	72	811	4.7	2.7	yes	SD	122x91x67	580	Sep 2009	no	B110 B111	
Pentax	K-020	APS-C	14.6	K mount	Pentamirr	96	46	11	100	6400	65	639	3	2.7	yes	SD	141x101x70	800	Jan 2008	no	B110	
Sigma	SD1 Merrill	15 x 3	SA mount	Pentaprism	96	77	11	100	6400	N/A	N/A	6	3	no	no	CF	146x113x89	790	February 2012	yes	B110 B111	
Sony	Alpha 900	Full frame	24.6	Sony Alpha 900 A	Pentaprism	100	40	9	100	6400	79	1431	5	3	no	no	CF, IS	156x117x82	895	Sep 2008	yes	B110
Sony	Alpha 850	Full frame	24.6	Sony Alpha 850 A	Pentaprism	98	40	9	100	6400	79	1415	3	3	no	no	CF, IS	156x117x82				

# Product Comparison Matrix (PCM)

PCMs can be too complex for humans

6 3 6 1	4 9 5 2 9 3 9 5 1	9 3 7 8 4 3 0 1 9	3 9 1 3 0 7 2 0 0	2 7 5 0 4
7 5 2 6	8 7 4 3 9 7 0 2 2 7	4 9 6 7 4 7 7 9 6	7 1 5 3 6 8 2 2 5	3 0 8 4 9
7 7 9 4	6 1 2 7 2 9 4 3 1	2 5 5 8 4 7 0 4 0	8 0 4 7 6 1 0 3 0	0 7 6 8 8
7 8 9 4	3 6 4 1 4 0 4 1 7	5 6 8 1 6 8 0 0 4	0 7 5 6 1 4 8 3 6	8 5 6 3 1
1 9 5 2	0 2 1 3 1 5 9 6 9	1 0 4 6 7 9 9 4 5	1 7 7 9 0 2 0 1 9	2 4 1 7 7
3 9 1 1	4 3 7 0 7 0 7 6 8	0 6 4 6 4 0 4 0	2 3 1 5 1 1 5 4 4	8 3 0 7 8
4 3 0 1	9 4 5 0 7 2 8 9 0	6 2 7 9 7 5 2 7 2	0 4 4 0 5 1 7 1 4	2 8 6 4 5
4 1 6 8	6 0 0 3 8 5 0 3 4	2 3 1 4 4 4 7 7 5	3 2 1 2 1 6 4 6 8	6 9 0 6 3
8 5 7 7	8 0 8 6 0 8 4 9 0	4 0 3 3 9 0 5 9 7	8 3 9 2 8 5 7 8 3	4 6 8 7 1
1 8 9 2	3 8 4 6 5 5 2 5 8	0 SYSTEM FAILURE	4 1 0 3 0 9 0 8 4	2 6 1 2 5
0 7 3 9	7 1 2 7 8 9 7 9 8	3 2 3 9 8 0 3 6 0	5 2 8 9 0 0 8 9 6	9 0 5 3 3
4 3 5 3	9 5 1 5 2 0 4 7 2	5 1 9 8 7 8 2 4 4	3 4 0 4 0 9 1 0 6	2 6 2 1 3
4 0 7 0	0 1 8 5 0 3 8 1 6	8 7 0 0 5 2 4 7 9	4 2 1 7 4 9 1 0 3	8 6 9 1 3
8 9 0 9	1 2 5 0 0 3 6 5 2	7 5 9 7 8 5 6 5 3	3 8 3 4 6 4 3 6 6	7 9 9 8 9
0 8 5 3	1 8 3 2 3 2 5 9 2	4 4 2 0 7 6 4 8 8	6 2 4 9 4 5 6 9 5	3 9 2 9 3
3 6 5 1	2 8 5 6 3 6 5 2 5	6 0 4 0 3 3 3 3 3	0 3 4 4 4 4 2 2 0	6 4 8 7 7
2 0 7 8	1 0 7 2 8 5 7 8 0	8 9 3 3 7 1 3 4 9	7 7 9 6 1 1 2 5 6	4 2 9 2 4
2 0 4 3	1 5 4 3 8 2 7 3 0	3 2 2 2 0 2 5 0 6	2 1 9 9 9 1 5 9 4	2 9 1 7 6

# Software to the rescue



DSLR

Interchangeable Lens

Resolution: 20 MP

Optical Zoom: 1x

Max ISO: 400

Display: 1.4 inch

Video: 0 x 0

Weight: 81 ounces

Price: \$14 - \$30000

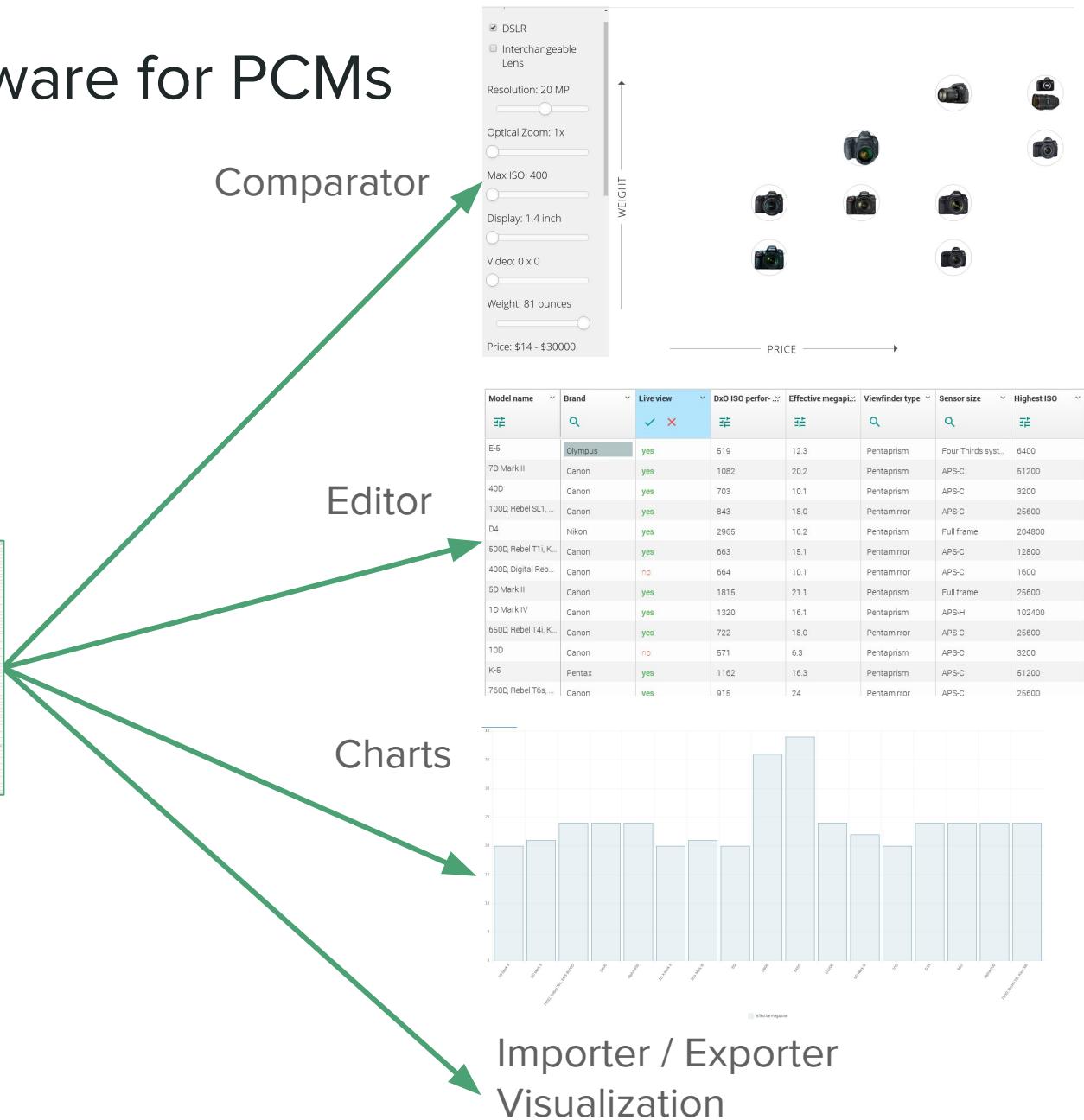
WEIGHT ↑

PRICE →



# Engineering Software for PCMs

# Product comparison matrices



# Diversity in PCMs = Complexity

Billions of PCMs on the Web

Numerous contributors

Organization of information

Format



WIKIPEDIA  
*The Free Encyclopedia*



Brand	Model name	Sensor size	Effective megapixels	Movie mode	Memory card type	Weight (g) <sup>[1]</sup>	In production
Canon	1D X Mark II	Full frame	20.2	yes	CF+CFast	1530	yes
Canon	1D X	Full frame	18.1	yes	CF (2x)	1530	no
Canon	1Ds Mark III	Full frame	21.1	no	CF+SD	1205	no
Canon	1D Mark IV	APS-H	16.1	yes	CF+SD	1180	no
Canon	5D Mark III	Full frame	22.3	yes	CF+SD	950 (860 without battery)	yes
Canon	5D Mark II	Full frame	21.1	yes	CF	810	no
Canon	6D	Full frame	20.2	yes	SD	755 (680 without battery)	no
Canon	7D Mark II	APS-C	20.2	yes	CF+SD	910 (820 without battery)	yes

# Diversity in PCMs = Complexity

Content of cells

Brand	Model name	Sensor size	Effective megapixels	Movie mode	Memory card type	Weight (g) <sup>[1]</sup>	In production
Canon	1D X Mark II	Full frame	20.2	yes	CF+CFast	1530	yes
Canon	1D X	Full frame	18.1	yes	CF (2x)	1530	no
Canon	1Ds Mark III	Full frame	21.1	no	CF+SD	1205	no
Canon	1D Mark IV	APS-H	16.1	yes	CF+SD	1180	no
Canon	5D Mark III	Full frame	22.3	yes	CF+SD	950 (860 without battery)	yes
Canon	5D Mark II	Full frame	21.1	yes	CF	810	no
Canon	6D	Full frame	20.2	yes	SD	755 (680 without battery)	no
Canon	7D Mark II	APS-C	20.2	yes	CF+SD	910 (820 without battery)	yes

Number

Boolean

Enumeration

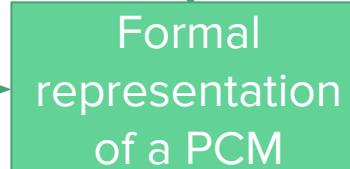
Multiple values

Condition

Gap between the representation of a PCM and its human interpretation

# How to Engineer Software for PCMs?

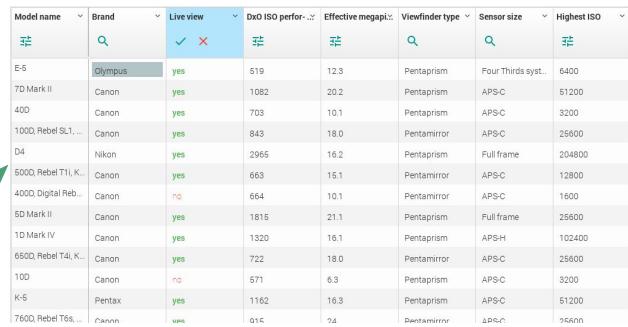
# Billions of product comparison matrices



## Comparator



Editor



## Charts

## Importer / Exporter Visualization

# How to formalize product comparison matrices?

# State of the Art

A product comparison matrix is a...

## Table

Relational model (Codd, 1970)

Wang's model (Wang, 1996)

Ontologies (Tijerino, 2005; Wang, 2012)

## Spreadsheet

ClassSheet (Engels, 2005)

Class diagrams (Hermans, 2010)

Data flow diagrams (Hermans, 2011)

Type system (Abraham, 2007)

**No formalism dedicated to PCMs**

# State of the Art

**Product comparison matrix** = view on a product line



**Product line** = set of products presenting commonalities and variabilities

Feature model (Kang, 1990)

+ extensions (Czarnecki, 2004; Classen, 2011; Czarnecki, 2008)

Feature model synthesis from

- Propositional formula (Andersen, 2012; She, 2011; Janota, 2008; Acher, 2013)
- Source code (Nadi, 2014; Vacchi, 2014)
- Textual requirements (Alves, 2008; Chen, 2005; Niu, 2009; Weston, 2009)
- Product descriptions (Davril, 2013)
- Product comparison matrices (Acher, 2012)

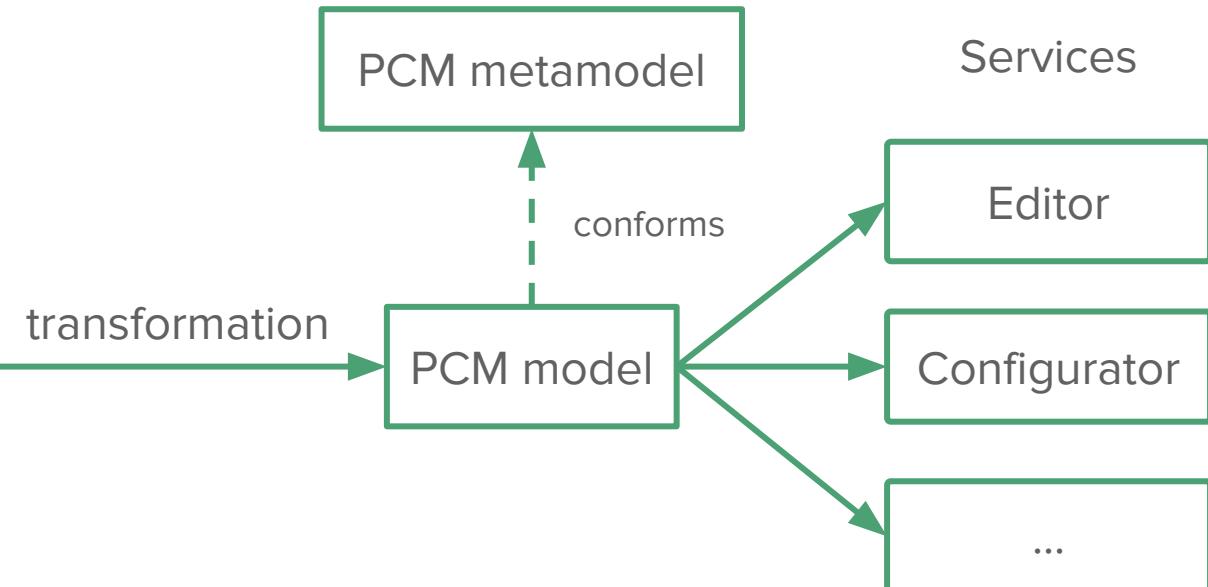
**Limited support of product comparison matrices**

# Complementary Approaches to Formalize PCMs

Billions of product comparison matrices

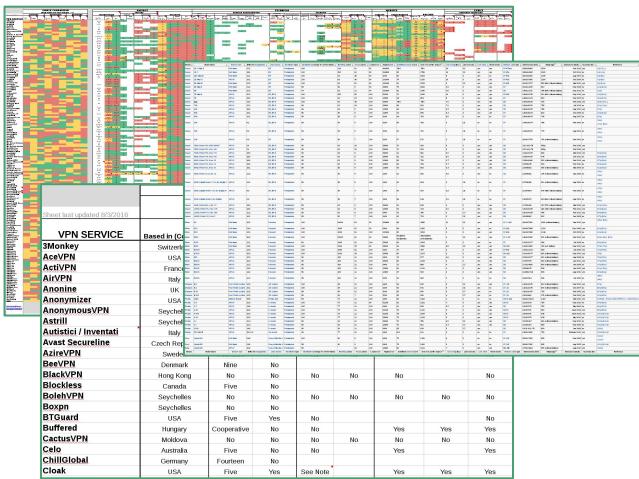
	VPN SERVICE	Based In (Country)	Server Count	USA	France	Italy	UK	USA	Seychelles	Italy	UK	Sweden	Denmark	Hong Kong	Croatia	Seychelles	Seychelles	USA	Hungary	Moldova	Australia	Germany	USA	
Surfshark	Surfshark	Surfshark	100000	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	
ExpressVPN	ExpressVPN	ExpressVPN	100000	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
AcetVPN	AcetVPN	AcetVPN	100000	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
ActivPN	ActivPN	ActivPN	100000	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
AlivVPN	AlivVPN	AlivVPN	100000	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Windscribe	Windscribe	Windscribe	100000	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Anonymous	Anonymous	Anonymous	100000	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
AnonymousVPN	AnonymousVPN	AnonymousVPN	100000	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Avast Secureline	Avast Secureline	Avast Secureline	100000	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
AztreVPN	AztreVPN	AztreVPN	100000	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
BeeVPN	BeeVPN	BeeVPN	100000	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
BlackVPN	BlackVPN	BlackVPN	100000	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Blockless	Blockless	Blockless	100000	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
BolehVPN	BolehVPN	BolehVPN	100000	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Bxpn	Bxpn	Bxpn	100000	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
BTGuard	BTGuard	BTGuard	100000	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Buffalo	Buffalo	Buffalo	100000	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
CactusVPN	CactusVPN	CactusVPN	100000	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Celo	Celo	Celo	100000	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
CloudGlobal	CloudGlobal	CloudGlobal	100000	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Cloak	Cloak	Cloak	100000	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No

Create a language for modeling PCMs



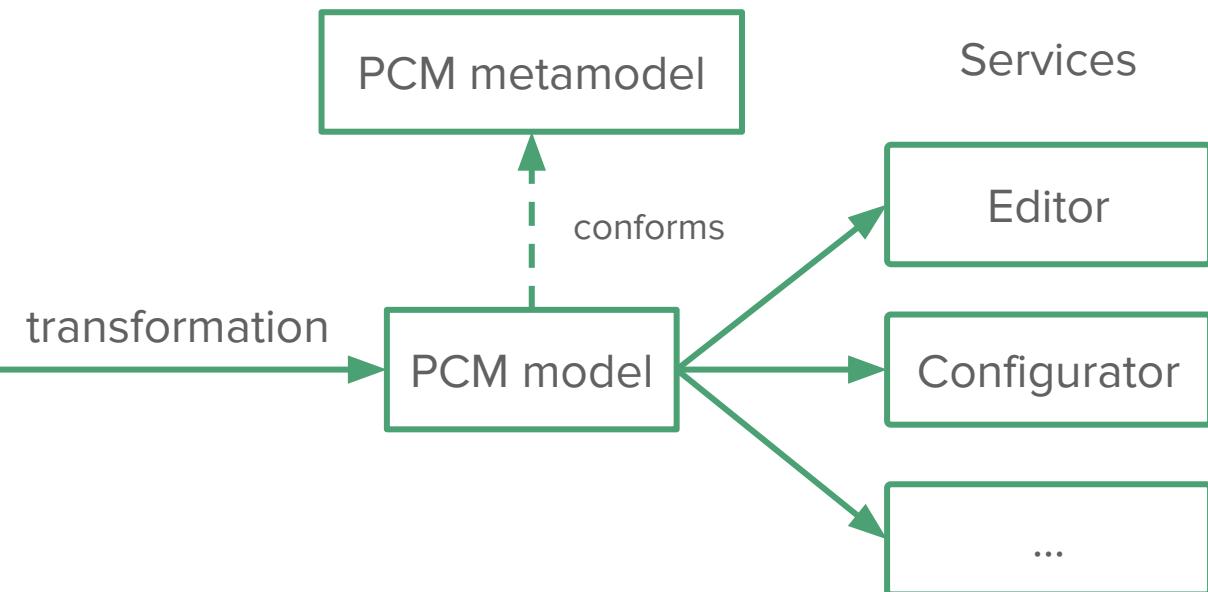
# Complementary Approaches to Formalize PCMs

Billions of product comparison matrices



A screenshot of a Microsoft Excel spreadsheet titled "Sheet1 last updated 03/29/2016". The table has a header row with various service names like "VPN SERVICE", "Smart", "AcetVPN", "ActivVPN", "AirVPN", "Airtel", "Anonymous", "AnonymousVPN", "Astro", "Astrolic", "Avast Secureline", "AztreVPN", "BeeVPN", "BlackVPN", "Blockless", "BoleVPN", "Boxpi", "BTGuard", "Buffalo", "CactusVPN", "Celo", "CloudGlobal", and "Cloak". The columns represent different countries or regions such as "Based in (Country)", "USA", "France", "Italy", "UK", "USA", "Seychelles", "Seychelles", "Italy", "Sweden", "Denmark", "Hong Kong", "China", "Seychelles", "Seychelles", "Five", "Five", "Five", "Five", "Fourteen", and "USA". The data consists of binary values (Yes/No) indicating the presence of specific features or services.

VPN SERVICE	Based in (Country)	USA	France	Italy	UK	USA	Seychelles	Seychelles	Italy	Sweden	Denmark	Hong Kong	China	Seychelles	Seychelles	Five	Five	Five	Five	Fourteen	USA
Smart	Smart	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
AcetVPN	AcetVPN	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
ActivVPN	ActivVPN	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
AirVPN	AirVPN	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Airtel	Airtel	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Anonymous	Anonymous	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
AnonymousVPN	AnonymousVPN	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Astro	Astro	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Astrolic	Astrolic	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Avast Secureline	Avast Secureline	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
AztreVPN	AztreVPN	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
BeeVPN	BeeVPN	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
BlackVPN	BlackVPN	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Blockless	Blockless	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
BoleVPN	BoleVPN	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Boxpi	Boxpi	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
BTGuard	BTGuard	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Buffalo	Buffalo	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
CactusVPN	CactusVPN	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Celo	Celo	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
CloudGlobal	CloudGlobal	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Cloak	Cloak	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No



Create a language for modeling PCMs

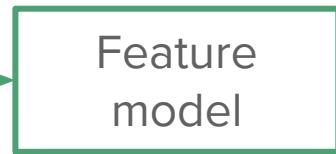
- Define the concepts of the domain
- Cover the diversity of PCMs
- May not be suitable for some services

# Complementary Approaches to Formalize PCMs

Billions of product comparison matrices

VPN SERVICE	Based in (C)
3Monkey	Singapore
AvPN	USA
BoxVPN	France
AirVPN	India
Anonime	UK
AnonymouseVPN	USA
Astrill	Service
Autisticci / Inventati	Seychelles
Avtast Secureline	Italy
Avtast Secureline	Czech Rep.
BeeVPN	Denmark
BlackVPN	Hong Kong
Blockless	Canada
BoxyPN	Five
Boxpn	No
BTGuard	Singapore
Buffered	No
CactusVPN	Seychelles
Celo	No
ChilliGlobal	USA
Cloak	Fourteen
	No
	See Note
	Yes
	Yes
	Yes

synthesis



Services

Editor

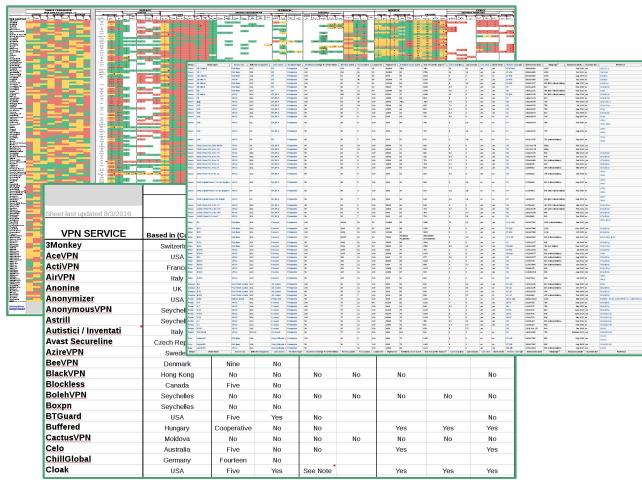
Configurator

• • •

Use an existing language for modeling product lines: **feature modeling**

# Complementary Approaches to Formalize PCMs

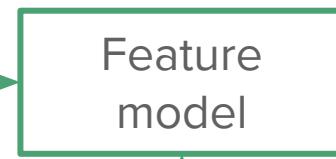
Billions of product comparison matrices



A screenshot of a complex product comparison matrix (PCM) for various VPN services. The matrix is color-coded and shows numerous features across different service providers. A legend on the left identifies the colors for different service types. The matrix includes columns for service names like 3Monkey, AirVPN, Anonime, and Cloak, and rows for geographical locations such as Australia, Canada, France, Germany, Hong Kong, India, Italy, Japan, Malaysia, Mexico, Singapore, Switzerland, Thailand, UK, USA, and Vietnam. The matrix contains binary values (Yes/No) indicating the presence or absence of specific features.

VPN SERVICE		Geographical Locations									
		Denmark	Hong Kong	Canada	France	India	UK	USA	Singapore	Malaysia	Thailand
Based in (C)	3Monkey	No	No	No	No	No	No	No	No	No	No
	AirVPN	Yes	No	No	No	No	No	No	No	No	No
Autonome	Yes	No	No	No	No	No	No	No	No	No	No
Anonymizer/VPN	Yes	No	No	No	No	No	No	No	No	No	No
Astrill	Yes	No	No	No	No	No	No	No	No	No	No
Autisticci / Inventati	Yes	No	No	No	No	No	No	No	No	No	No
Avast Secureline	Yes	No	No	No	No	No	No	No	No	No	No
Astroline	Yes	No	No	No	No	No	No	No	No	No	No
BeeVPN	Yes	No	No	No	No	No	No	No	No	No	No
BlackVPN	Yes	No	No	No	No	No	No	No	No	No	No
Blockless	Yes	No	No	No	No	No	No	No	No	No	No
BoxyPN	Yes	No	No	No	No	No	No	No	No	No	No
BTGuard	Yes	No	No	No	No	No	No	No	No	No	No
Buffalo	Yes	No	No	No	No	No	No	No	No	No	No
CactusVPN	Yes	No	No	No	No	No	No	No	No	No	No
Celo	Yes	No	No	No	No	No	No	No	No	No	No
ChilliGlobal	Yes	No	No	No	No	No	No	No	No	No	No
Cloak	Yes	No	No	No	No	No	No	No	No	No	No

synthesis



Services

Editor

Configurator

$\varphi$

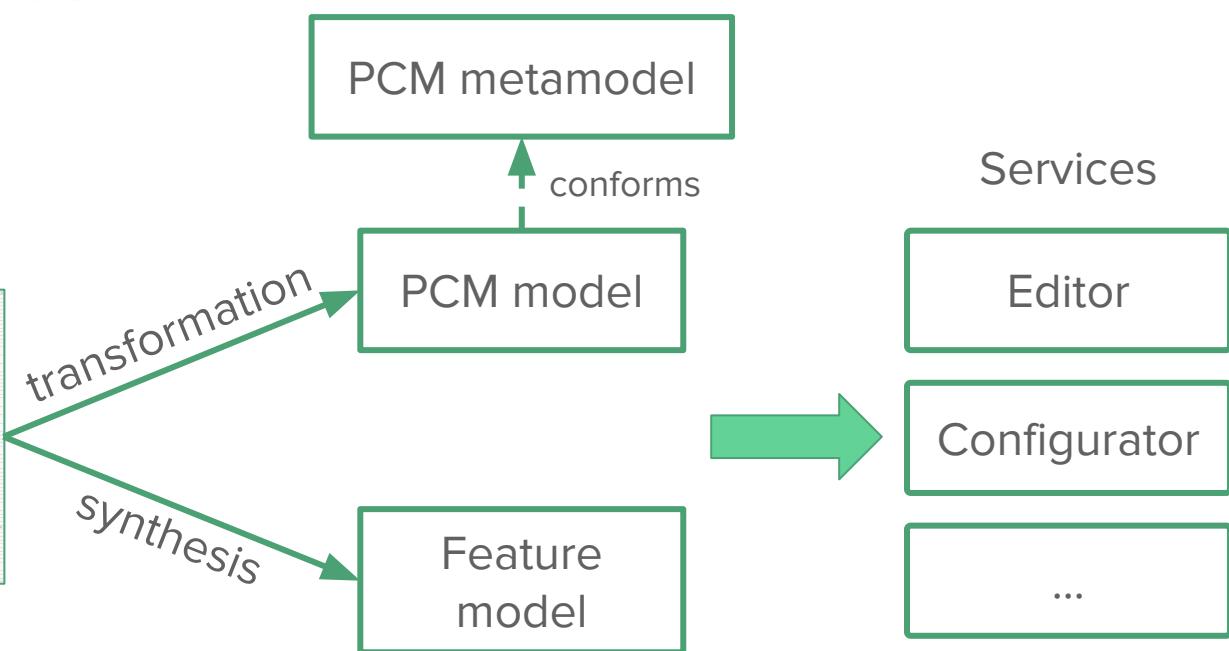
Use an existing language for modeling product lines: **feature modeling**

- Support a class of PCMs
- Feature oriented
- Precise semantics (propositional logic)
- Existing advanced reasoning techniques

# Complementary Approaches to Formalize PCMs

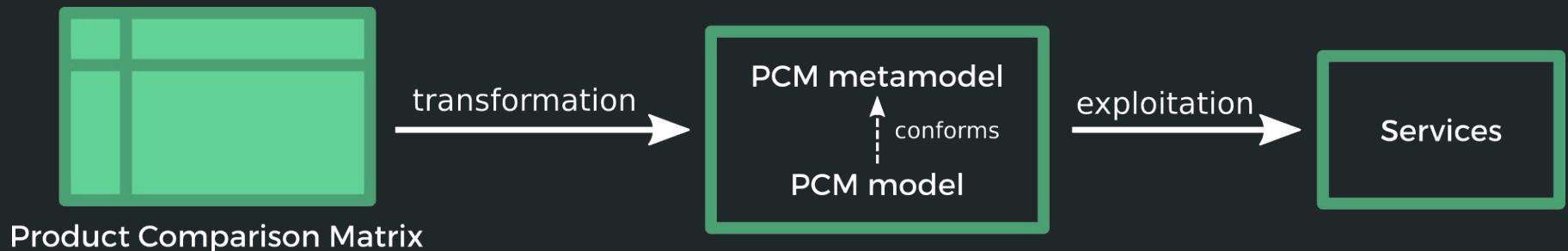
Billions of product comparison matrices

Product Comparison Matrix (PCM) Examples									
VPN SERVICE	Based in (Country)	Supports (Protocol)							
BlackVPN	Hong Kong	No							
CloudVPN	Canada	Five	No						
CloudVPN	Seychelles	No							
CloudVPN	Seychelles	Five	Yes	No	No	Yes	Yes	Yes	Yes
CloudVPN	USA	No							
CloudVPN	Moldova	Five	No	No	No	Yes	No	No	No
CloudVPN	Australia	Five	No	No	No	Yes	Yes	Yes	Yes
CloudVPN	Germany	Fourteen	No	No	No	Yes	Yes	Yes	Yes
CloudVPN	USA	Five	Yes	See Note					



	Model-based approach	Feature model synthesis
<b>Main concept</b>	Product	Feature
<b>Support of PCMs</b>	All	A class of PCMs
<b>Gap with PCMs</b>	Small	Large
<b>Services</b>	Generic (e.g. editor)	Specific (e.g. configurator)

# Model-Based Approach

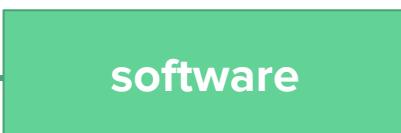


# Design of a model-based approach for PCMs

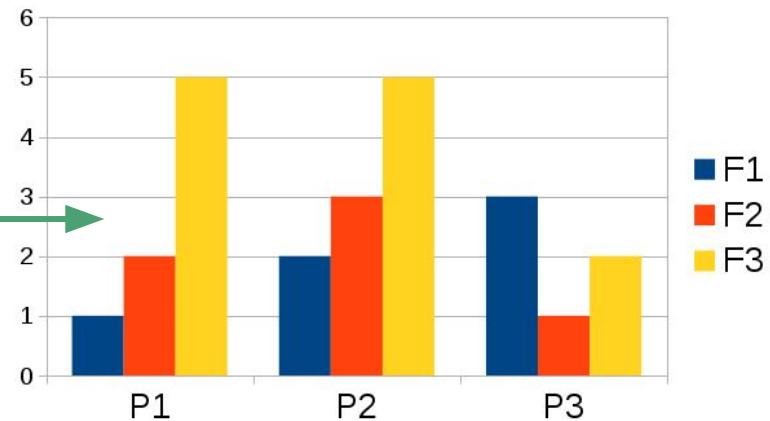
Let's pick a first PCM and service and build the appropriate metamodel

Product comparison matrix

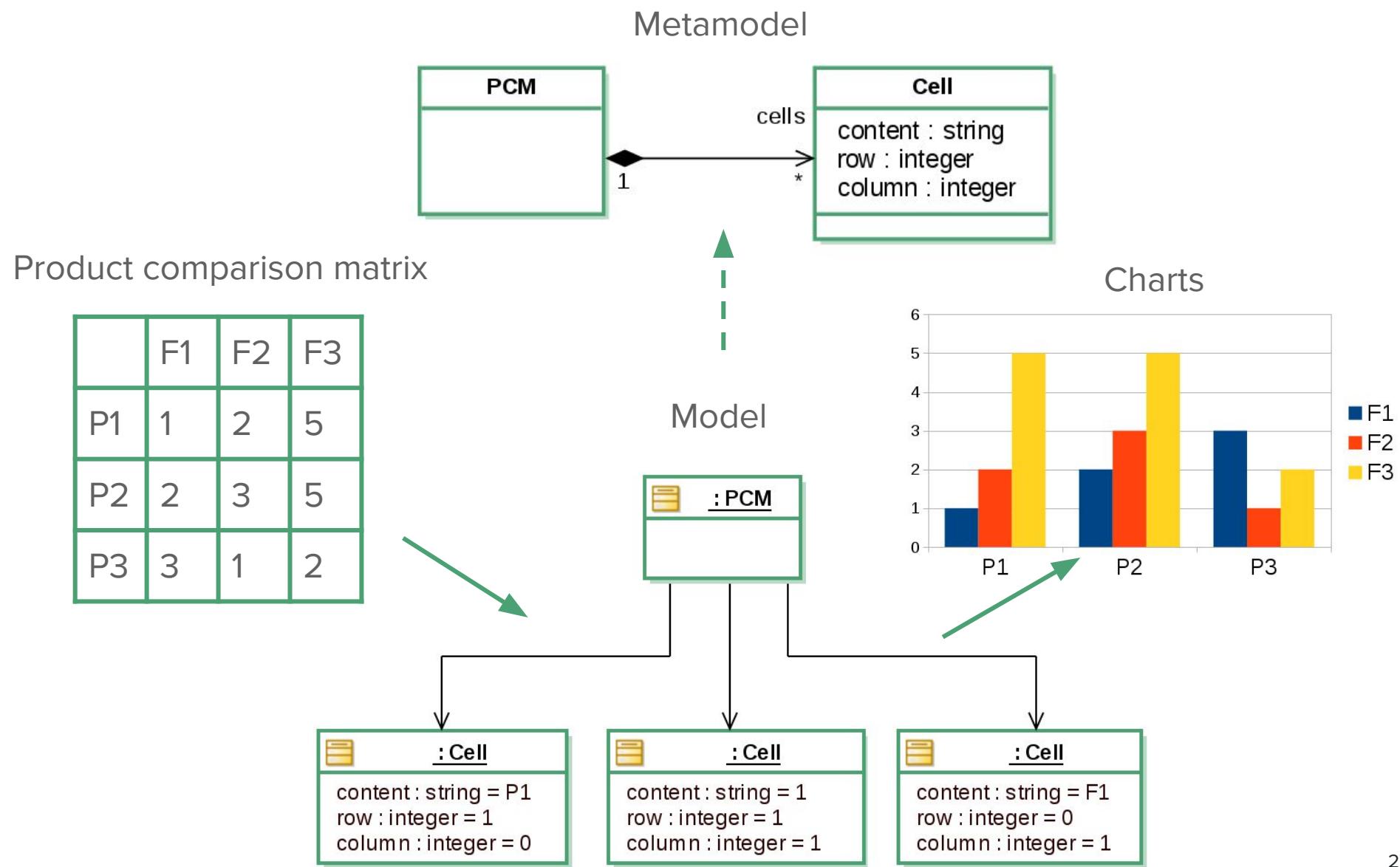
	F1	F2	F3
P1	1	2	5
P2	2	3	5
P3	3	1	2



Charts



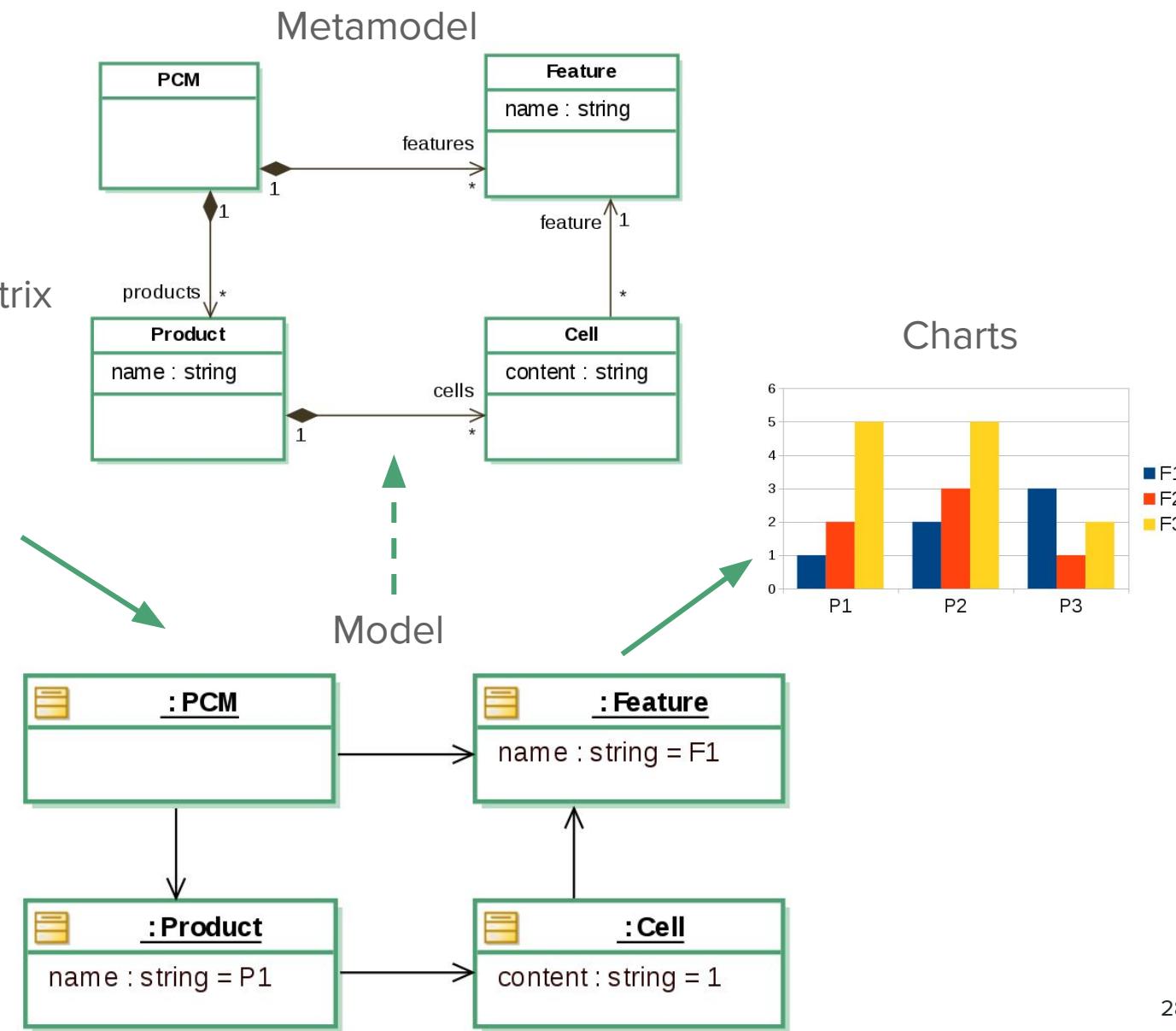
# Design of a model-based approach for PCMs



# Design of a model-based approach for PCMs

Product comparison matrix

	F1	F2	F3
P1	1	2	5
P2	2	3	5
P3	3	1	2

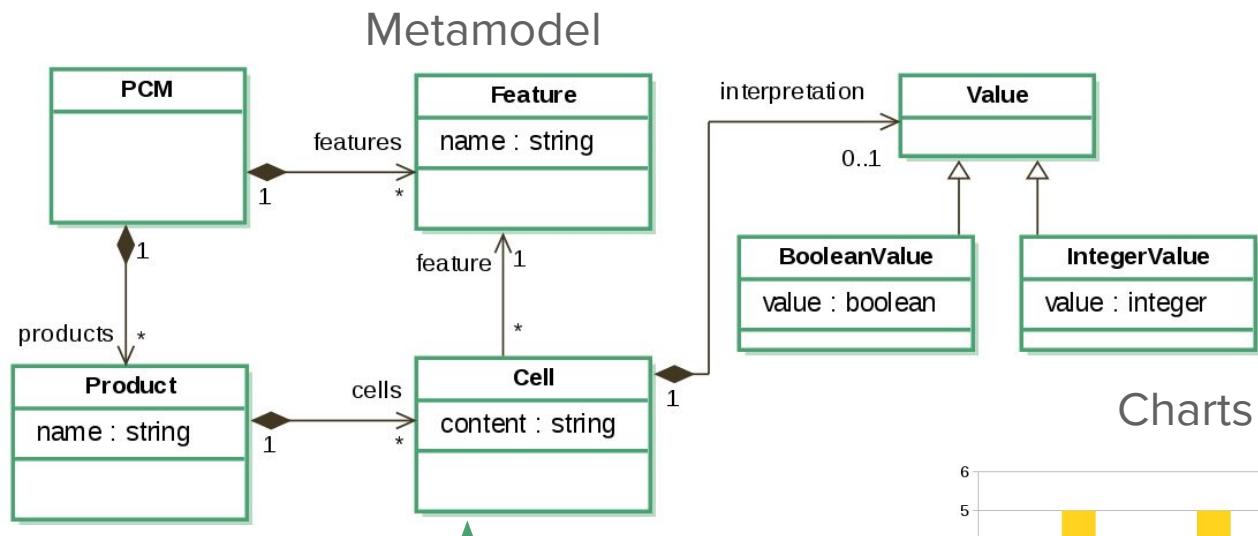


# Design of a model-based approach for PCMs

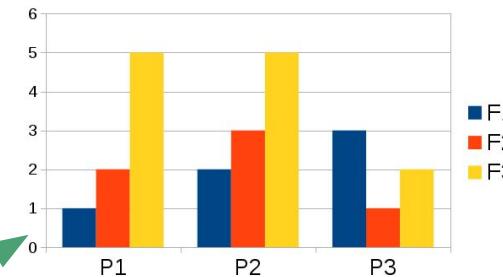
Product comparison matrices

	F1	F2	F3
P1	1	2	5
P2	2	3	5
P3	3	1	2

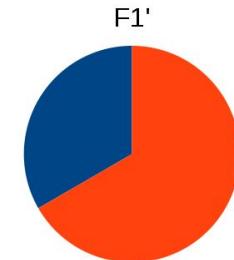
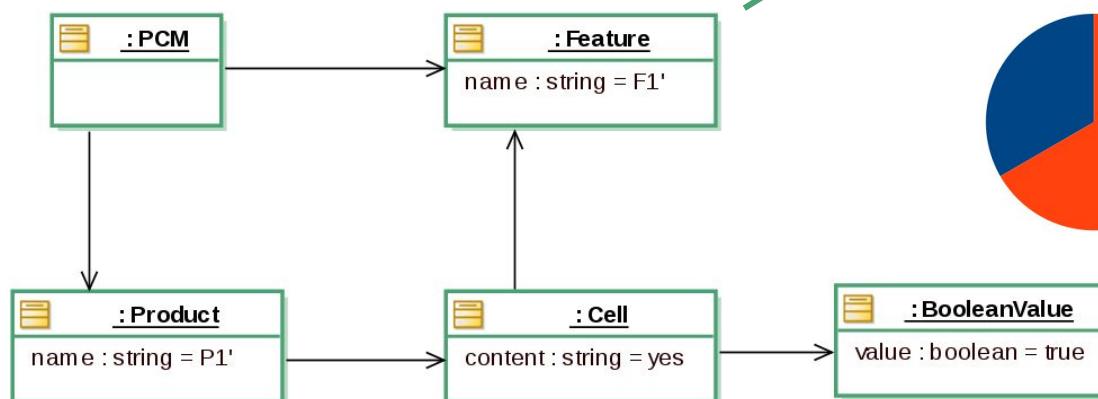
	F1'	F2'
P1'	yes	no
P2'	no	no
P3'	no	yes



Charts



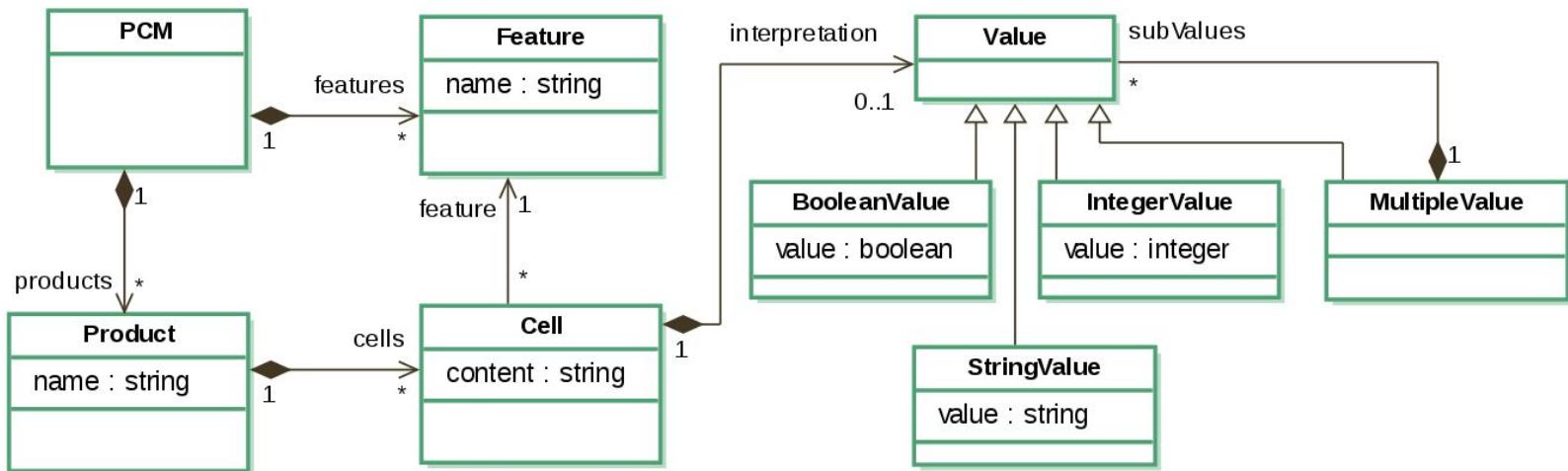
Model



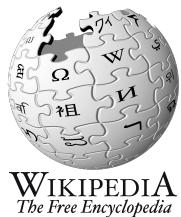
yes  
no

# Design of a model-based approach for PCMs

Metamodel



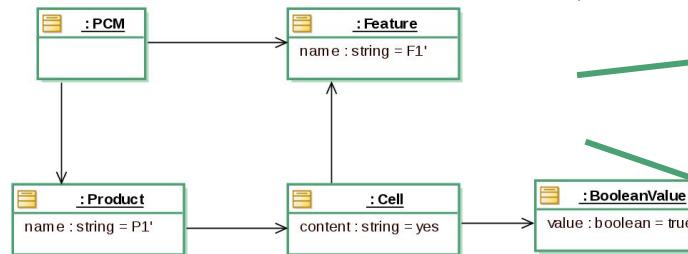
Product comparison  
matrices



10 PCMs

transformation

Model



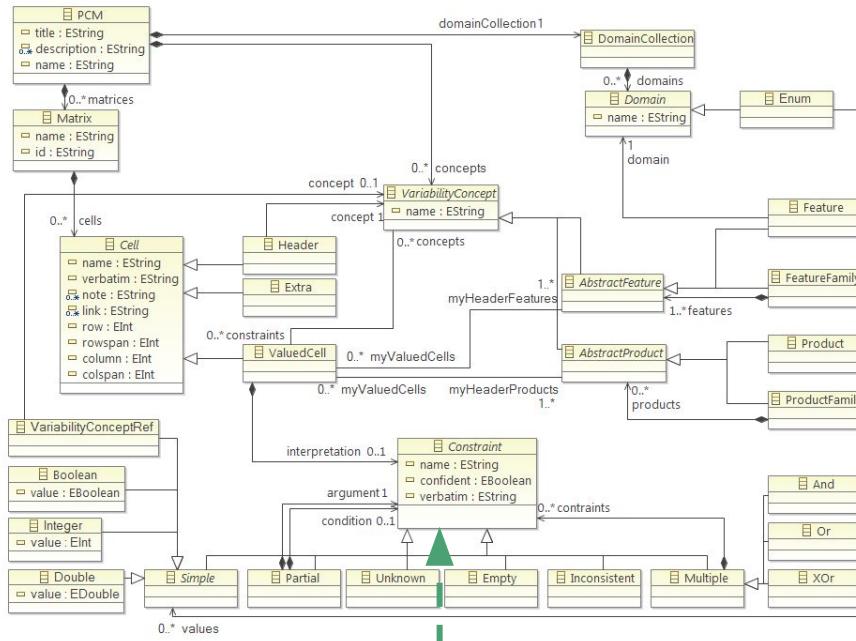
Charts

Editor

...

# Design of a model-based approach for PCMs

## Metamodel



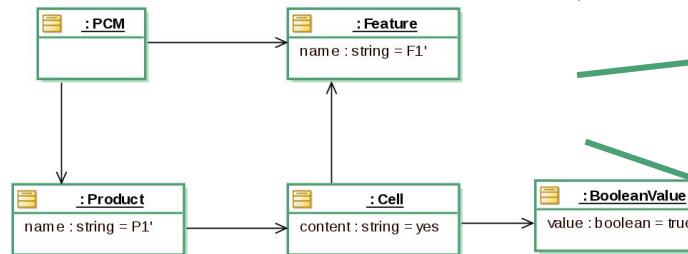
Product comparison  
matrices



50 PCMs

transformation

## Model



Charts

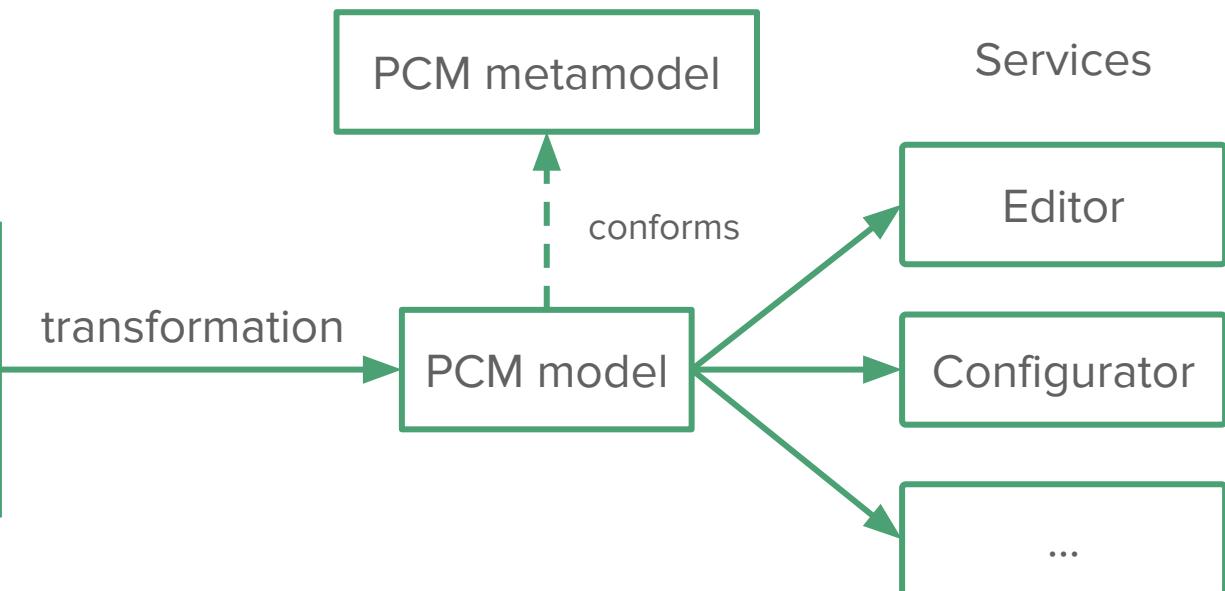
Editor

...

# How to evaluate our approach?

Billions of product comparison matrices

VPN SERVICE	Based in (Country)	Latency (ms)	Encryption	Protocol	Bandwidth (Mbps)	Latency (ms)	Encryption	Protocol	Bandwidth (Mbps)
3Monkey	Switzerland	Nine	No	No	No	No	No	No	No
AceVPN	USA	Five	No	No	No	No	No	No	No
AdGuardVPN	France	Five	No	No	No	No	No	No	No
AirVPN	Italy	Five	No	No	No	No	No	No	No
Anonmine	UK	Five	No	No	No	No	No	No	No
AnonymousVPN	USA	Five	No	No	No	No	No	No	No
Astrial	Spain	Five	No	No	No	No	No	No	No
Autistic / Inventati	Sweden	Five	No	No	No	No	No	No	No
Avast Secureline	Czech Rep.	Five	No	No	No	No	No	No	No
BeeVPN	Denmark	Nine	No	No	No	No	No	No	No
BlackVPN	Hong Kong	No	No	No	No	No	No	No	No
Blockless	Canada	Five	No	No	No	No	No	No	No
BobVPN	No	No	No	No	No	No	No	No	No
Boxpn	Seychelles	No	No	No	No	No	No	No	No
BTGuard	USA	Five	Yes	No	No	No	No	No	No
Buffered	Hungary	Cooperative	No	No	Yes	Yes	Yes	Yes	Yes
CastorVPN	Azerbaijan	Five	No	No	No	No	No	No	No
Celo	Australia	Fifteen	No	No	Yes	Yes	Yes	Yes	Yes
ChillGlobal	Germany	Five	Yes	See Note	Yes	Yes	Yes	Yes	Yes
Coak	USA								



No oracle to check that our metamodel and transformation are correct, complete and relevant

Validate our approach with users

- Not modeling experts
- Build services to show the concepts of the metamodel in action

# Evaluation: Settings

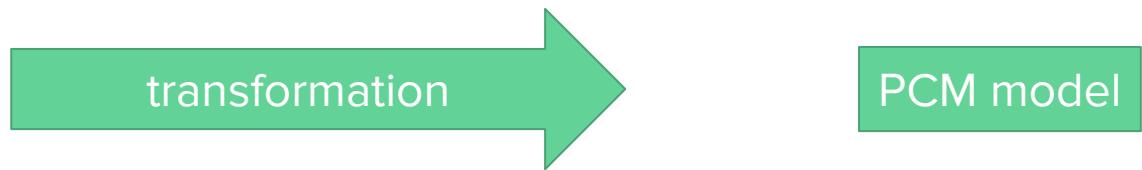
RQ1: Is our metamodel **complete** and **relevant** to the formalization of PCMs?

RQ2: What is the **precision** of our transformation of raw PCMs to PCM models?



WIKIPEDIA  
The Free Encyclopedia

75 PCMs



Completely automated except for indicating cells that contain features or products

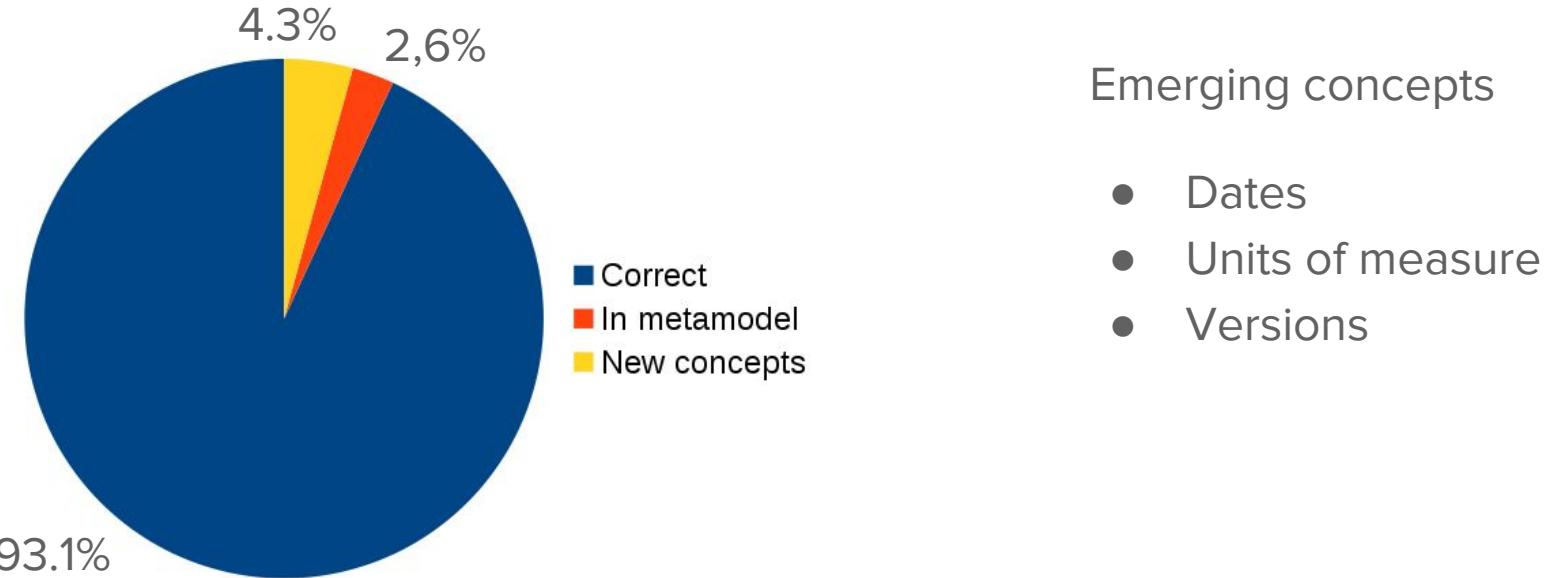
Manual review of PCMs

- 20 evaluators (researchers and engineers)
- Task: check interpretation of each cell
- Development of an editor to support the evaluation

feature	Type	Number of doors	Color	Automatic head
Car 1	4x4	5	black, white, red	yes
Car 2	sedan	5	black, white, red,...	no
Car 3	sedan	5	grey	yes
Car 4	city car	3	black, white, green	no
Car 5	sport	3	Integer(3) :	

# Evaluation

30,061 evaluated cells



Emerging concepts

- Dates
- Units of measure
- Versions

RQ1: Is our metamodel **complete** and **relevant** to the formalization of PCMs?

Almost but we can add the 3 missing concepts to the metamodel

RQ2: What is the **precision** of our transformation of raw PCMs to PCM models?

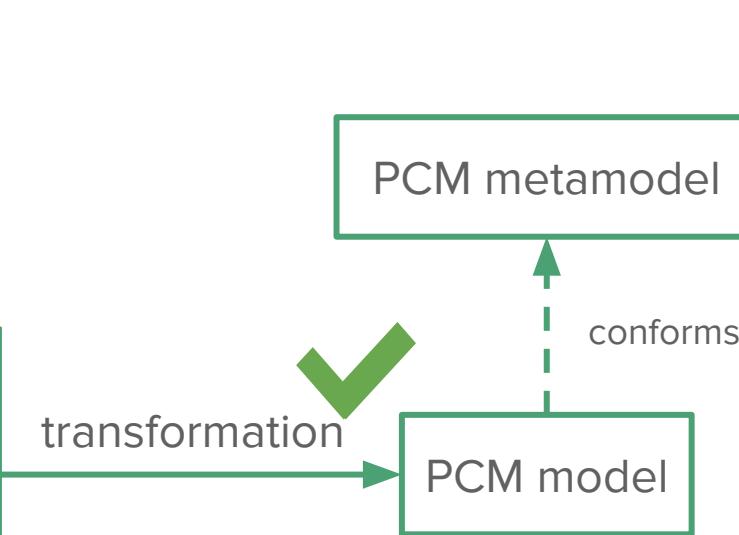
High precision of the transformation (93.1%)

# Evaluation

Billions of product comparison matrices



VPN SERVICE		Based in (Country)		Protocol		Encryption		P2P		Bandwidth		Latency		Reliability		Security		Logs		Support		Price		Overall	
3Monkey	Switzerland	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	
AceVPN	USA	Five	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	
AdGuardVPN	France	Five	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	
AirVPN	Italy	Five	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	
Anonmine	UK	Five	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	
AnonymousVPN	USA	Five	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	
Astrill	Sweden	Five	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	
Autistic / Inventati	Italy	Five	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	
Avast Secureline	Czech Rep.	Five	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	
BeeVPN	Denmark	Nine	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	
BlackVPN	Hong Kong	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	
Blockless	Canada	Five	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	
Boxpn	Singapore	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	
BTGuard	Seychelles	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	
Buffered	USA	Five	Yes	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	
CastorVPN	Hungary	Cooperative	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	
Cefo	Australia	Five	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	
ChillGlobal	Australia	Fifteen	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	
cloak	Germany	Five	Yes	See Note *	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	



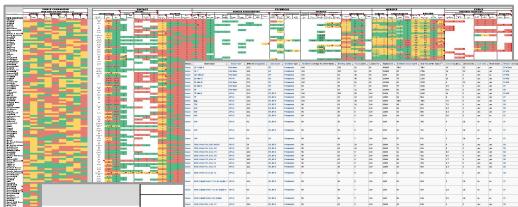
Data-driven metamodeling (ASE, 2014)

- Iterative process
- Numerous instances
- Co-design of metamodel and transformation

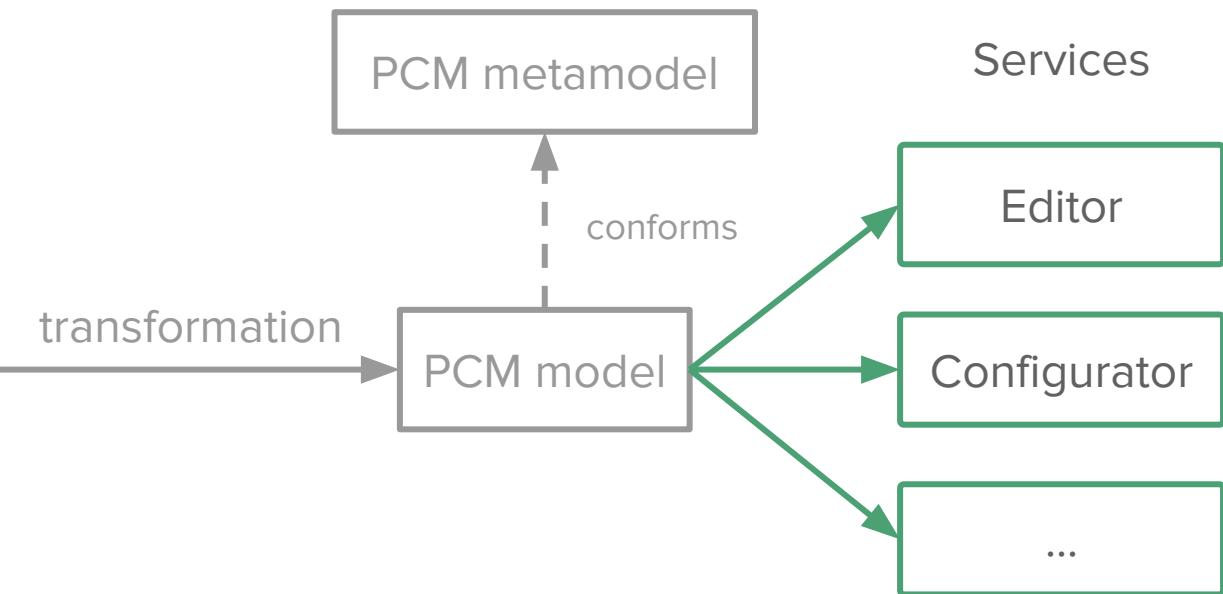
Development of services to validate the approach by end-users

# Evaluation

Billions of product comparison matrices



VPN SERVICE		Based in (Country)									
3Monkey	Switzerland	No	No	No	No	No	No	No	No	No	No
AceVPN	USA	Five	No	No	No	No	No	No	No	No	No
AdGuardVPN	France	No	No	No	No	No	No	No	No	No	No
AirVPN	Italy	No	No	No	No	No	No	No	No	No	No
Anonmine	UK	No	No	No	No	No	No	No	No	No	No
Anonymizer	USA	No	No	No	No	No	No	No	No	No	No
Astirill	Spain	No	No	No	No	No	No	No	No	No	No
Autistic / Inventati	Italy	No	No	No	No	No	No	No	No	No	No
Avast Secureline	Czech Rep.	No	No	No	No	No	No	No	No	No	No
BeeVPN	Denmark	Nine	No	No	No	No	No	No	No	No	No
BlackVPN	Hong Kong	No	No	No	No	No	No	No	No	No	No
Blockless	Canada	Five	No	No	No	No	No	No	No	No	No
Boxpn	Seychelles	No	No	No	No	No	No	No	No	No	No
BTGuard	USA	Five	Yes	No	No	No	No	No	No	No	No
Buffered	Hungary	Cooperative	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
CastorVPN	Australia	Five	No	No	No	No	No	No	No	No	No
Cefo	Australia	Five	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
ChillGlobal	Germany	Fourteen	No	See Note *	Yes						
cloak	USA	Five	Yes	See Note *	Yes						



Success factors (Hermans, 2009):

- Learnability
- Usability
- Expressiveness

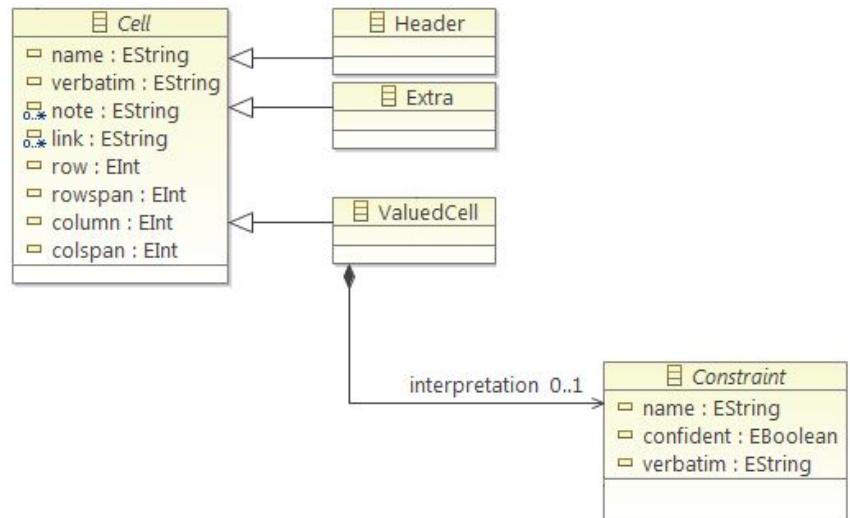
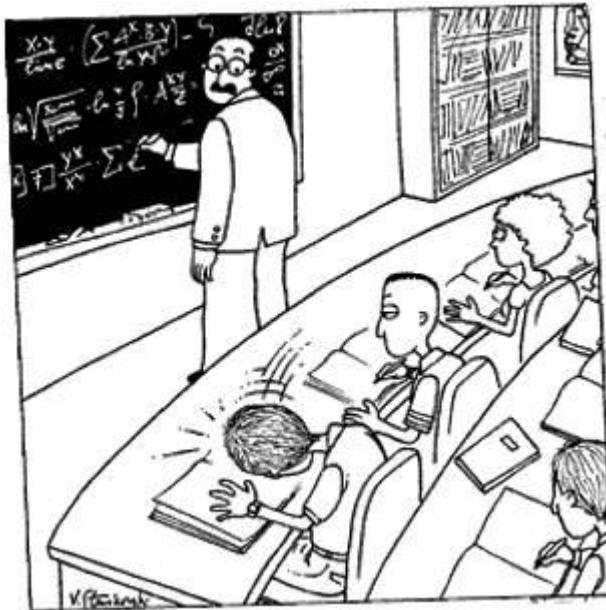
# Evaluation

RQ3: Is our metamodel suited for **engineering** automated analyses and **services** dedicated to PCMs?

Experiment: development of an API and an editor (20 Master students)



Learnability: the metamodel is hard to explain



# Evaluation

RQ3: Is our metamodel suited for **engineering** automated analyses and **services** dedicated to PCMs?

Experiment: development of the editor to support the evaluation (1 researcher)

feature	Type	Number of doors	Color	Automatic headlights	0 to 100 km/h
Car 1	4x4	5	black, white, red	yes	9.5
Car 2	sedan	5	black, white, red,...	no	6.3
Car 3	sedan	5	grey	yes	5.6
Car 4	city car	3	black, white, green	no	8.3
Car 5	sport	3	Integer(3) :		3.9
Car 6	sedan		2 doors	yes (premium ve...	N/A

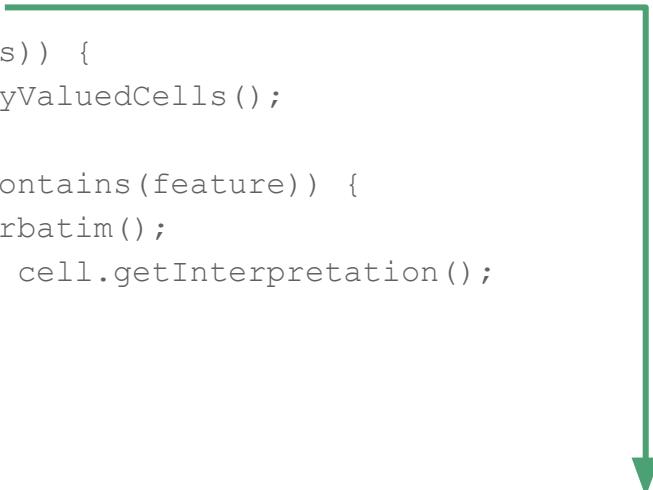


A simplified metamodel was internally used for the editor

# Evaluation

Central operation for services: iterate over each cell

```
for (Product product : getProducts(concepts)) {  
    for (Feature feature : getFeatures(concepts)) {  
        List<ValuedCell> cells = product.getMyValuedCells();  
        for (ValuedCell cell : cells) {  
            if (cell.getMyHeaderFeatures().contains(feature)) {  
                String content = cell.getVerbatim();  
                Constraint interpretation = cell.getInterpretation();  
            }  
        }  
    }  
  
}  
  
private List<Product> getProducts(List<? extends VariabilityConcept> concepts) {  
    List<Product> products = new ArrayList<>();  
    for (VariabilityConcept concept : concepts) {  
        if (concept instanceof ProductFamily) {  
            List<AbstractProduct> subProducts = ((ProductFamily) concept).getProducts();  
            products.addAll(getProducts(subProducts));  
        } else if (concept instanceof Product) {  
            products.add((Product) concept);  
        }  
    }  
    return products;  
}
```



call

**X** Expressiveness: complex navigation to main concepts

# Evaluation

RQ3: Is our metamodel suited for **engineering** automated analyses and **services** dedicated to PCMs?

## Feedback



Concepts of features and products



The metamodel is hard to explain



Rows / Columns  
VS  
Features / Products



A simplified metamodel is used internally for the editor



Influence of Wikipedia



Complex navigation to main concepts



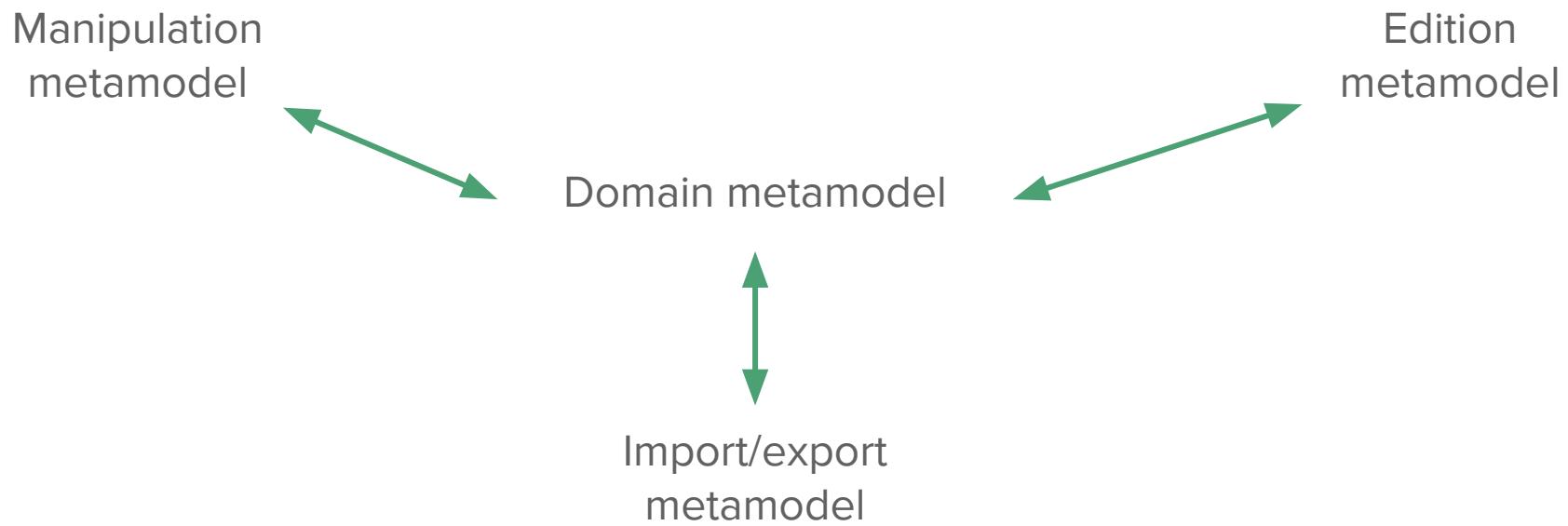
Semantics of the cells

**Our metamodel is not satisfying for engineering services**

**We need another solution !**

# From One Metamodel to Multiple Metamodels

**Objective:** separate each concern in a dedicated metamodel



**Task-specific metamodels:** reduce the gap between the representation of a PCM and the service

# Evaluation

RQ3: Is our metamodel suited for **engineering** automated analyses and **services** dedicated to PCMs?

## Experiments

- Web editor
- Comparator
- Configurator
- Visualization
- Importer / exporter (HTML, CSV and Wikitext)
- Browser extensions
- Machine learning
- Statistics
- Summary
- MatrixMiner (Ben Nasr, 2015)

Informal feedback from

- 4 researchers
- 85 Master students in different courses

# Evaluation

RQ3: Is our metamodel suited for **engineering** automated analyses and **services** dedicated to PCMs?



Explaining the metamodel is easier



Students can program the extraction of statistics in a few minutes

Central operation for services: iterate over each cell

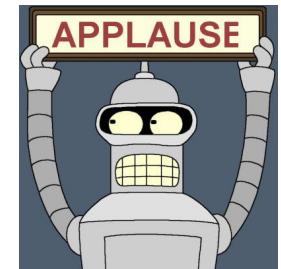
```
for (Product product : pcm.getProducts()) {  
    for (Feature feature : pcm.getConcreteFeatures()) {  
        Cell cell = product.findCell(feature);  
        String content = cell.getContent();  
        Value interpretation =  
cell.getInterpretation();  
    }  
}
```



} Implementing an editor is straightforward (the metamodel is close to the concrete syntax)



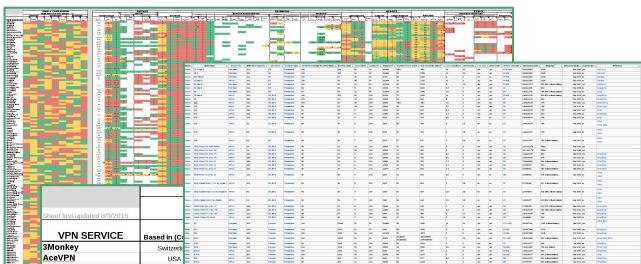
Comes with well known synchronization problems



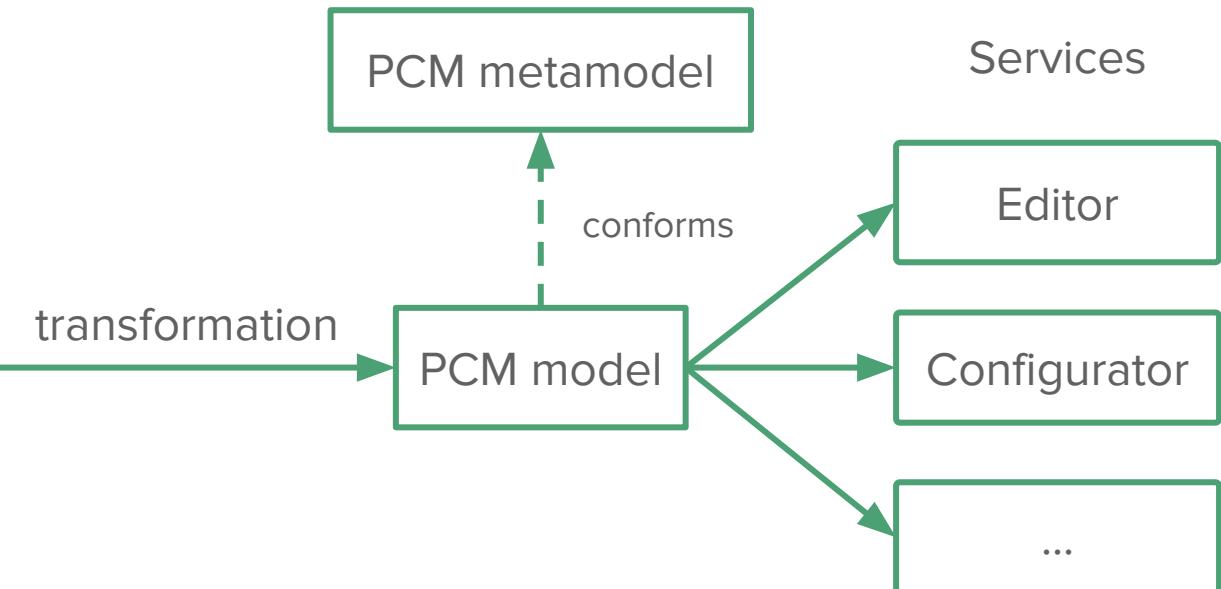
**Separating concerns in dedicated metamodels eases the engineering of services**

# Conclusion

Billions of product comparison matrices

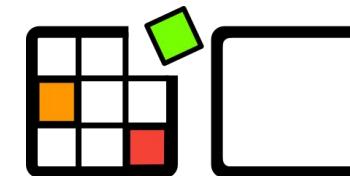
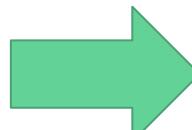


VPN SERVICE	Based in IC	Switzerland	USA	France	Germany	UK	USA	Sweden	Czech Rep.	Denmark	Hong Kong	Canada	Seychelles	Spain	USA	Hungary	Moldova	Australia	Portugal	BTGuard	Buffered	CactusVPN	Cato	ChillGlobal	Cloak
3Monkey	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No		
AceVPN	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No		
Avast	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No		
Anonymizer	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No		
AnyAnonymousVPN	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No		
AntiT	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No		
Antivirus	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No		
Avast Secureline	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No		
AzureVPN	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No		
Balvi	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No		
BlackVPN	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No		
Blockless	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No		
BolehVPN	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No		
Bug	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No		
BTGuard	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No		
Buffered	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No		
CactusVPN	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No		
Cato	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No		
ChillGlobal	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No		
Cloak	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No		



Model-based approach for formalizing PCMs

- Completeness and relevance of concepts
- Precise transformation
- Dedicated services



OpenCompare

framework and website  
dedicated to PCMs

# Lessons learned

Data-driven metamodeling is hard

- Numerous examples (1,500,000 PCMs in Wikipedia)
- Iterative process
- Co-design of metamodel and transformation

Metamodeling is not just data-driven

- Validation of metamodel and transformation requires end-users
- Requires the development of services for end-users
- Services influence the design of the metamodel

Metamodeling is a multi-objective problem with conflicting forces (data, users, services)

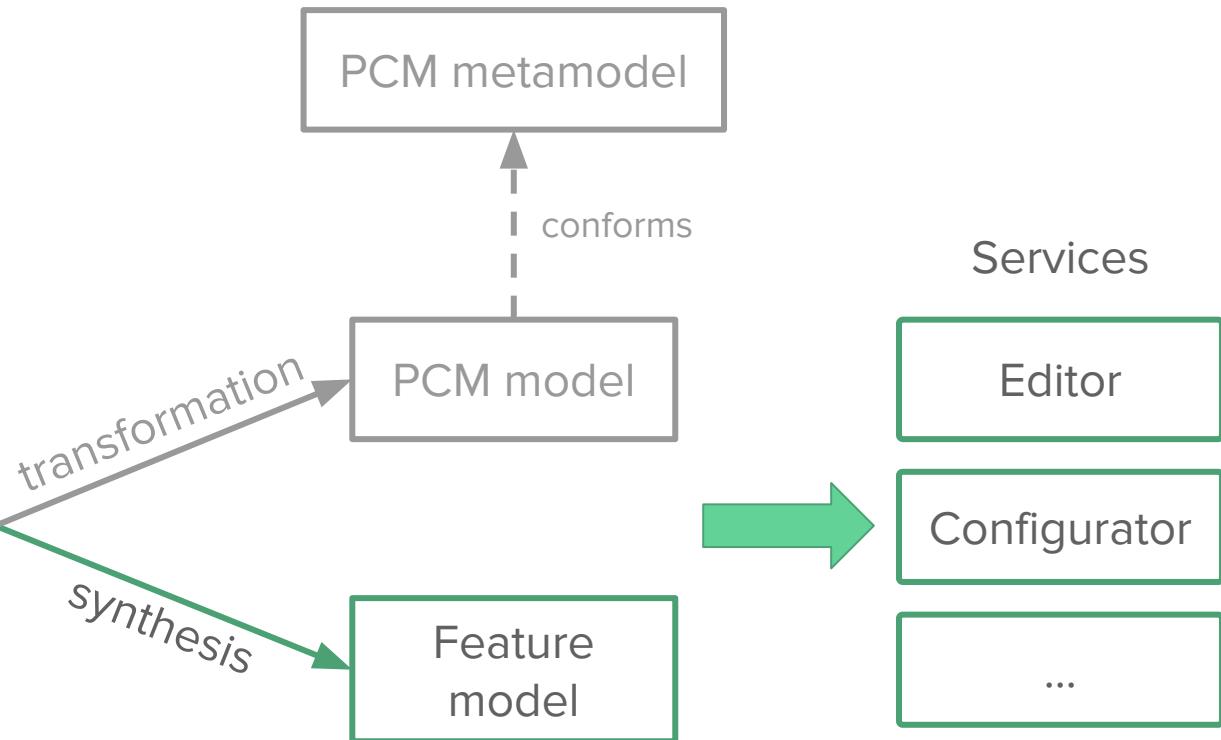
# Synthesis of Feature Models



# Formalizing PCMs

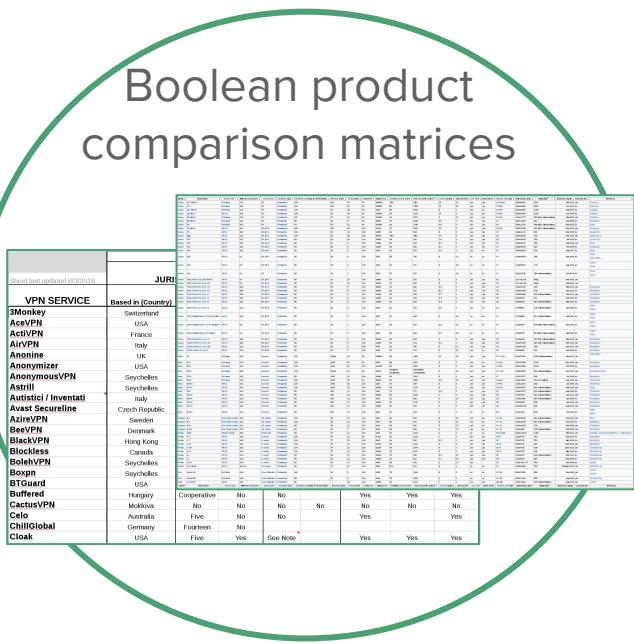
Billions of product comparison matrices

VPN SERVICE		Comparison Metrics									
3Monkey	Switzerland	No	No	No	No	No	No	No	No	No	No
ActiVPN	USA	No	No	No	No	No	No	No	No	No	No
AirVPN	France	No	No	No	No	No	No	No	No	No	No
Anonymizer	Italy	No	No	No	No	No	No	No	No	No	No
Astrill	UK	No	No	No	No	No	No	No	No	No	No
Autoneti i Inventati	USA	No	No	No	No	No	No	No	No	No	No
Avast SecureLine	Italy	No	No	No	No	No	No	No	No	No	No
AzireyVPN	Cyprus	No	No	No	No	No	No	No	No	No	No
BeeVPN	Denmark	No	No	No	No	No	No	No	No	No	No
BlackVPN	Hong Kong	No	No	No	No	No	No	No	No	No	No
Blockless	Canada	Five	No	No	No	No	No	No	No	No	No
BokehVPN	Seychelles	No	No	No	No	No	No	No	No	No	No
Boxpn	Seychelles	No	No	No	No	No	No	No	No	No	No
BTGuard	USA	Five	Yes	No	No	No	No	No	No	No	No
Buffered	Hungary	Cooperative	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
CiscoVPN	Moldova	Five	No	No	No	No	No	No	No	No	No
Celo	Australia	Five	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
ChillGlobal	Germany	Fourteen	No	See Note	Yes						
Cloak	USA	Five	Yes	See Note	Yes						



# Feature Model Synthesis

Product comparison  
matrices



Boolean product  
comparison matrices

synthesis

Feature model

Services

Editor

Configurator

...

$\Phi$

The semantics of a feature model is defined by a propositional formula

# Feature Modeling

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
PostgreSQL	✓	✗	✗	✗	✗	✗	✗	✗	✗	✗
MySQL	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓
License	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Domain	✓	✗	✗	✗	✗	✓	✗	✓	✓	✗
Proprietary License	✓	✗	✗	✗	✗	✗	✗	✗	✗	✗
Local	✗	✗	✓	✗	✓	✗	✗	✗	✗	✓
Programming Language	✗	✓	✗	✗	✓	✓	✓	✓	✗	✓
Java	✗	✓	✗	✗	✗	✓	✗	✗	✗	✓
Storage	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
PHP	✗	✗	✗	✗	✓	✗	✓	✓	✗	✗
Open Source	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓
Wiki	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Hosting	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Hosted Service	✓	✓	✗	✓	✗	✓	✓	✓	✗	✗

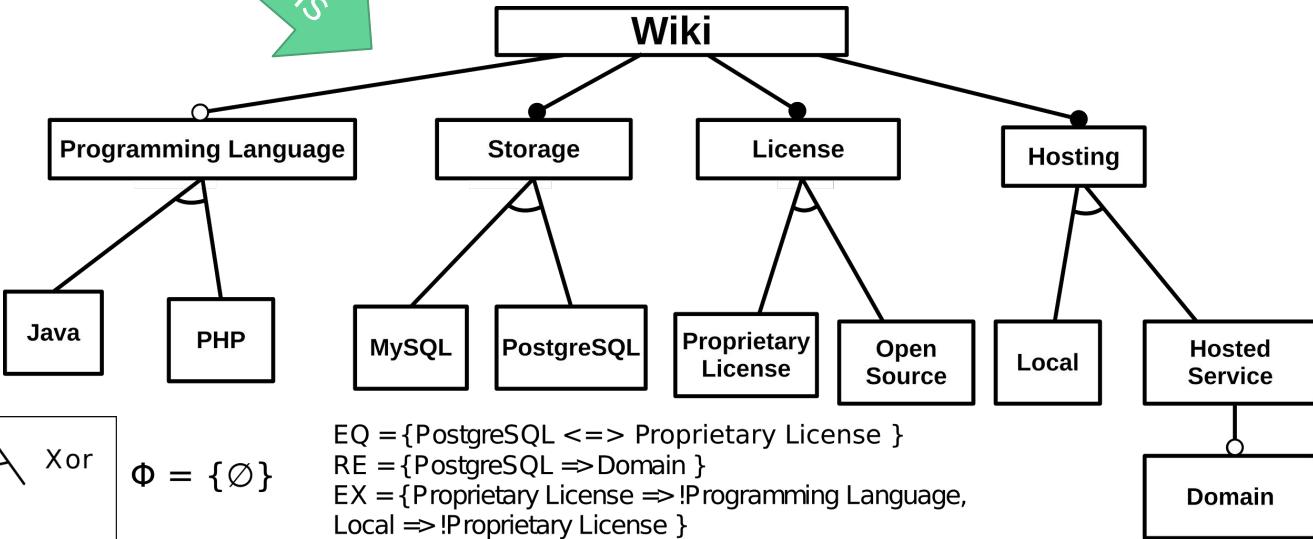
## Configuration semantics

Set of valid products

synthesis

What is the gap between PCMs and feature models?  
 Can we build an algorithm to synthesize feature models from PCMs?

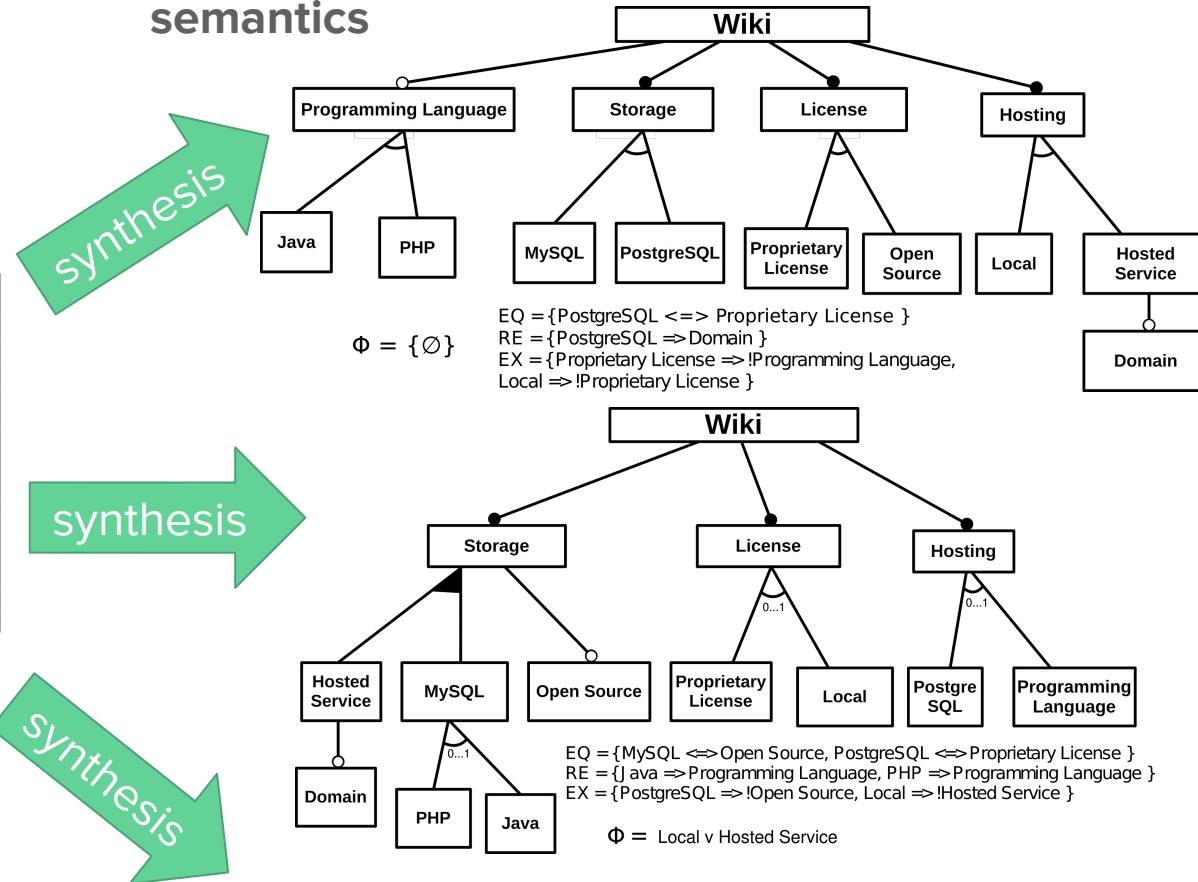
## Feature model



# Importance of Ontological Semantics

Multiple feature models can be synthesized with the same configuration semantics

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
PostgreSQL	✓	✗	✗	✗	✗	✗	✗	✗	✗	✗
MySQL	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓
License	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Domain	✓	✗	✗	✗	✗	✓	✗	✓	✓	✗
Proprietary License	✓	✗	✗	✗	✗	✗	✗	✗	✗	✗
Local	✗	✗	✓	✗	✓	✗	✗	✗	✗	✓
Programming Language	✗	✓	✗	✗	✓	✓	✓	✓	✗	✓
Java	✗	✓	✗	✗	✗	✓	✗	✗	✗	✓
Storage	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
PHP	✗	✗	✗	✗	✓	✗	✓	✓	✗	✗
Open Source	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓
Wiki	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Hosting	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Hosted Service	✓	✓	✗	✓	✗	✓	✓	✓	✓	✗



## Ontological semantics

Express the relationships between features  
Hierarchy of features + feature groups

...

# Importance of Ontological Semantics

## Ontological semantics

Express the relationships between features

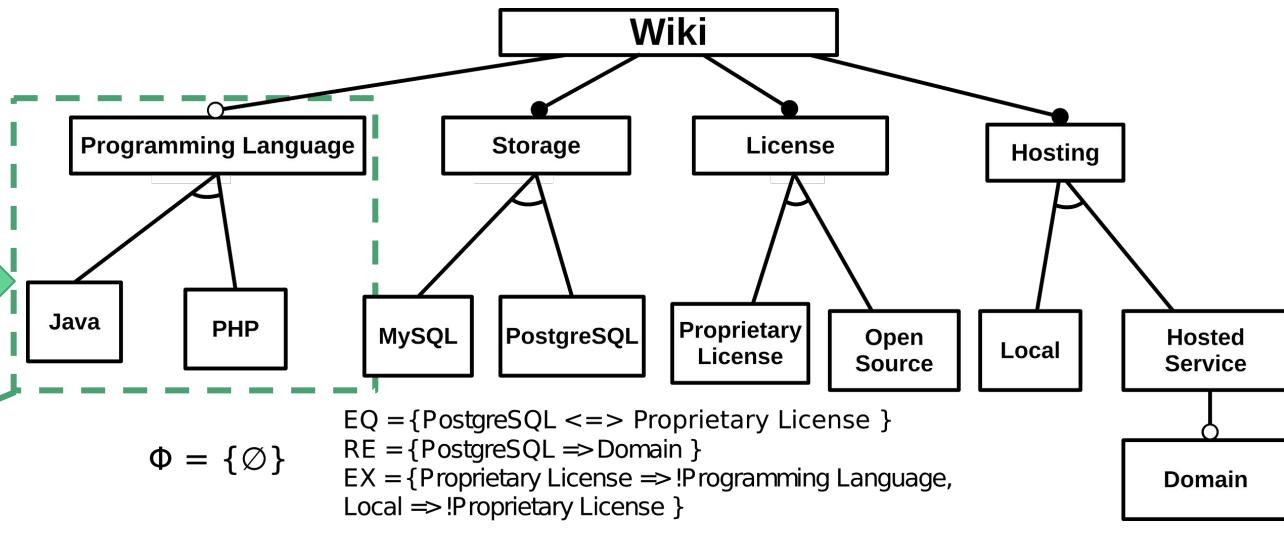
Hierarchy of features + feature groups

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
PostgreSQL	✓	✗	✗	✗	✗	✗	✗	✗	✗	✗
MySQL	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓
License	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Domain	✓	✗	✗	✗	✗	✗	✓	✓	✓	✗
Proprietary License	✓	✗	✗	✗	✗	✗	✗	✗	✗	✗
Local	✗	✗	✗	✗	✗	✗	✓	✓	✓	✓
Programming Language	✗	✓	✗	✗	✓	✓	✓	✓	✓	✓
Java	✗	✓	✗	✗	✗	✓	✓	✓	✓	✓
Storage	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
PHP	✗	✗	✗	✗	✗	✗	✓	✓	✗	✗
Open Source	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓
Wiki	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Hosting	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Hosted Service	✓	✓	✗	✓	✓	✓	✓	✓	✓	✗

synthesis

$$\Phi = \{\emptyset\}$$

$$\begin{aligned} EQ &= \{PostgreSQL \Leftrightarrow \text{Proprietary License} \} \\ RE &= \{PostgreSQL \Rightarrow \text{Domain} \} \\ EX &= \{\text{Proprietary License} \Rightarrow \neg \text{Programming Language}, \\ &\quad \text{Local} \Rightarrow \neg \text{Proprietary License} \} \end{aligned}$$



## Configurator

Which programming language do you prefer?

Java

PHP



# Importance of Ontological Semantics

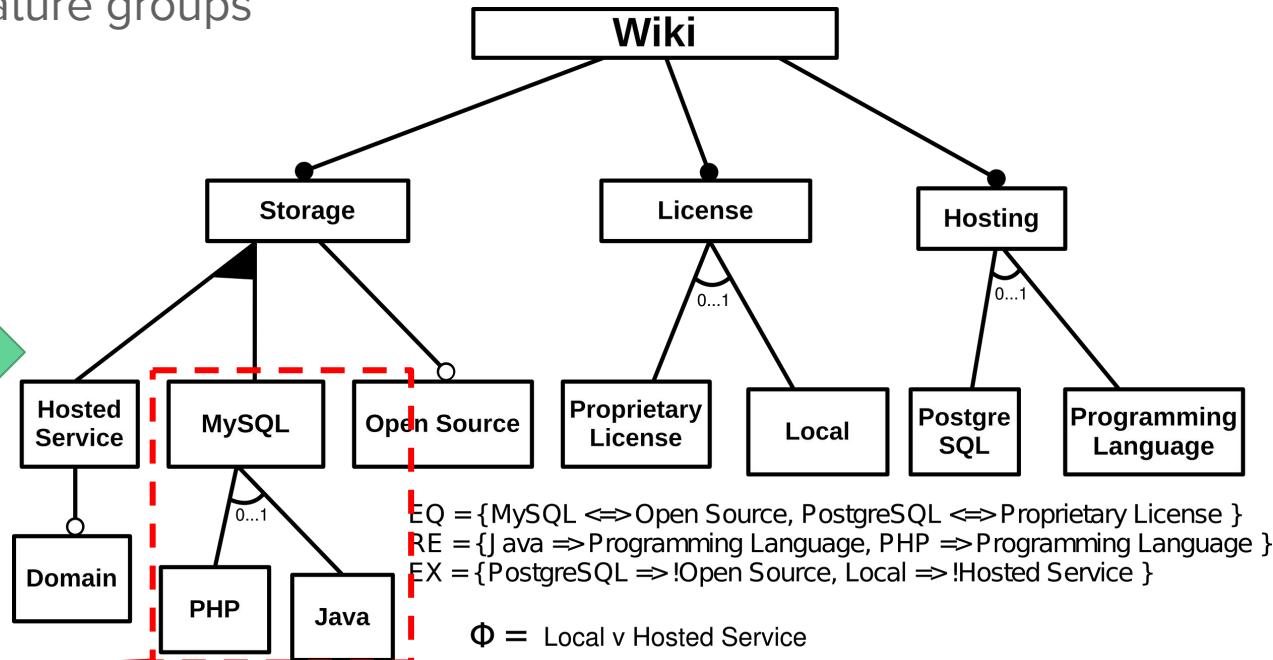
## Ontological semantics

Express the relationships between features

Hierarchy of features + feature groups

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
PostgreSQL	✓	X	X	X	X	X	X	X	X	X
MySQL	X	✓	✓	✓	✓	✓	✓	✓	✓	✓
License	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Domain	✓	✓	X	X	X	X	✓	✓	✓	X
Proprietary License	✓	X	X	X	X	X	✓	✓	X	X
Local	X	X	X	X	X	X	X	X	X	X
Programming Language	X	✓	X	X	✓	✓	✓	✓	✓	✓
Java	X	✓	X	X	X	✓	✓	✓	✓	✓
Storage	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
PHP	X	X	X	X	✓	✓	✓	✓	X	X
Open Source	X	✓	✓	✓	✓	✓	✓	✓	✓	✓
Wiki	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Hosting	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Hosted Service	✓	✓	X	✓	✓	✓	✓	✓	✓	X

synthesis



## Configurator

Which MySQL database do you prefer?

Java

PHP



**State of the art:** few works address both configuration and ontological semantics

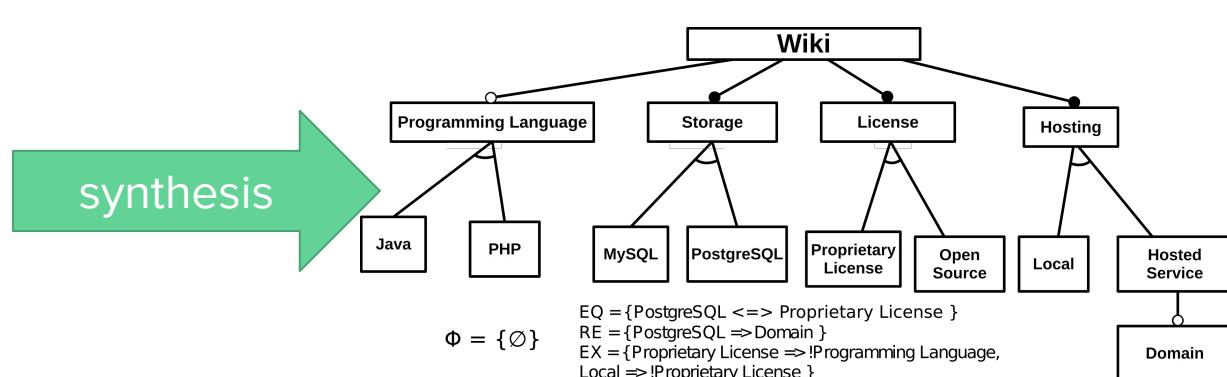
# Contribution

## Product comparison matrices

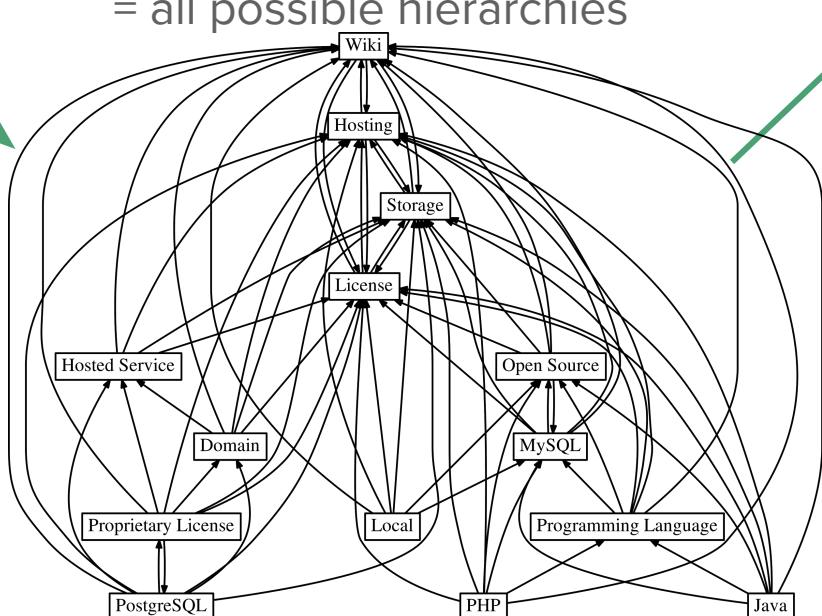
VPN SERVICE Based in IC  
3Monkey Switzerland  
AceVPN USA  
AcetechVPN France  
AirVPN Italy  
Anonmine UK  
AnonymousVPN Spain  
Antivirus / Internet Security Avast Secureline  
Avast Secureline Czech Rep  
Avast Secureline Switzerland  
BlackVPN Hong Kong  
Blockless Canada  
BolehVPN Seychelles  
BTGuard Singapore  
Buffered Hungary  
CactusVPN Moldova  
Celo Australia  
CloudGlobal Germany  
Cloak USA

	Name	No	No	No	No	No	No
3Monkey	No	No	No	No	No	No	No
AceVPN	Five	No	No	No	No	No	No
AcetechVPN	No	No	No	No	No	No	No
AirVPN	No	No	No	No	No	No	No
Anonmine	Five	Yes	No	Yes	Yes	Yes	Yes
AnonymousVPN	Cooperative	No	No	No	No	No	No
Antivirus / Internet Security	Five	No	No	Yes	Yes	Yes	Yes
Avast Secureline	Fourteen	No	No	No	No	No	No
Avast Secureline Czech Rep	Five	Yes	See Note	Yes	Yes	Yes	Yes
Avast Secureline Switzerland	Five	Yes	See Note	Yes	Yes	Yes	Yes
BlackVPN	Hong Kong	Canada	France	Italy	Spain	UK	USA
Blockless	Czech Rep	Canada	France	Italy	Spain	UK	USA
BolehVPN	Singapore	Canada	France	Italy	Spain	UK	USA
BTGuard	Hungary	Canada	France	Italy	Spain	UK	USA
Buffered	Moldova	Canada	France	Italy	Spain	UK	USA
CactusVPN	Australia	Canada	France	Italy	Spain	UK	USA
Celo	Australia	Canada	France	Italy	Spain	UK	USA
CloudGlobal	Germany	Canada	France	Italy	Spain	UK	USA
Cloak	USA	Canada	France	Italy	Spain	UK	USA

## Feature Model

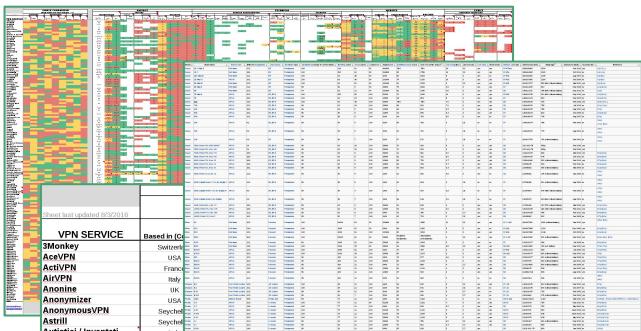


Binary implication graph  
= all possible hierarchies



# Contribution

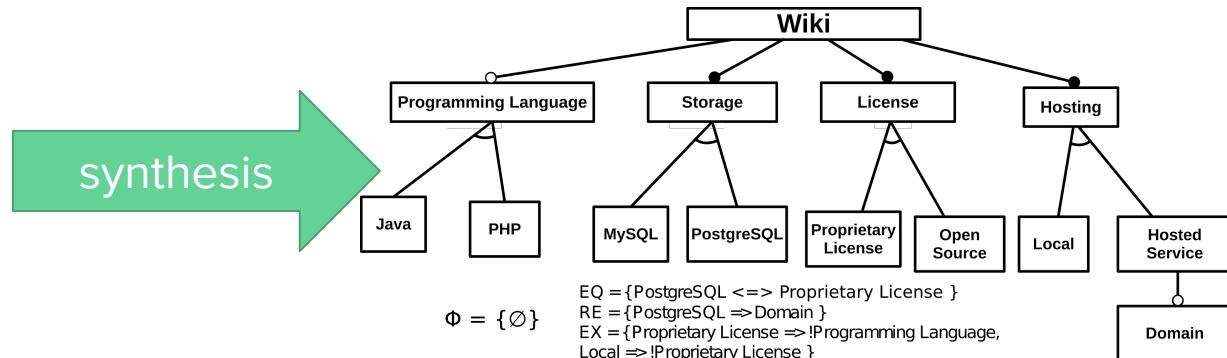
## Product comparison matrices



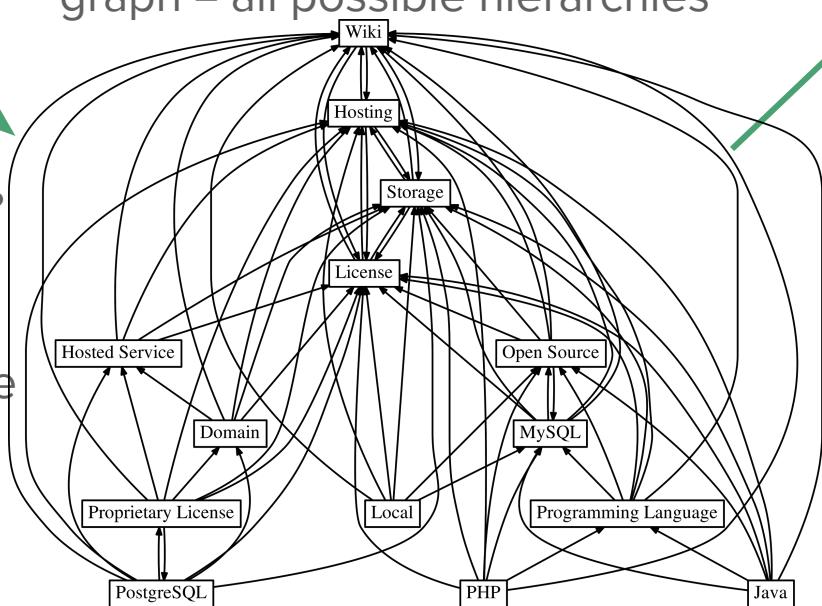
**VPN SERVICE**

	3Monkey	AceVPN	AcetechVPN	AvPN	Anonime	AnonymousVPN	Antonii	Antistatic / Inventati	Avast Secureline	AzireVPN	Balys	BlackVPN	Blockless	BolehVPN	Bsp	BTGuard	Buffeted	CactusVPN	Celo	CloudGlobal	Cloak
Based in IC	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Switzerland	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
USA	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
France	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Belgium	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
UK	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
USA	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Scyld	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Spain	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Italy	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Czech Rep.	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Sweden	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Denmark	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Hong Kong	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Canada	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Seychelles	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Sierra Leone	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
USA	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Hungary	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Moldova	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Australia	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Germany	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
USA	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Fourteen	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
USA	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Five	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
See Note	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

## Feature Model



**Weighted binary implication graph = all possible hierarchies**

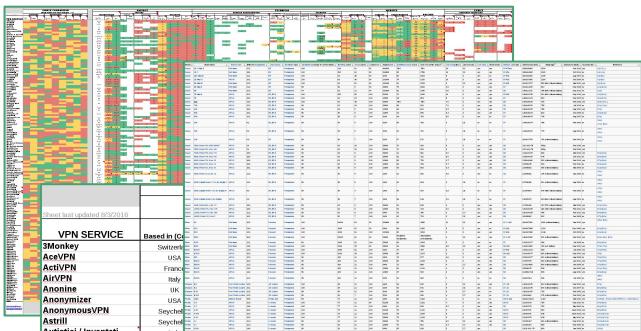


How to define the weights?

- User input
- Ontological heuristics (similarity of the feature names)
- Logical heuristics

# Contribution

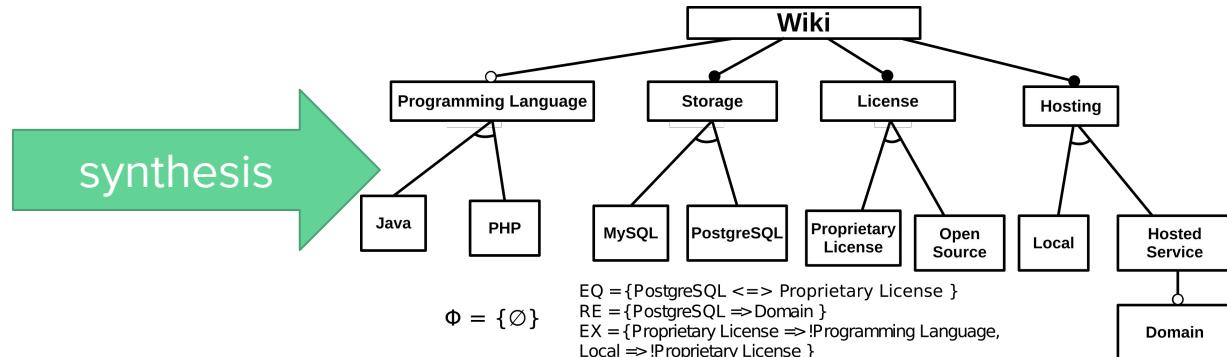
## Product comparison matrices



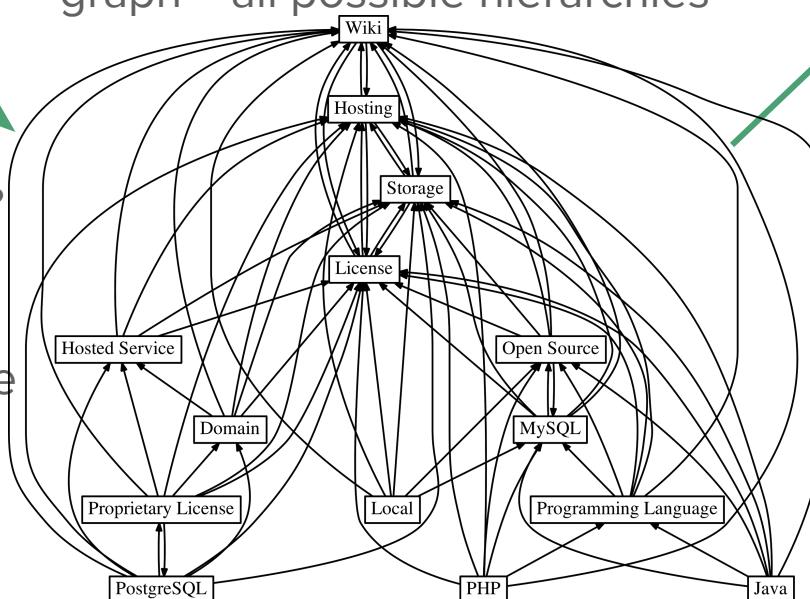
VPN SERVICE Based in IC  
3Monkey Switzerland  
AceVPN USA  
AcetechVPN France  
AvPN UK  
Anonime USA  
AnonymousVPN Spain  
Antivirus / Internet Security Avast Secureline  
AzireVPN  
B4V  
BlackVPN  
Blockless  
BolehVPN  
Bsp  
BTGuard  
Buffered  
CactusVPN  
Celo  
ChinaGlobal  
Cloak

	Name	No	No	No	No	No	No
3Monkey	No	No	No	No	No	No	No
AceVPN	Five	No	No	No	No	No	No
AcetechVPN	No	No	No	No	No	No	No
AvPN	Five	No	No	No	No	No	No
Anonime	No	No	No	No	No	No	No
AnonymousVPN	Five	No	No	No	No	No	No
Antivirus / Internet Security	Cooperative	No	No	No	Yes	Yes	Yes
Avast Secureline	Five	No	No	Yes	No	No	No
AzireVPN	Fourteen	No	No	Yes	Yes	Yes	Yes
B4V	Five	Yes	See Note	Yes	Yes	Yes	Yes

## Feature Model



**Weighted binary implication graph = all possible hierarchies**

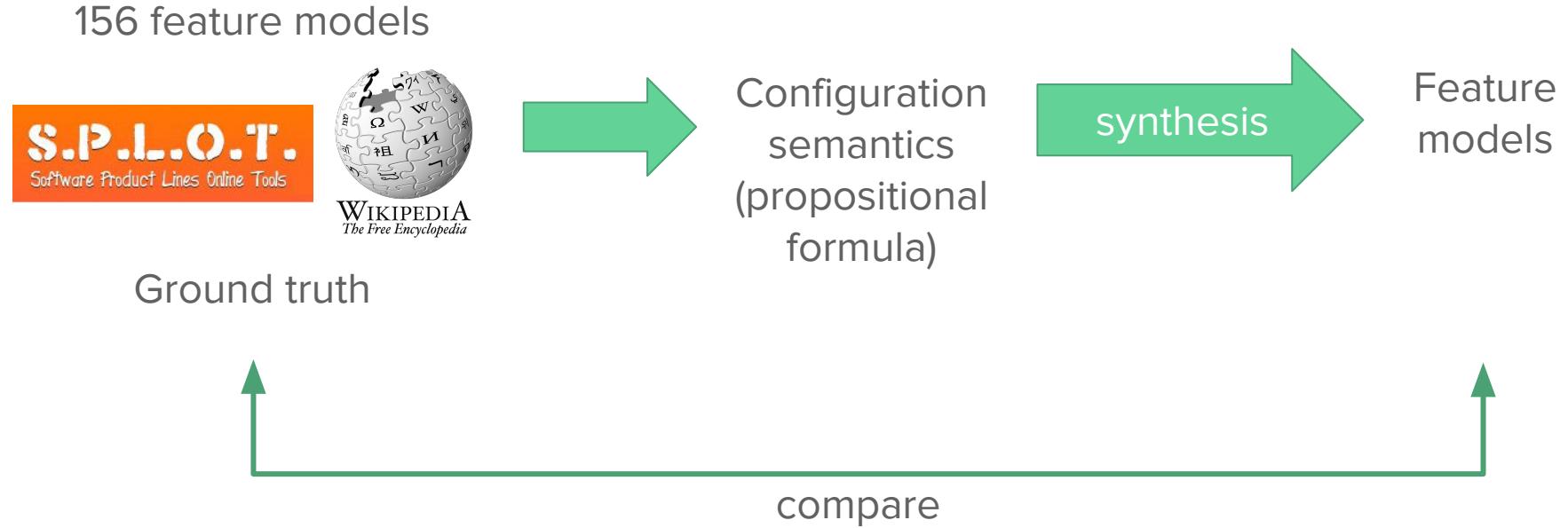


Optimum branching algorithm to extract the best hierarchy

How to define the weights?

- User input
- Ontological heuristics (similarity of the feature names)
- Logical heuristics

# Evaluation of the Ontological Semantics



**Challenge:** synthesize the same feature model as the ground truth

- Hierarchy
- Configuration semantics

# Evaluation of the Ontological Semantics

Key results (Empirical Software Engineering, 2015)

- A fully automated synthesis produces feature models far from ground truth
- User input is necessary to have high quality feature models
- Ontological heuristics improve the quality of ranking lists
- Ontological heuristics produce less clusters than random but they are more accurate

Development of an interactive environment for feature model synthesis (SPLC, 2014)

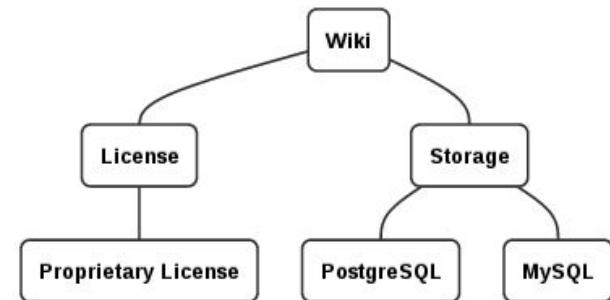
Ranking lists of parent candidates

- + Storage
- Proprietary License
- License ▾
  - PostgreSQL ▾
  - Storage ▾
  - Wiki ▾
- + License
- + PostgreSQL
- + PHP
- + MySQL

Clusters of features

- >
  - Storage
  - Programming Language
- >
  - License
  - Proprietary License
- Storage ▾
  - ✓
  - ✗
- PostgreSQL
- MySQL

Preview of hierarchy



# Feature Model Synthesis

Product comparison  
matrices

Boolean product  
comparison matrices

VPN SERVICE	JURIS	Based in (Country)
3Monkey	Switzerland	USA
AceVPN	France	France
ActiVPN	Italy	UK
AvPN	USA	USA
Anonmine	Seychelles	Seychelles
Anonymizer	Seychelles	Seychelles
AnonymousVPN	Seychelles	Seychelles
Avast!	Czech Republic	USA
Avast! Internet Security	Sweden	Denmark
AzireVPN	Hong Kong	Canada
Betternet	Malta	Seychelles
BlackVPN	Moldova	Seychelles
Blockless	Hungary	Cooperative
BolehVPN	Australia	No
Bogbit	Germany	No
BTGuard	USA	No
Buffered	Moldova	No
CactusVPN	Australia	No
Celo	Germany	No
CloudGlobal	USA	Fourteen
Cloak	USA	Five
		See Note *
		Yes
		Yes
		Yes

synthesis

Feature  
model

Services

Editor

Configurator

• • •

# Feature Model Synthesis

Product comparison  
matrices

Boolean product  
comparison matrices

VPN SERVICE	JURIS	Based in (Country)	...
3Monkey	Singapore	USA	
AceVPN		France	
ActiVPN		Italy	
AnyVPN		UK	
Anonymizer		USA	
AnonymousVPN		Seychelles	
Avast SecureLine		Seychelles	
Autistic Internet		USA	
Avast Secureline	Czech Republic		
AzireVPN		Sweden	
Betternet		Denmark	
BlackVPN		Hong Kong	
Blockless		Canada	
BolehVPN		Seychelles	
Bogon		Seychelles	
BTGuard		USA	
Buffered		Hungary	Cooperative
CactusVPN		Moldova	No No
Celo		Australia	No No
ChiliGlobal		Germany	No No
Cloak		USA	Fourteen
			Five Yes
			See Note
			Yes Yes
			Yes Yes
			Yes Yes

synthesis

attributed  
feature model

Services

Editor

Configurator

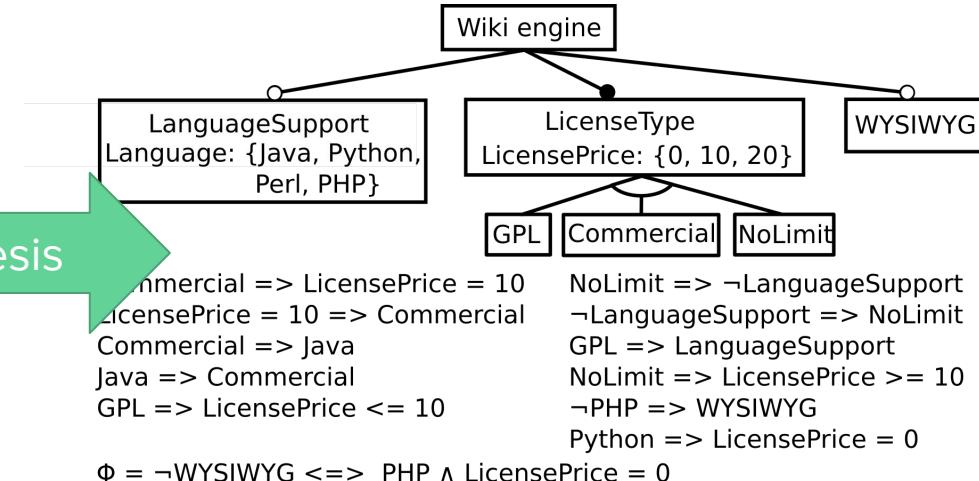
...

**Objective:** increase the number of PCMs supported by feature model synthesis

# From Feature Models to Attributed Feature Models

**Objective:** increase the number of PCMs supported by feature model synthesis

Id	License Type	License Price	Language Support	Lang uage	WYSI WYG
W1	Commercial	10	Yes	Java	Yes
W2	NoLimit	20	No	-	Yes
W3	NoLimit	10	No	-	X
W4	GPL	0	Yes	Python	
W5	GPL	0	Yes	Perl	
W6	GPL	10	Yes	Perl	Yes
W7	GPL	0	Yes	PHP	No
W8	GPL	10	Yes	PHP	Yes



**Attributed feature model** = feature model

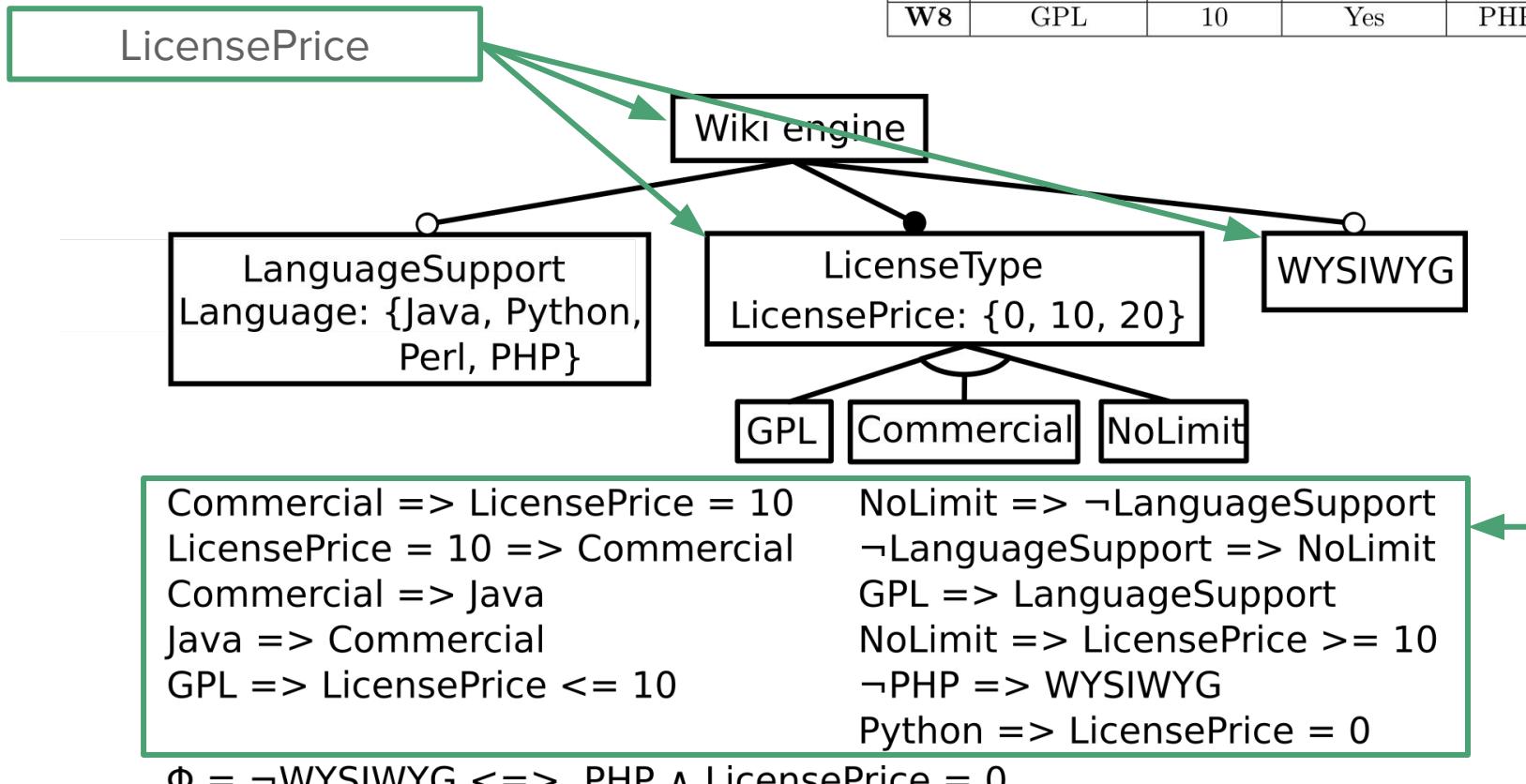
- + Attributes (Booleans, numbers and enumerations)
- + Constraints on attributes

**State of the art:** no synthesis algorithm for attributed feature models

# Attributed Feature Model Synthesis

Key differences with feature model synthesis

- Place of attributes
- Generation of constraints on attributes



# Contributions

First algorithm for synthesizing attributed feature models from PCMs (SPLC, 2015)

The diagrammatic part of an attributed feature model is an over-approximation of the PCM

Computation of sound and complete data structures that can guide the user during the synthesis

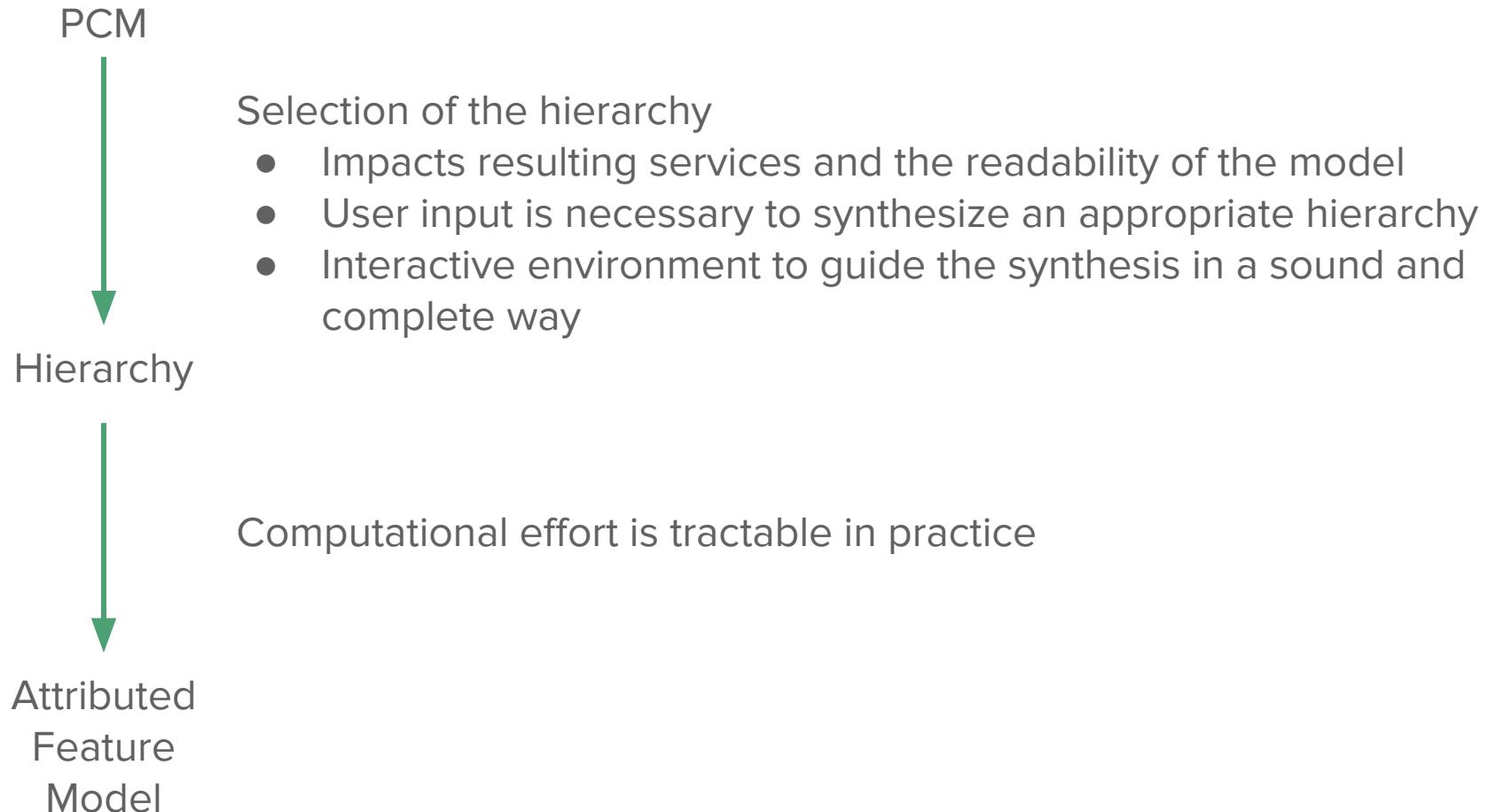


Scalability of the synthesis algorithm may be impacted by attributes

- Computation of or-groups does not scale
- The rest of the computation has a polynomial complexity in practice

# Conclusion

Characterization of the gap between PCMs and attributed feature models

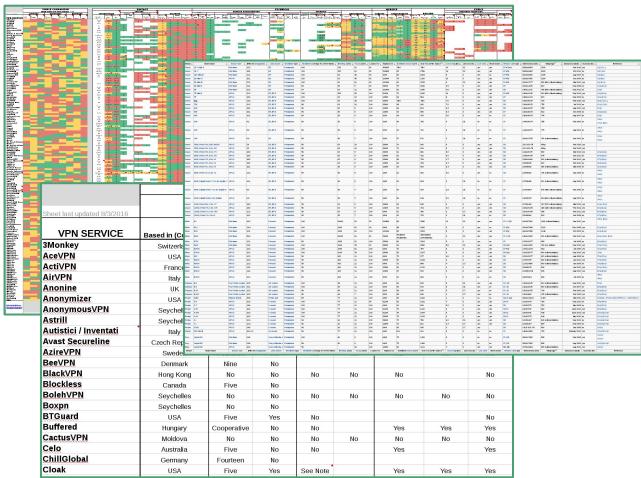


# Conclusion and Perspectives

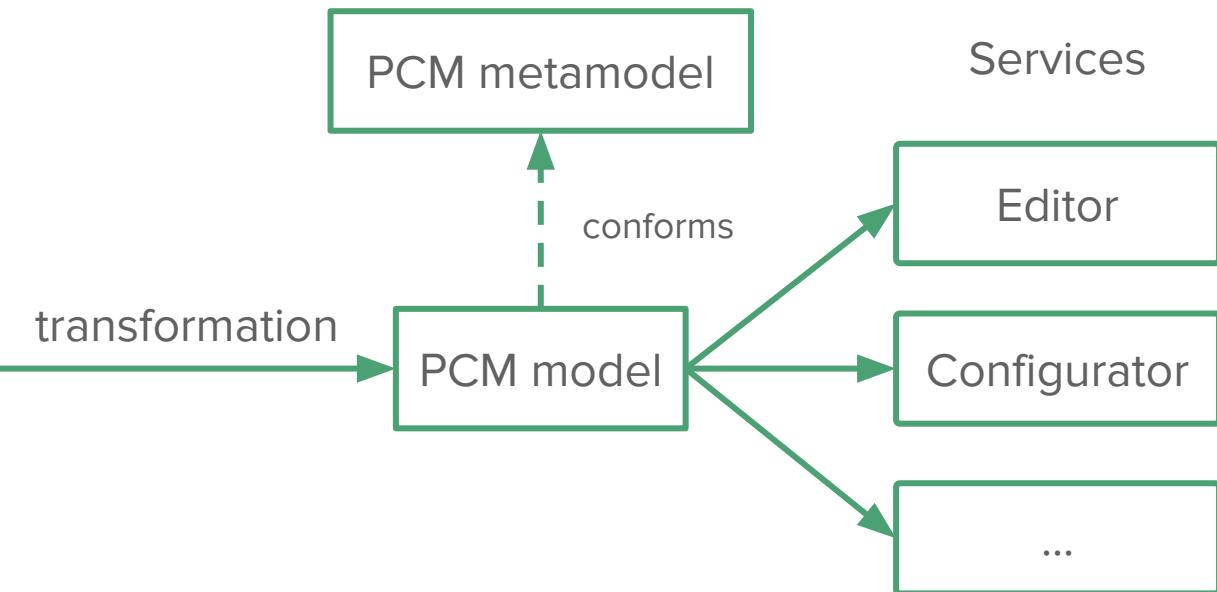
---

# Contributions

Billions of product comparison matrices



A screenshot of a Microsoft Excel spreadsheet titled "Sheet1 last updated 03/29/2016". The spreadsheet contains two main sections: "VPN SERVICE" and "Smartphones". The "VPN SERVICE" section has a legend for colors (red, green, yellow, blue) and lists various services like AceVPN, ActiVPN, AirVPN, and Anonymizer. The "Smartphones" section lists phones from brands like Apple, Samsung, Google, and LG, each with a color-coded cell. Below these are two tables. The first table compares features across phones, with columns for "Screen", "Processor", "RAM", "Storage", "Battery", and "Price". The second table is a smaller subset with columns for "Screen", "Processor", "RAM", "Storage", and "Price". Both tables have rows for different phones.



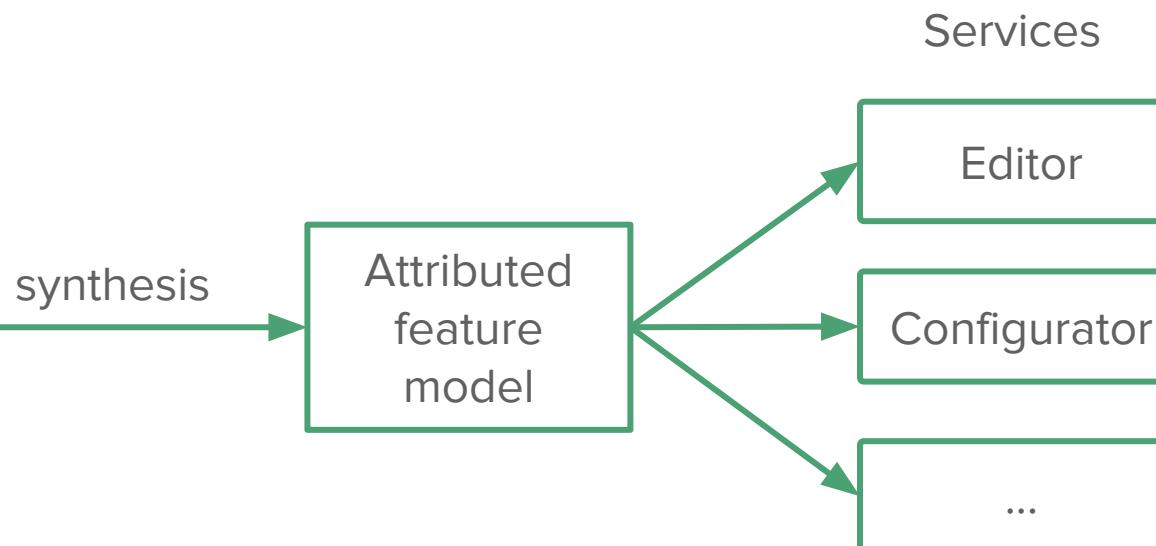
Metamodelling is a multi-objective problem

- Iterative process to cover a maximum of data
- Validation of the metamodel and transformation by end-users through dedicated tools
- Developers of services influence the design of the approach

# Contributions

Billions of product comparison matrices

Comparison Matrix (Partial View)						
VPN SERVICE	Based in (C)	Country	Support	Country	Support	Country
3Monkey	Singapore	Hong Kong	No	No	No	No
ActivPN	USA	Canada	No	No	No	No
AirVPN	France	Germany	No	No	No	No
Anonime	India	UK	No	No	No	No
Anonymizer/VPN	USA	USA	No	No	No	No
Astrill	Sweden	Sweden	No	No	No	No
Autistic / Inventati	Seychelles	Italy	No	No	No	No
Awest Secureline	Italy	Czech Republic	No	No	No	No
Astrofile	Denmark	Denmark	No	No	No	No
BeeVPN	Hong Kong	Hong Kong	No	No	No	No
BlackVPN	Canada	Canada	No	No	No	No
Blockless	Singapore	Singapore	No	No	No	No
BoxyPN	Seychelles	Seychelles	No	No	No	No
BTPGuard	USA	Five	Yes	No	No	No
Buferred	Hungary	Cooperative	No	No	Yes	Yes
CactusVPN	Malta	Malta	No	No	No	No
Celo	Australia	Five	No	No	Yes	Yes
ChilliGlobal	Germany	Fourteen	No	See Note	Yes	Yes
Cloak	USA	Five	Yes	See Note	Yes	Yes



Characterization of the gap between PCMs and attributed feature models

Algorithms and interactive environment to synthesize a sound and complete feature model with an appropriate hierarchy

- Opens to advanced reasoning techniques for PCMs
- Contribute to the community of software product line engineering

# Perspectives

Going beyond attributed feature models

Billions of product comparison matrices

VPN SERVICE	Attributed Feature Model								
	3Monkey	Acetech	AirVPN	Anonime	Anonymizer/VPN	Astirli	Autistic / Inventati	Aveast Secureline	Astrosoft
Based in (C)	Singapore	USA	France	India	UK	USA	Germany	Italy	Czech Rep.
3Monkey	No	No	No	No	No	No	No	No	No
Acetech	No	No	No	No	No	No	No	No	No
AirVPN	No	No	No	No	No	No	No	No	No
Anonime	No	No	No	No	No	No	No	No	No
Anonymizer/VPN	No	No	No	No	No	No	No	No	No
Astirli	No	No	No	No	No	No	No	No	No
Autistic / Inventati	No	No	No	No	No	No	No	No	No
Aveast Secureline	No	No	No	No	No	No	No	No	No
Astrosoft	No	No	No	No	No	No	No	No	No
BeeVPN	Denmark	None	No	No	No	No	No	No	No
BlackVPN	Hong Kong	No	No	No	No	No	No	No	No
Blitzless	FYR Macedonia	No	No	No	No	No	No	No	No
BolehVPN	Senegal	No	No	No	No	No	No	No	No
Boxpn	Seychelles	No	No	No	No	No	No	No	No
BTGuard	USA	Five	Yes	No	No	No	No	No	No
Buffered	Hungary	Cooperative	No	No	Yes	Yes	Yes	Yes	Yes
CactusVPN	Moldova	Five	No	No	No	No	No	No	No
Celo	Australia	Five	No	No	Yes	No	No	No	Yes
ChillGlobal	Germany	Fourteen	No	See Note	Yes	Yes	Yes	Yes	Yes
Cloak	USA	Five	Yes	See Note	Yes	Yes	Yes	Yes	Yes

synthesis

Evaluate user effort

Attributed  
feature  
model

Services

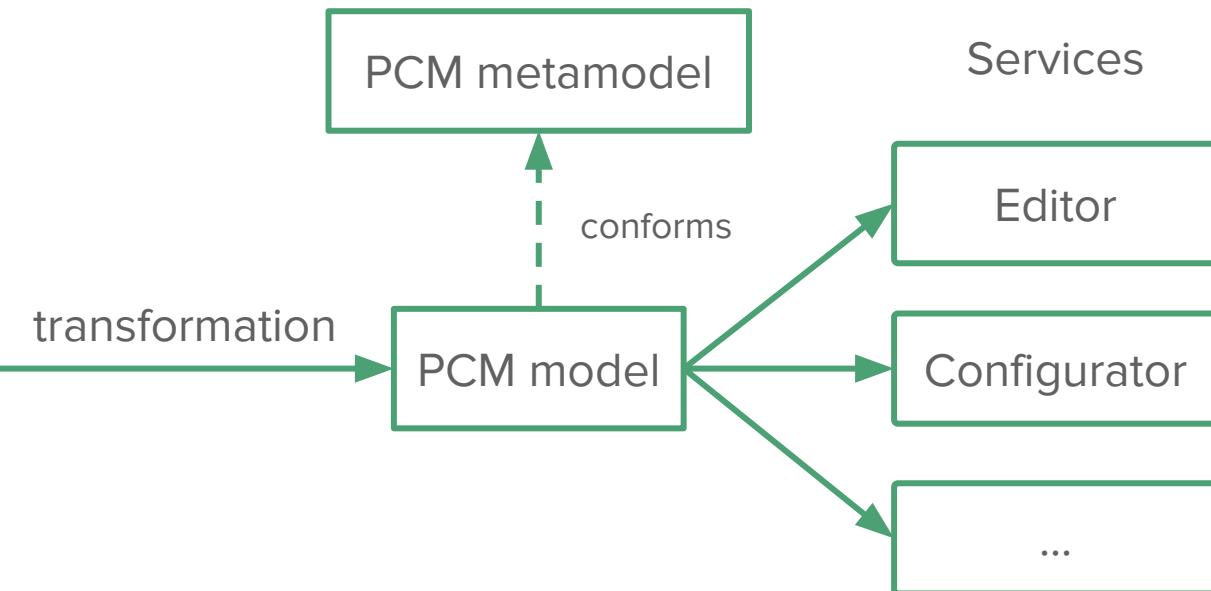
Editor

Configurator

• • •

# Perspectives

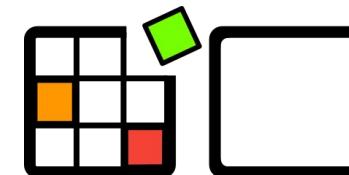
# Billions of product comparison matrices



## Data driven: How to use machine learning to discover and classify new types of cells?

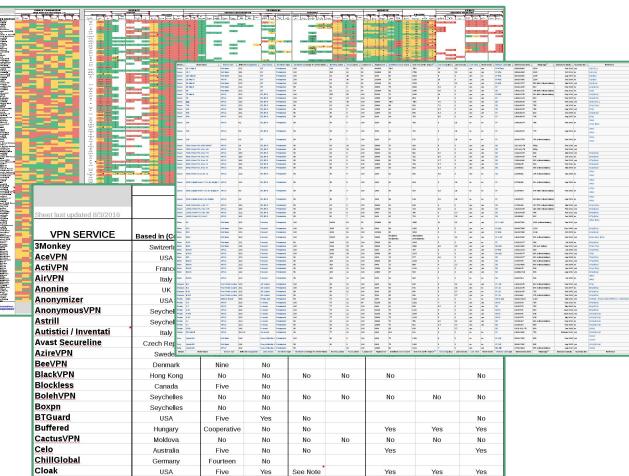
## **Service driven:** How to evaluate a metamodel or compare multiple metamodels according to a set of objectives?

## User driven: How to involve users in OpenCompare?



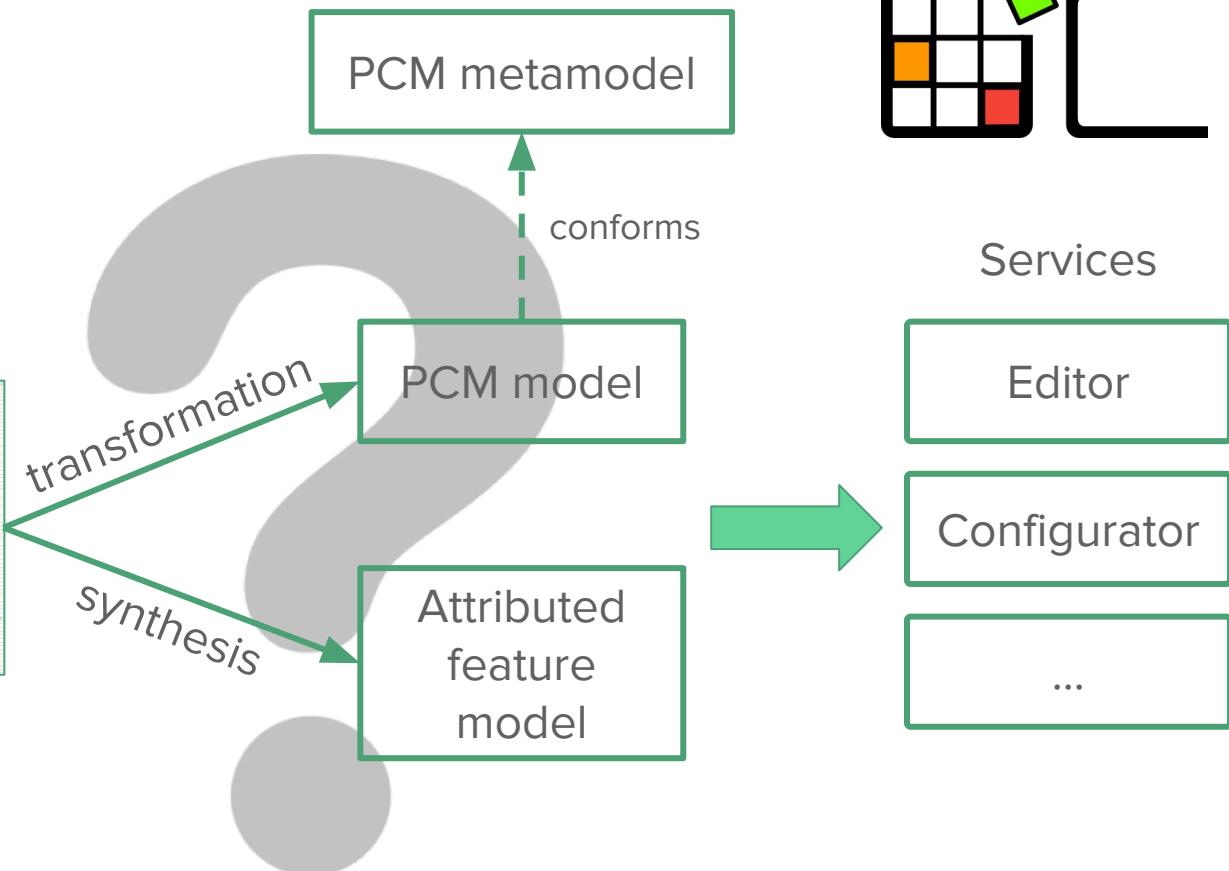
# How to formalize product comparison matrices?

Billions of product comparison matrices



A screenshot of a web-based application showing a large product comparison matrix. The top section displays a heatmap of service names. Below is a detailed table with columns for 'VPN SERVICE' (e.g., 3Monkey, ActiVPN, Anonymizer, Astrill, AzireyVPN, BeeVPN, BlackVPN, Blockless, BokehVPN, Boxpn, BTGuard, Buffered, CasteVPN, Celso, Cloak) and rows for various countries (e.g., Denmark, Hong Kong, Canada, Seychelles, USA, Hungary, Maldives, Australia, Germany). The table contains binary values (Yes/No) indicating the presence of specific features.

VPN SERVICE	Based in (C)									
3Monkey	No	No	No	No	No	No	No	No	No	No
ActiVPN	USA									
Anonymizer	France									
Astrill	Italy									
AzireyVPN	UK									
BeeVPN	USA									
BlackVPN	Seychelles									
Blockless	Denmark	Yes	No	No	No	No	No	No	No	No
BokehVPN	Hong Kong	No	No	No	No	No	No	No	No	No
Boxpn	Canada	Five	No	No	No	No	No	No	No	No
BTGuard	Seychelles	No	No	No	No	No	No	No	No	No
Buffered	USA	Five	Yes	No	No	No	No	No	No	No
CasteVPN	Hungary	Cooperative	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Celso	Maldives	No	No	No	No	No	No	No	No	No
Cloak	Australia	Five	No	No	No	Yes	Yes	Yes	Yes	Yes
	Germany	Fourteen	No	See Note	Yes	Yes	Yes	Yes	Yes	Yes



# Appendix

---

# OpenCompare

# Case Study

5,000,000+  
articles



**WIKIPEDIA**  
*The Free Encyclopedia*

Written in  
Wikitext

Various domains

Numerous  
contributors

1,500,000+  
tables

Shopping List				
Bread & Butter	Pie	Buns	Danish	Croissant
	Cheese	Ice cream		Butter Yogurt

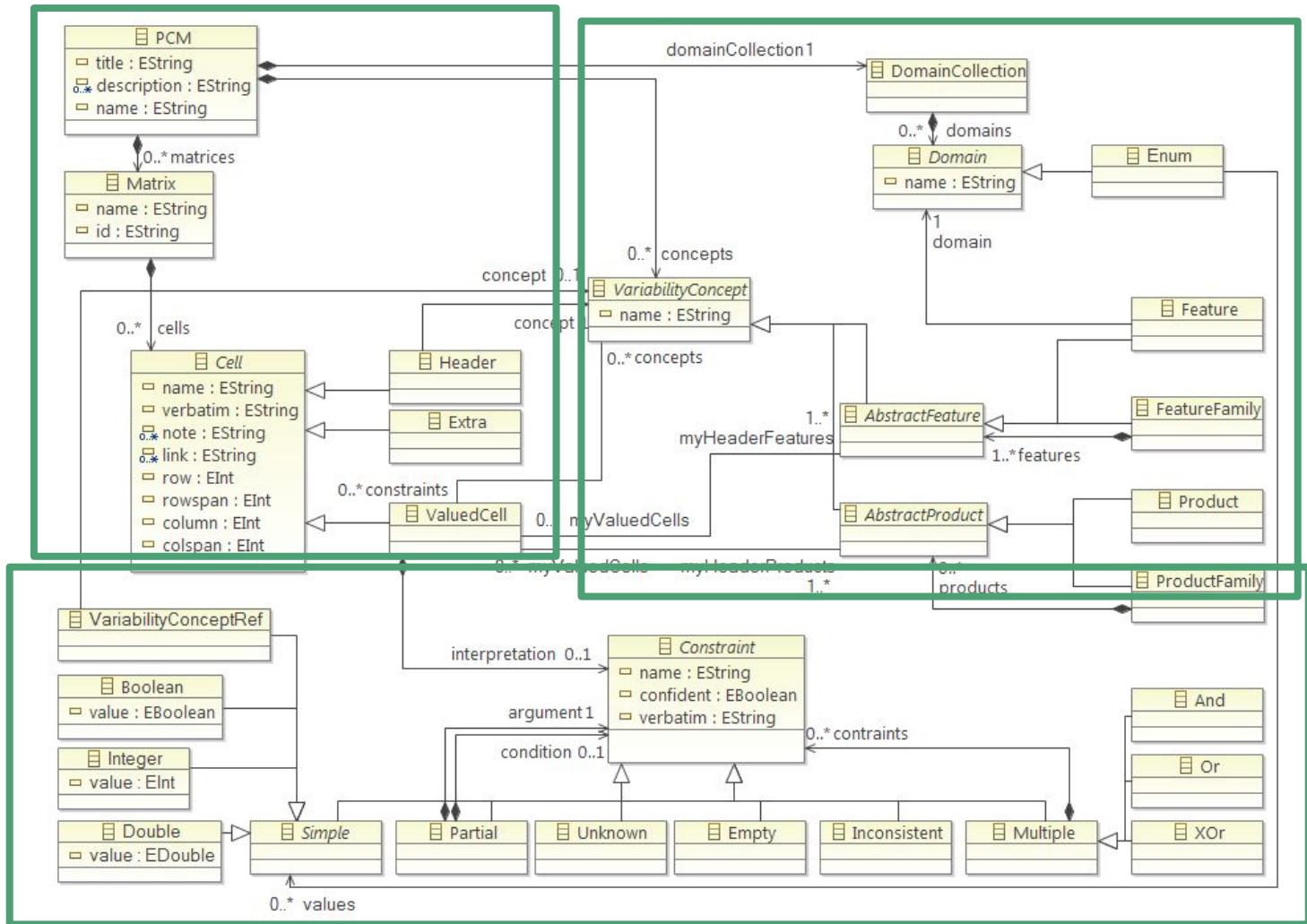
```
{| class="wikitable"
!colspan="6" | Shopping List
|- rowspan="2" | Bread & Butter
| Pie
| Buns
| Danish
| colspan="2" | Croissant
|- rowspan="2" | Ice cream
| Butter
| Yogurt
|}
```

# Product Comparison Matrix Metamodel

**Dataset:** 50 PCMs (“Comparison of...”)

Concepts of product lines

Structure



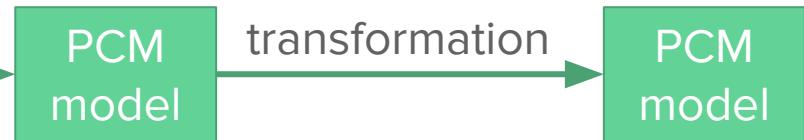
Semantics

# First Iteration: Transformation Chain

**Dataset:** 50 PCMs (“Comparison of...”)

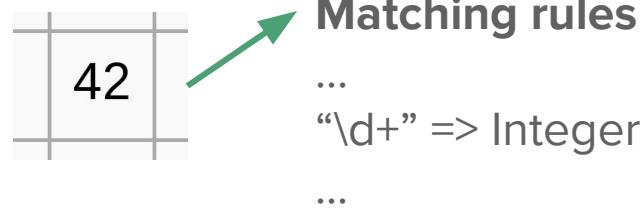
```
{| class="wikitable sortable"
|-  
! Brand  
! Model name  
! [[Image sensor format|Sensor size]]  
! Effective [[megapixel]]s  
! [[Lens mount]]  
! [[Viewfinder]] type
```

parsing  
preprocessing

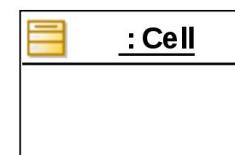


**Main step:** interpret cells based on a set of syntactic rules (regular expressions)

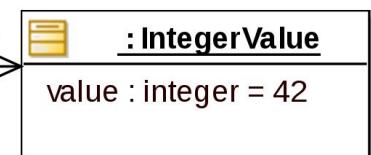
Use the iterative process to define the syntactic rules



match

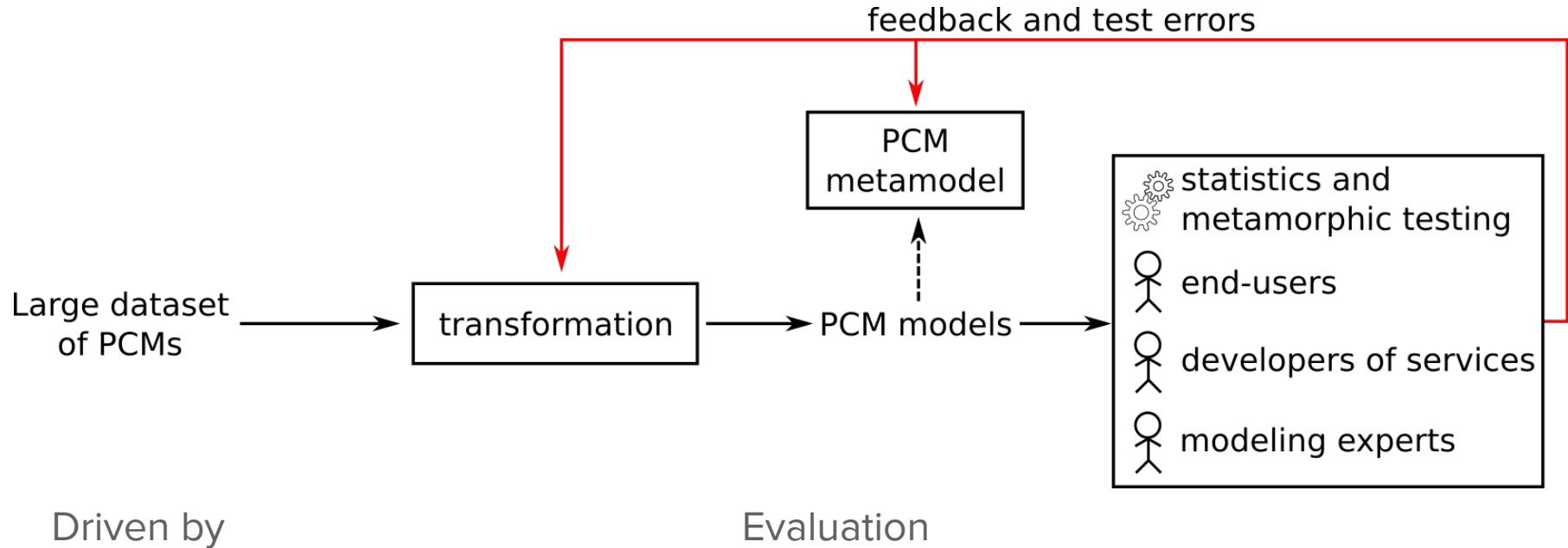


interpretation



# Design through an Iterative Process

**Objective:** define a metamodel and a transformation



Driven by

- Data
- Users
- Services

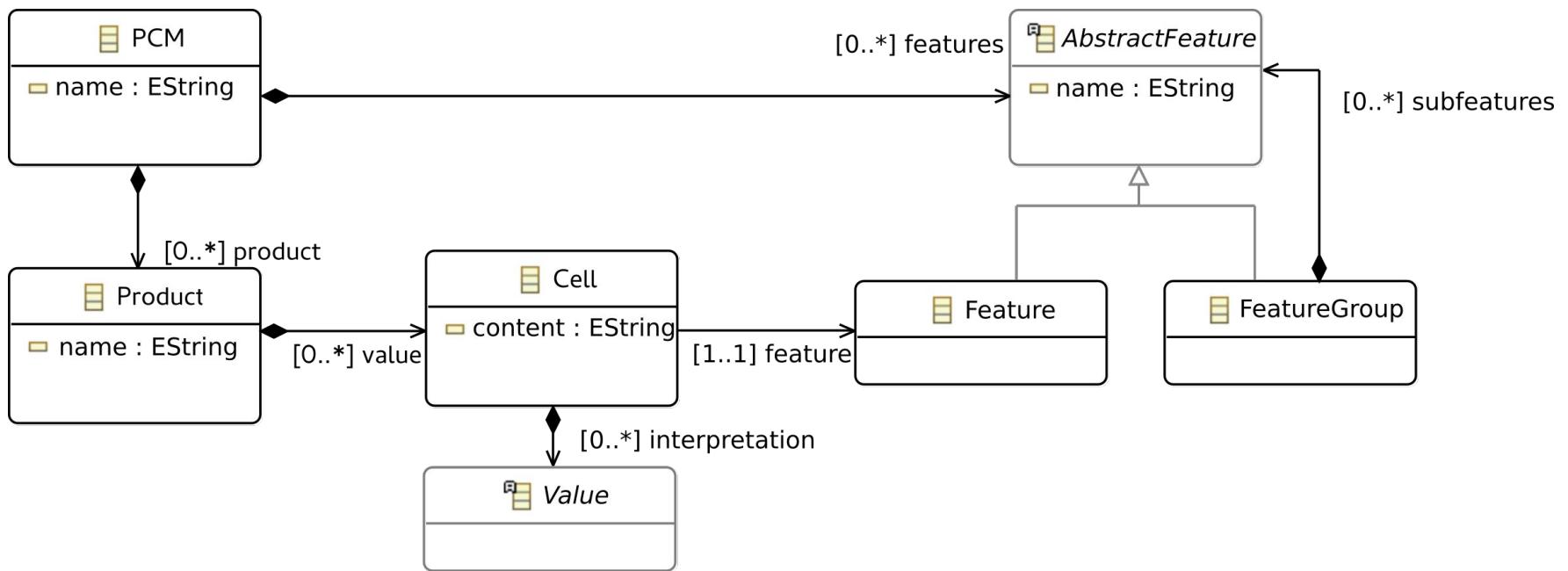
Evaluation

- Automated analyses on the PCM models
- Manual checking by end-users
- Feedback from developers of services based on the metamodel

# Second Iteration: Multiple Metamodels

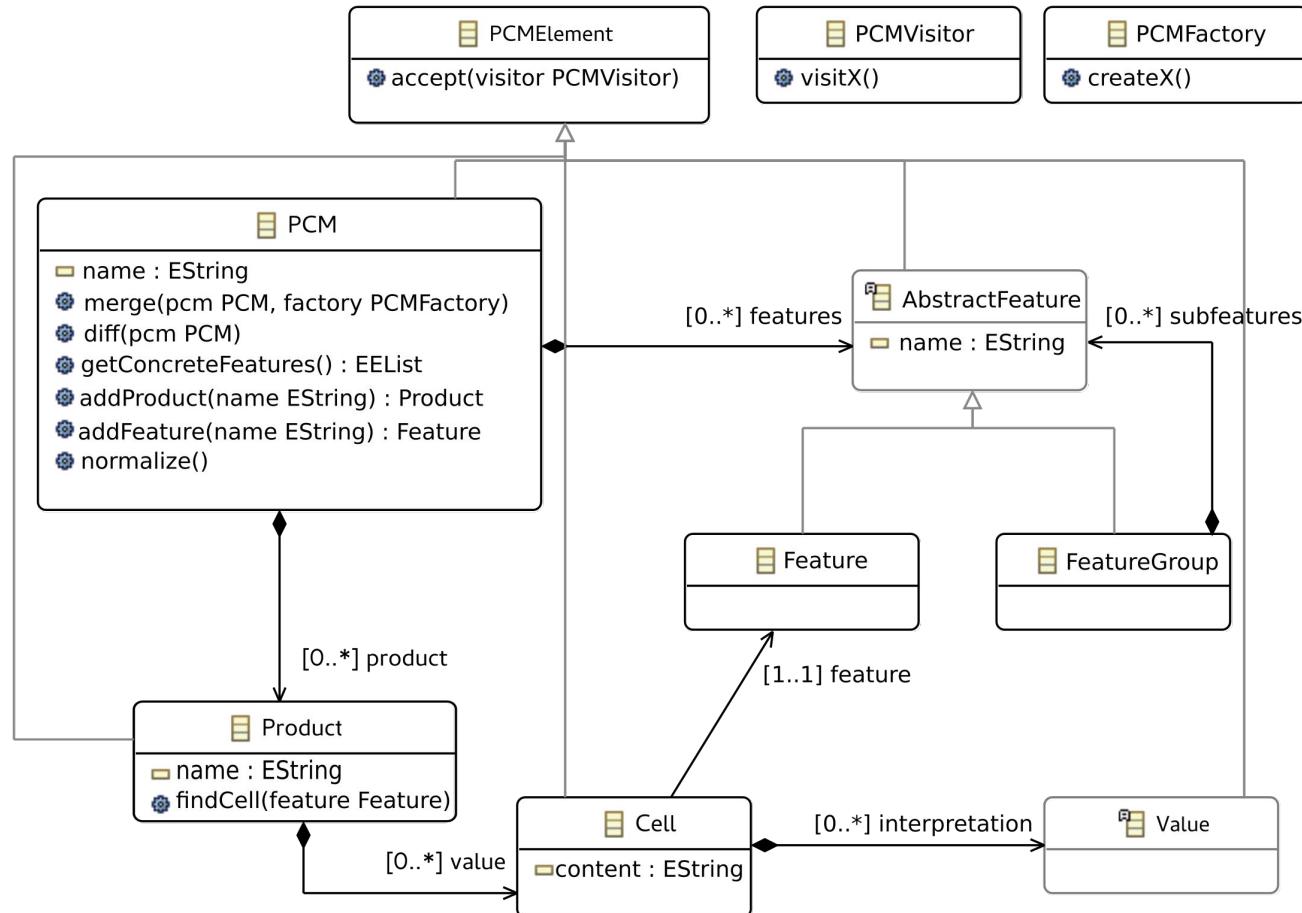
**Objective:** separate each concern in a dedicated metamodel

**Domain metamodel:** description of the domain



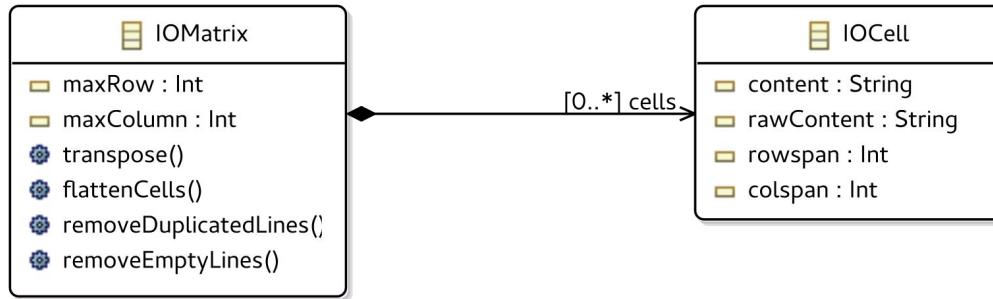
# Second Iteration: Multiple Metamodels

**Manipulation:** generic operations on PCMs

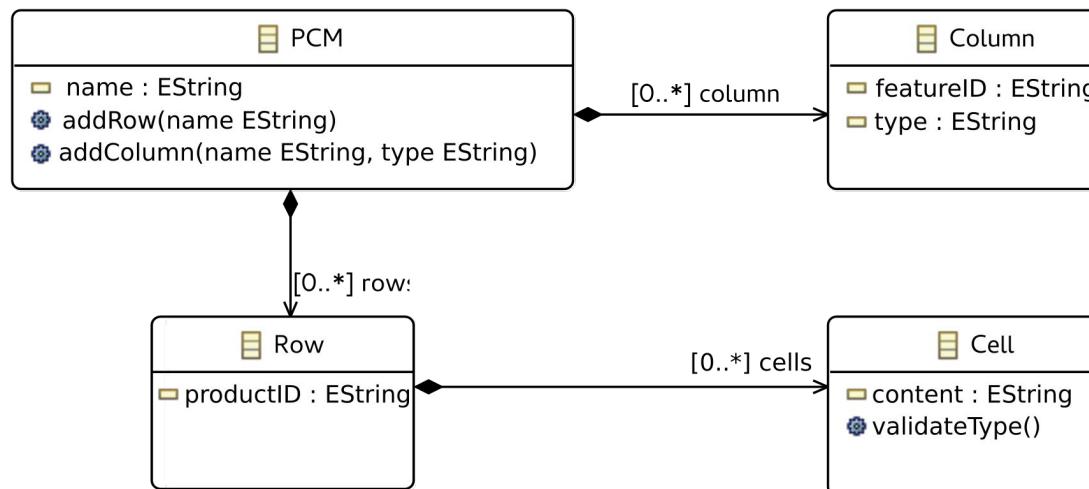


# Second Iteration: Multiple Metamodels

**Import / export:** structure of common table formats (CSV, HTML, Wikitext)



**Edition:** focus on the position of the cells



# Metamorphic testing



WIKIPEDIA  
The Free Encyclopedia

```
{| class="wikitable sortable"
|-  
! Brand  
! Model name  
! [[Image sensor format|Sensor  
size]]  
! Effective [[megapixel]]s  
! [[Lens mount]]  
! [[Viewfinder]] type
```

1,500,000 PCMs

import

PCM model

export

equals?

```
{| class="wikitable sortable"
|-  
! Brand  
! Model name  
! [[Image sensor format|Sensor  
size]]  
! Effective [[megapixel]]s  
! [[Lens mount]]  
! [[Viewfinder]] type
```

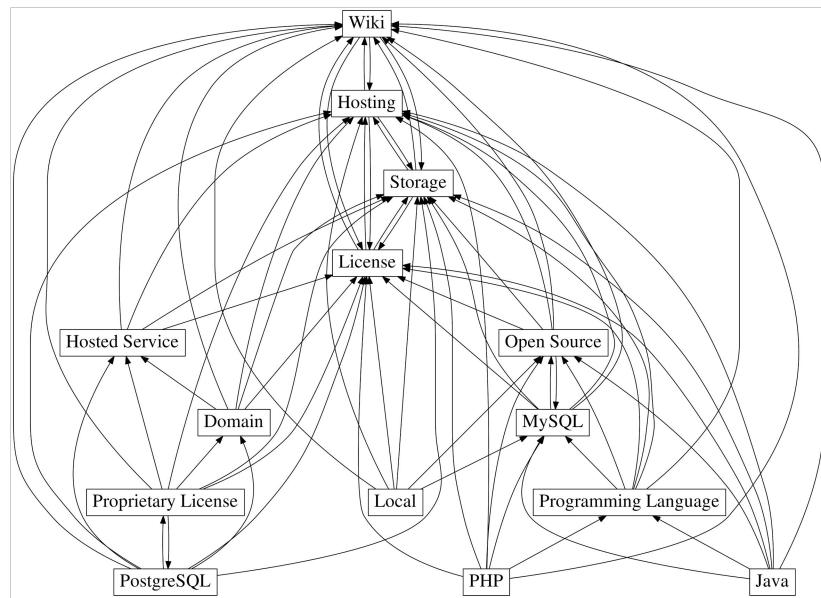
import

PCM model

# Feature Model Synthesis

# Synthesis of the Hierarchy of a Feature Model

**Weight:** similarity of features



Binary implication graph  
= all possible hierarchies

**Heuristics to define the weights**

Ontological: compute the similarity of the feature names

- Wordnet (Wu & Palmer + PathLength)
- Wikipedia, Wiktionary
- Edit distance (Levenshtein, Smith-Waterman)

Logical

- Transitive reduction of the binary implication graph

Hybrid = ontological + logical

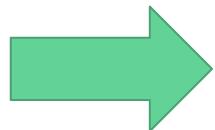
Optimum branching algorithm to extract a valid hierarchy that optimize the weights

# How to compute constraints on attributes?

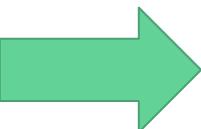
root	F0	F1	F2	F3	A1	A2
Yes	Yes	Yes	No	Yes	3	0
Yes	Yes	No	Yes	Yes	2	2
Yes	Yes	Yes	No	No	2	0
Yes	Yes	No	Yes	No	0	8

## Binary implications

$$\begin{array}{ll} A1 = 0 \Rightarrow A2 \in \{8\} & A2 = 0 \Rightarrow A1 \in \{2,3\} \\ A1 = 2 \Rightarrow A2 \in \{0,2\} & A2 = 2 \Rightarrow A1 \in \{2\} \\ A1 = 3 \Rightarrow A2 \in \{0\} & A2 = 8 \Rightarrow A1 \in \{0\} \end{array}$$



## Merge



Readable constraints  
 $A1 > 0 \Rightarrow A2 \leq 2$   
 $A2 < 8 \Rightarrow A1 \geq 2$

## Domain knowledge

Interesting values for attributes

$A1 \rightarrow 0$

$A2 \rightarrow 8$

## Grammar of readable constraints

```
RC      ::= bool_factor '⇒' bool_factor
bool_factor ::= feature_name | '¬' feature_name | rel_expr
rel_expr  ::= attribute_name rel_op num_literal
rel_op    ::= '>' | '<' | '≥' | '≤' | '='
```

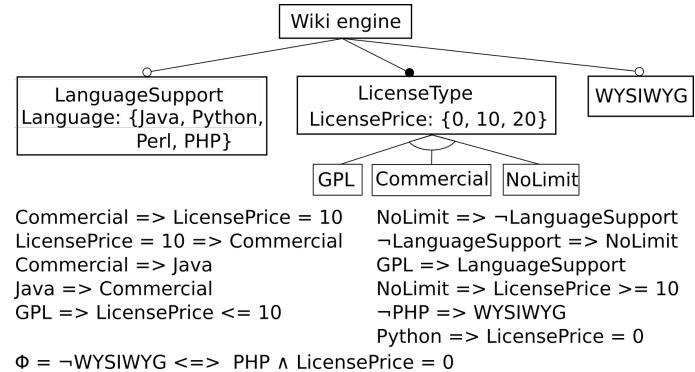
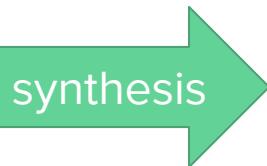
# Evaluation of the Scalability

## Generator of random PCMs

- Number of variables (features + attributes)
- Number of configurations (products)
- Maximum domain size (number of distinct values in a column)

Id	License Type	License Price	Language Support	Language	WYSI WYG
W1	Commercial	10	Yes	Java	Yes
W2	NoLimit	20	No	—	Yes
W3	NoLimit	10	No	—	Yes
W4	GPL	0	Yes	Python	Yes
W5	GPL	0	Yes	Perl	Yes
W6	GPL	10	Yes	Perl	Yes
W7	GPL	0	Yes	PHP	No
W8	GPL	10	Yes	PHP	Yes

Random PCMs

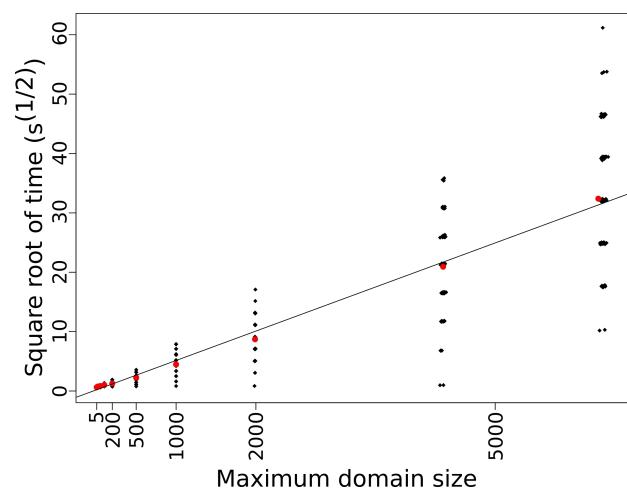
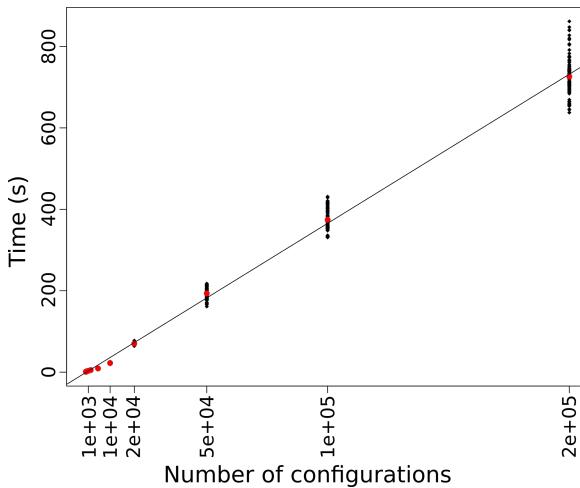
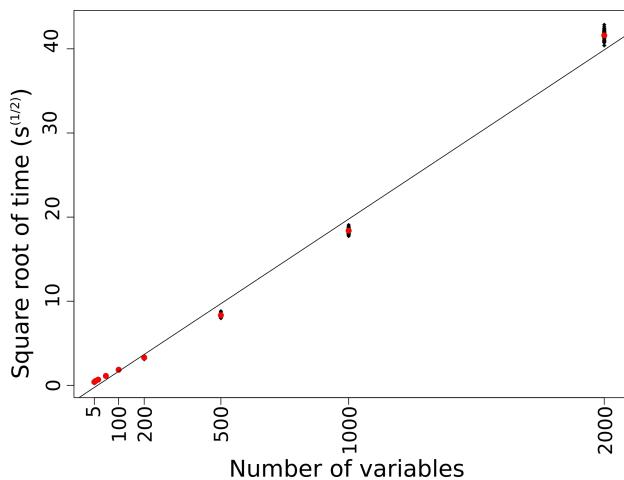


Measure execution time

Computation of Or-groups is deactivated because it does not scale (NP-complete problem)

# Evaluation of the Scalability

## Execution time (no or-groups)



- Up to 2,000 variables
- 1,000 configurations
- Max domain size of 10

- 100 variables
- Up to 200,000 configurations
- Max domain size of 10

- 10 variables
- 10,000 configurations
- Up to 6,000 distinct values

(features + attributes)<sup>2</sup>



configurations



(maximum domain size)<sup>2</sup>

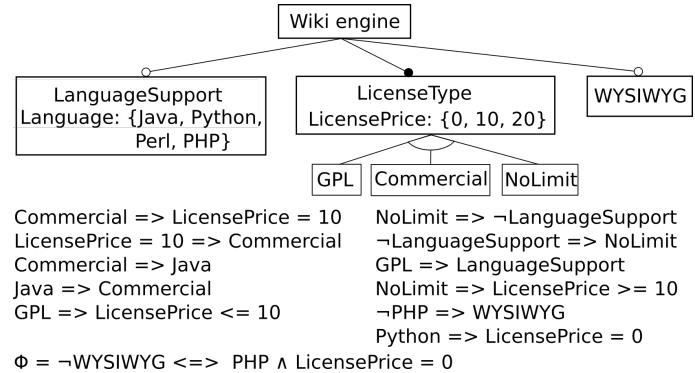
On all experiments: 2.6 min in average, 62 min max

# Evaluation of the Scalability

## Realistic matrices



242 PCMs



Execution time: 0.5s in average

Execution time of 2.1s for the most challenging matrix:

- 77 variables (feature + attributes)
- 185 configurations
- Maximum domain size of 185

Execution time is similar to the random dataset