



# Model Checking Contest results for 2022

Fabrice Kordon — LIP6, Sorbonne Université, France

Pierre Bouvier — Inria/LIG, France

Hubert Garavel — Inria/LIG, France








Francis Hulin-Hubard — LIP6, CNRS/Sorbonne Université, France

Emmanuel Paviot-Adet — LIP6 & Université Paris Descartes, France

Loïc Jézequel — LS2N, Université de Nantes, France

# INCC 2022

## Objective 1 : promoting model checking tools

-  Asynchronous systems
-  Compare and debug
-  Enhance reproducibility of results
-  Encourage tools
  -  To enhance reliability
  -  To increase their features
-  Encourage interoperability among tools

## Objective 2 : creating a common collection of benchmarks

-  <https://mcc.lip6.fr/models.php>

**Hubert Garavel**  
(Inria)



**Pierre Bouvier**  
(Inria)

**Managing  
Models**

**Managing  
Execution +  
analysis**

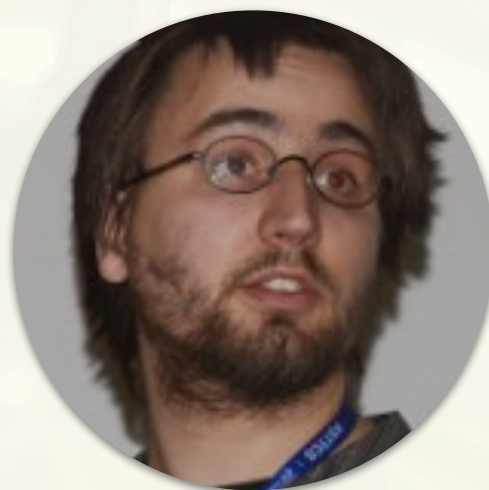
**Fabrice Kordon**  
(SU)



**Francis Hulin-Hubard**  
(CNRS)



**Loïg Jezequel**  
(U. Nantes)



**Emmanuel Paviot-Adet**  
(UP5)



**Managing  
Formulas**

# New Models etc.

## «known» models

- Those from past years
  - ▶ Test the tool as used by its developers

Model  $\Rightarrow$  one Petri net, possibly with scaling parameter

Instance  $\Rightarrow$  one Petri-net with scaling parameters instantiated

## «Surprise» models

- New models proposed by the community this year
  - ▶ Test the tool as used by «non-expert» of the tool
  - ▶ new situations for the tool

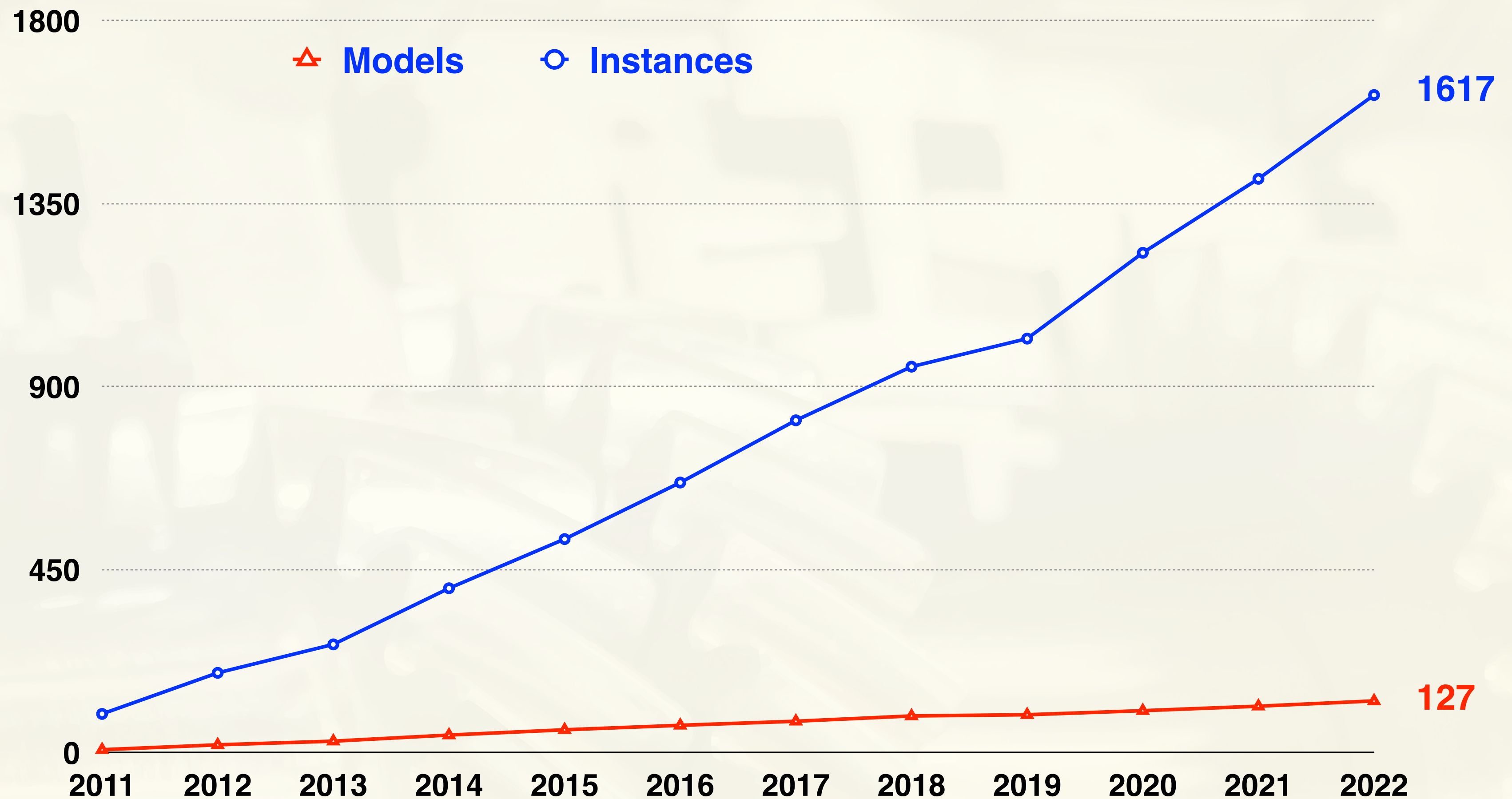
### Coefficients in 2022

1411 «known» instances (x 1)  
206 «surprise» instances (x 6)

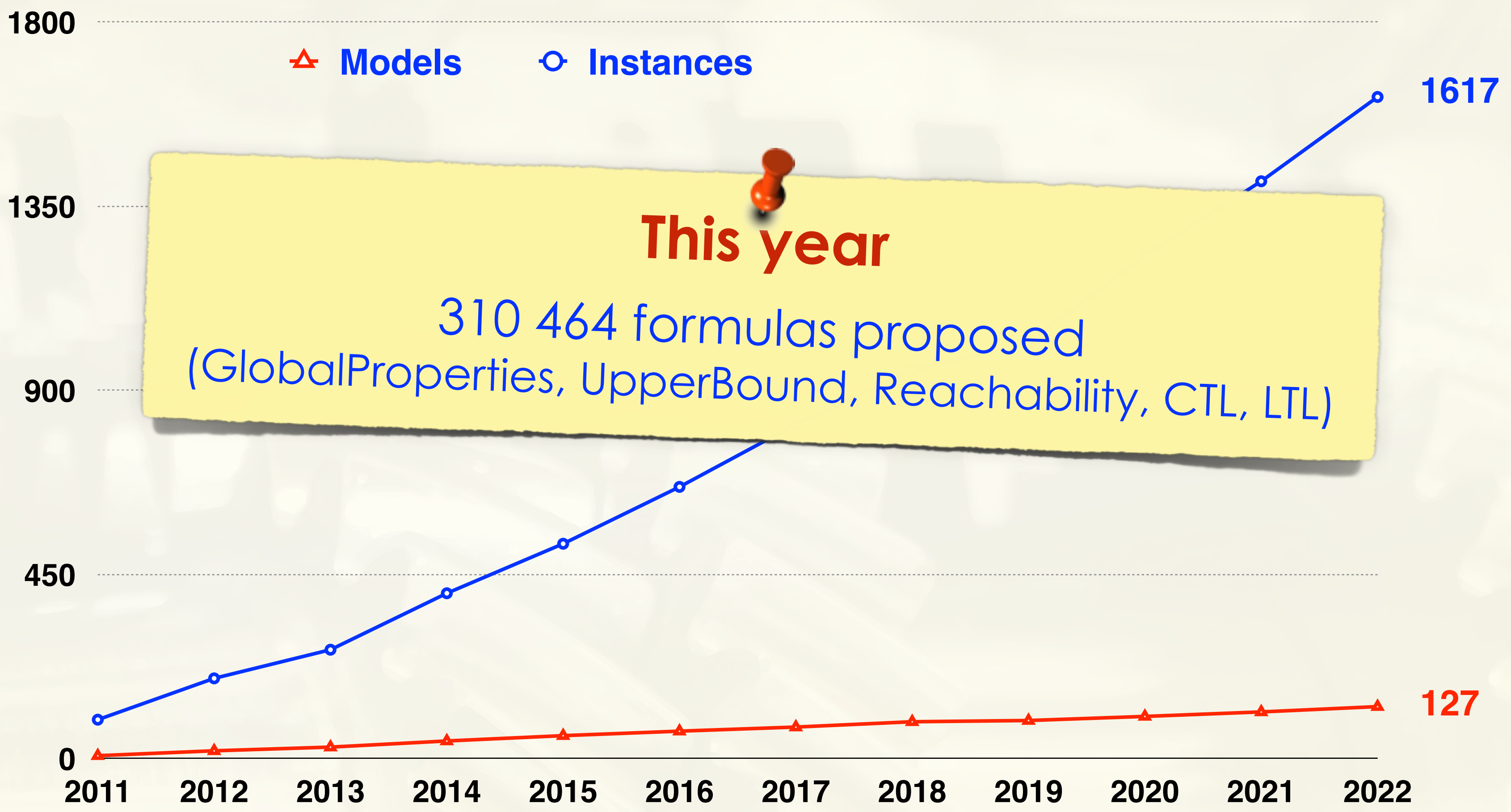
 **Y. Thierry-Mieg** BugTracking, DBSingleClientW, DoubleLock, FunctionPointer, Medical, RingSingleMessageInMbox, SieveSingleMsgMbox, Szymanski **M. Jasper, F. Kordon & H. Garavel** RERS'2022 **R. Mendas & F. kordon** CANInsertWithFailure **L. Muller & H. Garavel** AutonomousCar **L. Di Stefano & H. Garavel** StigmergyCommit, StigmergyElection  
**Thanks!!!**

We really need different models

## Extending the collection of benchmarks every year



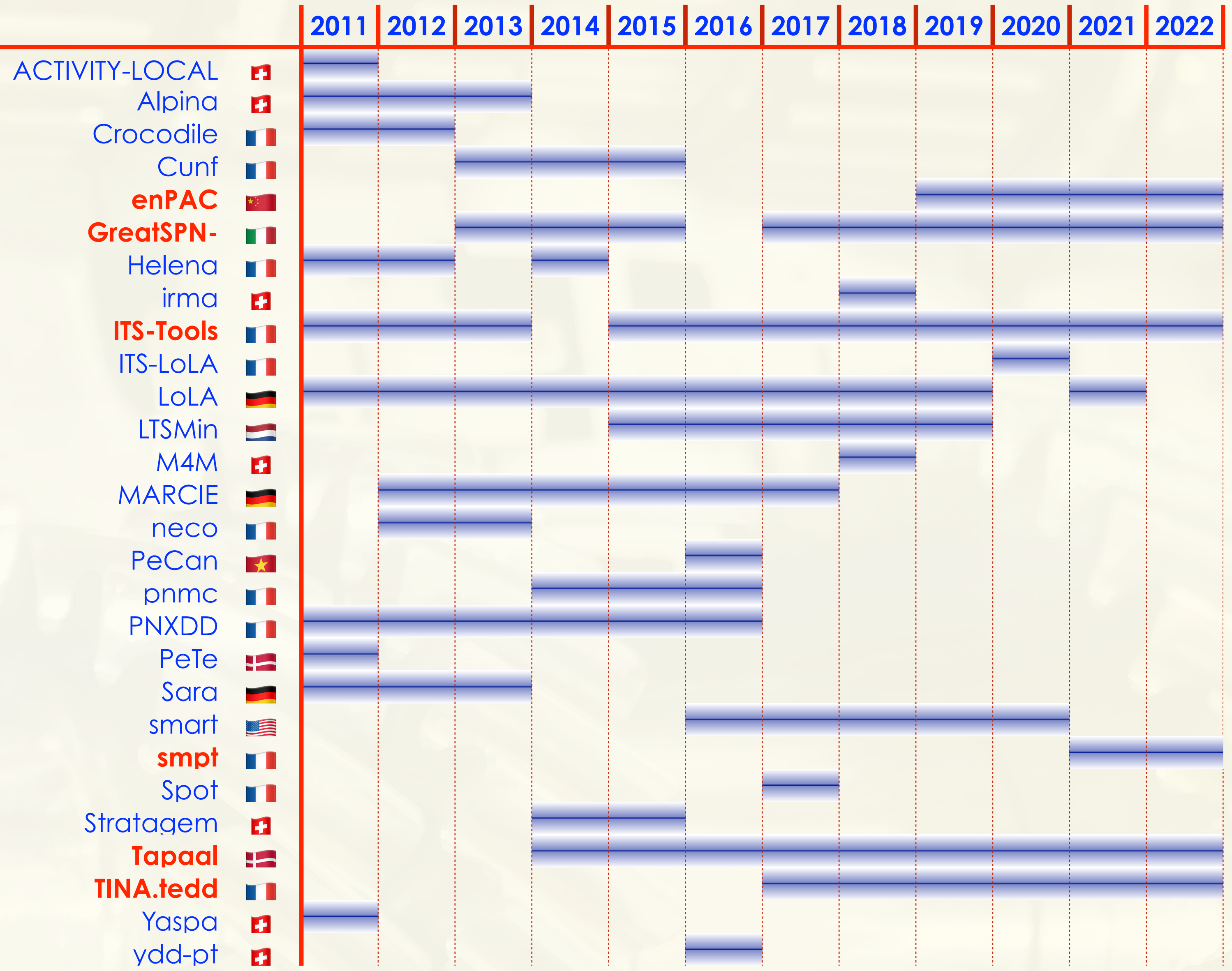
## Extending the collection of benchmarks every year














# Participating Tools

 **26 tools**



 **26 tools**

	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
ACTIVITY-LOCAL 												
Alpina 												
Cocodile 												
M4M 												
MARCIE 												
neco 												
PeCan 												
Yaspa 												
ydd-pt 												

**2022**

6 tools and variants  
 + comparison to 2021 winner  
 + Best Virtual Tool

**All VMs published**  
 For reproducibility

Tools	PN type	parallelism	Techniques
<b>enPAC</b>	P/T	/	ABSTRACTIONS EXPLICIT SEQUENTIAL_PROCESSING STATE_COMPRESSION STRUCTURAL_REDUCTION USE_NUPN
<b>Tina.tedd-s</b>	P/T & Col	/	DECISION_DIAGRAMS EXPLICIT IMPLICIT LATTICE_POINTS_COUNTING LINEAR_EQUATIONS SEQUENTIAL_PROCESSING STRUCTURAL_REDUCTION TOPOLOGICAL UNFOLDING_TO_PT USE_NUPN
<b>GreatSPN-Meddy</b>	P/T & Col	Colat	DECISION_DIAGRAMS PARALLEL_PROCESSING TOPOLOGICAL UNFOLDING_TO_PT USE_NUPN
<b>ITS-Tools</b>	P/T & Col	Colat	BESTFIRST_WALK CL_INSENSITIVE COLLATERAL_PROCESSING CPN_APPROX DEADLOCK_TEST DECISION_DIAGRAMS EXHAUSTIVE_WALK EXPLICIT_INITIAL_STATE INVARIANTS KNOWLEDGE K_INDUCTION LTSMIN MARKED_SUFFIX_TEST OVER_APPROXIMATION PARIKH_WALK PARTIAL_ORDER PROBABILISTIC_WALK QUASILIVENESS_TEST RANDOM_WALK REACHABILITY_KNOWLEDGE SAT_SMT SCC_TEST SIPHON_TEST SKELETON_TEST SL_INSENSITIVE STACK_TEST STRUCTURAL STRUCTURAL_REDUCTION STUTTER_TEST TAJAN COV TOPOLOGICAL TRIVAL UNMARKED_COV_TEST USE_NUPN
<b>smpt</b>	P/T & Col	Colat	BMC COLLATERAL_PROCESSING IMPLICIT K-INDUCTION NET_UNFOLDING PDR-COV PDR-REACH-SATURATED SAT-SMT STATE_EQUATION STRUCTURAL_REDUCTION TOPOLOGICAL LINEAR_EQUATIONS TO_PT USE_NUPN WALK
<b>Tappal</b>	P/T & Col	Colat	AUTOMATON_HEUR AUT_STUB COLLATERAL_PROCESSING CPN_APPROX CTL_CZERO DIST_HEUR EXPLICIT HEURISTIC LOGFIRECOUNT_HEUR(5000) LP_APPROX NDFS OPTIM-1 QUERY_REDUCTION RANDOM_HEUR SAT_SMT SIPHON_TRAP STATE_COMPRESSION STRUCTURAL_REDUCTION STUBBORN STUBBORN_SETS TARJAN TOPOLOGICAL TRACE ABSTRACTION REFINEMENT UNFOLDING_TO_PT WEAK SKIP
<b>TINA.tedd-c</b>	P/T & Col	Colat	COLLATERAL_PROCESSING DECISION_DIAGRAMS EXPLICIT IMPLICIT LATTICE_POINTS_COUNTING LINEAR_EQUATIONS STRUCTURAL_REDUCTION TOPOLOGICAL UNFOLDING_TO_PT USE_NUPN
<b>TINA.tedd-s</b>	P/T & Col	Colat	COLLATERAL_PROCESSING DECISION_DIAGRAMS EXPLICIT IMPLICIT LATTICE_POINTS_COUNTING LINEAR_EQUATIONS STRUCTURAL_REDUCTION TOPOLOGICAL UNFOLDING_TO_PT USE_NUPN

# Examinations

## StateSpace

## GlobalProperties

Liveness, OneSafe, QuasiLiveness, ReachabilityDeadlock, StableMarking

## UpperBound

## Reachability

ReachabilityCardinality → atomic propositions refer to tokens

ReachabilityFireability → atomic propositions refer to firing

## CTL

CTLCardinality → atomic propositions refer to tokens

CTLFireability → atomic propositions refer to firing

## LTL

LTLCardinality → atomic propositions refer to tokens






LTLFireability → atomic propositions refer to firing

# Schedule & Tool confidence

## ~April 15, delivery of disk images

-  Qualification phase (finished mid May)
  - ▶ ~37 000 test runs

## May 14, starting to operate tools

-  **168 168** runs distributed on 4 different machines (France & Germany)
-  VM with 4 cores / 16GB
  - ▶ GreatSPN-meddly, ITS-Tools, Tapa, TINA.tedd-c
-  VM with 1 core / 16 GB
  - ▶ enPAC, TINA.tedd-s
-  Time confinement, 1h (30mn for GlobalPoperties)
-  Only **physical cores** allocated to VMs

## Mid June, feedback sent to developers



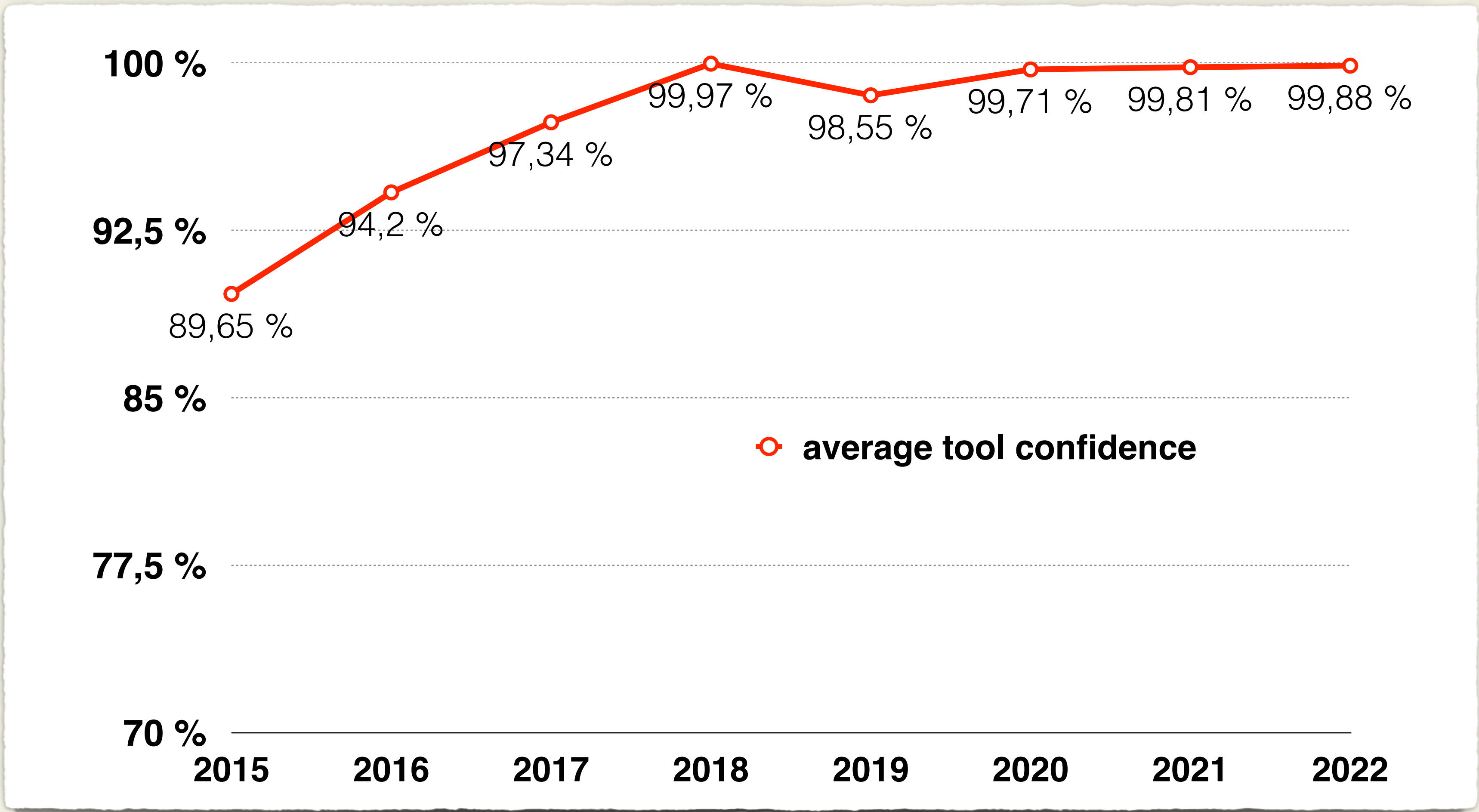
## Consistency checks

- Colored versus equivalent P/T nets

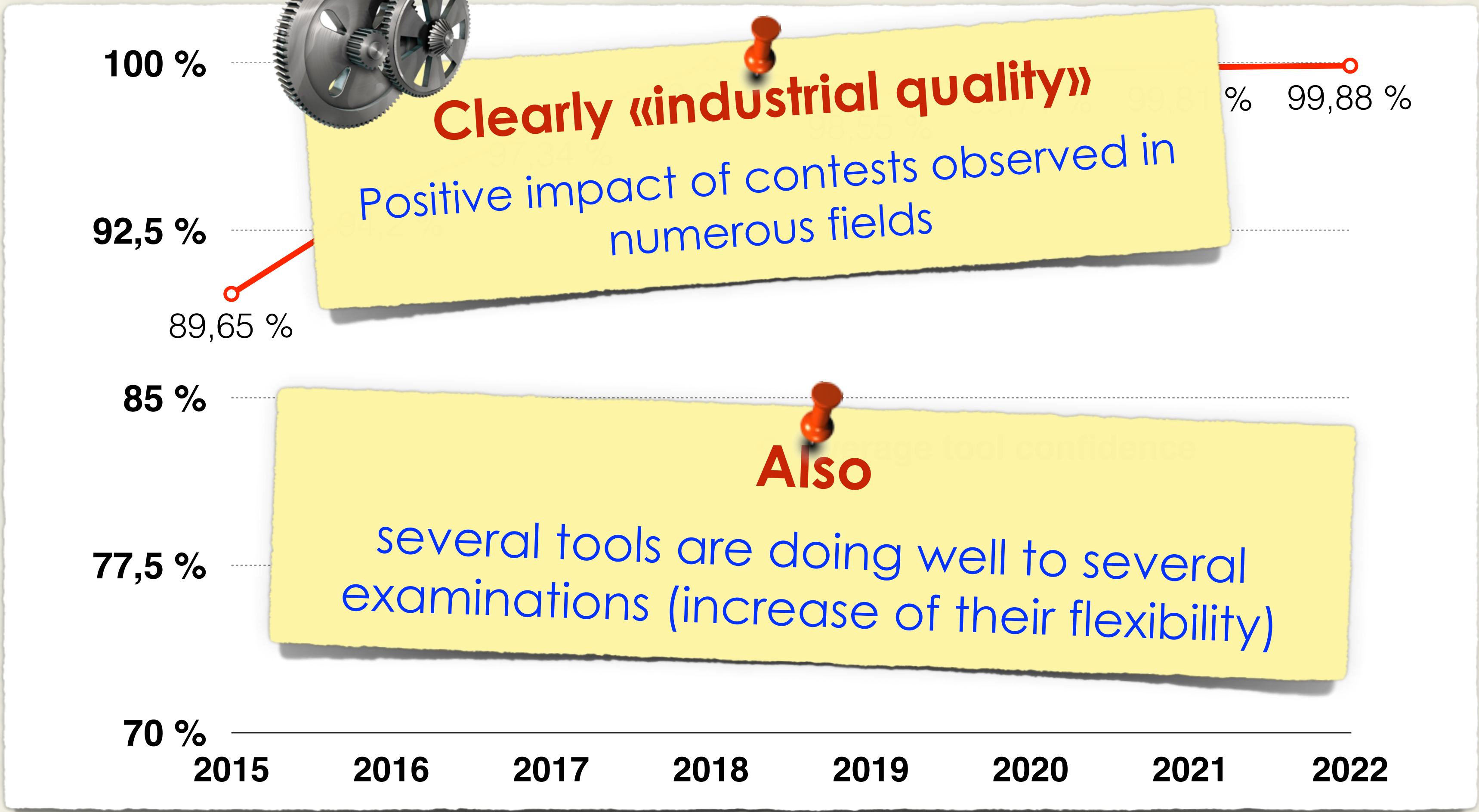
## Computing the “confidence rate”

- Section III.2 in <http://mcc.lip6.fr/rules.php>
- Computing  $V$ , the set of values with a **majority of 3 or more tools**
- For each tool  $t$ , selecting  $V_t$ , the values computed  $\in V$
- For each tool  $t$ , selecting  $V_{\#}$ , the correct values computed  $\in V_t$
- Confidence rate =  $\frac{|V_{\#}|}{|V_t|}$

## Evolution of the confidence since it was introduced







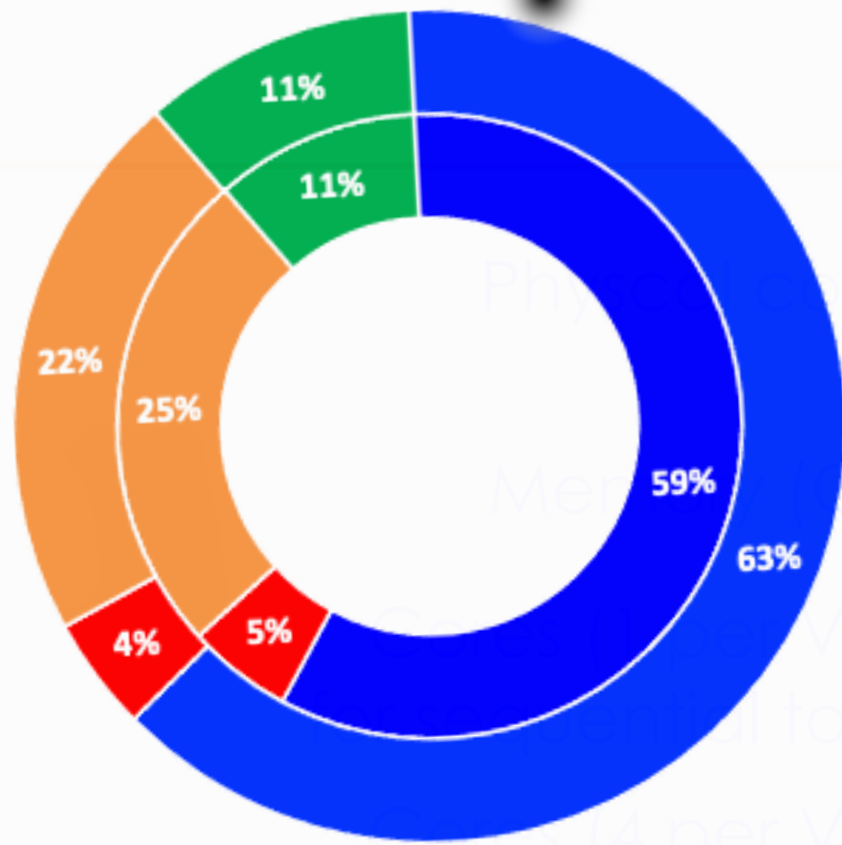
## Evolution of the confidence since it was introduced



Tools	Confidence	success	selected	Examinations
enPAC	99.965 %	42251	42266	2 – LTLCardinality, LTLFireability
tedd-s	99.762 %	3772	3781	1 – StateSpace
GreatSPN	99.569 %	83815	84178	13 – StateSpace, Liveness, OneSafe, QuasiLiveness, ReachabilityDeadlock, StableMarking, UpperBounds, ReachabilityCardinality, ReachabilityFireability, CTLCardinality, CTLFireability, LTLCardinality
ITS-Tools	99.999 %	149757	149758	13 – StateSpace, Liveness, OneSafe, QuasiLiveness, ReachabilityDeadlock, StableMarking, UpperBounds, ReachabilityCardinality, ReachabilityFireability, CTLCardinality, CTLFireability, LTLCardinality
smpt	100.000 %	45585	45585	2 – ReachabilityCardinality, ReachabilityFireability
Tapaal	99.999 %	149868	149870	12 – Liveness, OneSafe, QuasiLiveness, ReachabilityDeadlock, StableMarking, UpperBounds, ReachabilityCardinality, ReachabilityFireability, CTLCardinality, CTLFireability, LTLCardinality
tedd-c	99.762 %	3774	3783	1 – StateSpace
2021-gold	99.998 %	149561	149564	13 – StateSpace, Liveness, OneSafe, QuasiLiveness, ReachabilityDeadlock, StableMarking, UpperBounds, ReachabilityCardinality, ReachabilityFireability, CTLCardinality, CTLFireability, LTLCardinality
BVT-2022	100.000 %	155901	155901	13 – StateSpace, Liveness, OneSafe, QuasiLiveness, ReachabilityDeadlock, StableMarking, UpperBounds, ReachabilityCardinality, ReachabilityFireability, CTLCardinality, CTLFireability, LTLCardinality

# Computing Results in 2019

	 Tall	 Small	 Octoginta-2	 Tajo	Total
Physical cores	13x32 @ 2.4GHz	14x12 @ 2.4GHz	80 @ 2.4GHz	96 @ 2.4GHz	-
Memory (GB)	13x384	14x64	1536	2048	-
Cores (1 per VM) for sequential tools	13x23 13x23 VM in //	14x2 14x2 VM in //	79 79 VM in //	95 95 VM in //	-
Cores (4 per VM) for parallel tools	13x28, 13x7 VM in //	14x3 14x3 VM in //	76 19 VM in //	92, 23 VM in //	-
<b>Number of runs</b>	98592	42640	9152	17784	<b>168168</b>
<b>Total CPU required</b>	1613d, 20h, 56m, 41s	548d, 17h, 38m, 29s	112d, 17h, 14m, 33s	270d, 16h, 25m, 41s	<b>2546d, 0h, 15m, 25s</b>
<b>Total CPU</b>	<b>About 6 years, 11 months and 26 days</b>				-
<b>Time spent to complete benchmarks</b>	<b>about 19 days</b>				-
<b>VM boot time of VMs + management (overhead)</b>	<b>About 10d (Included in total CPU)</b>				-

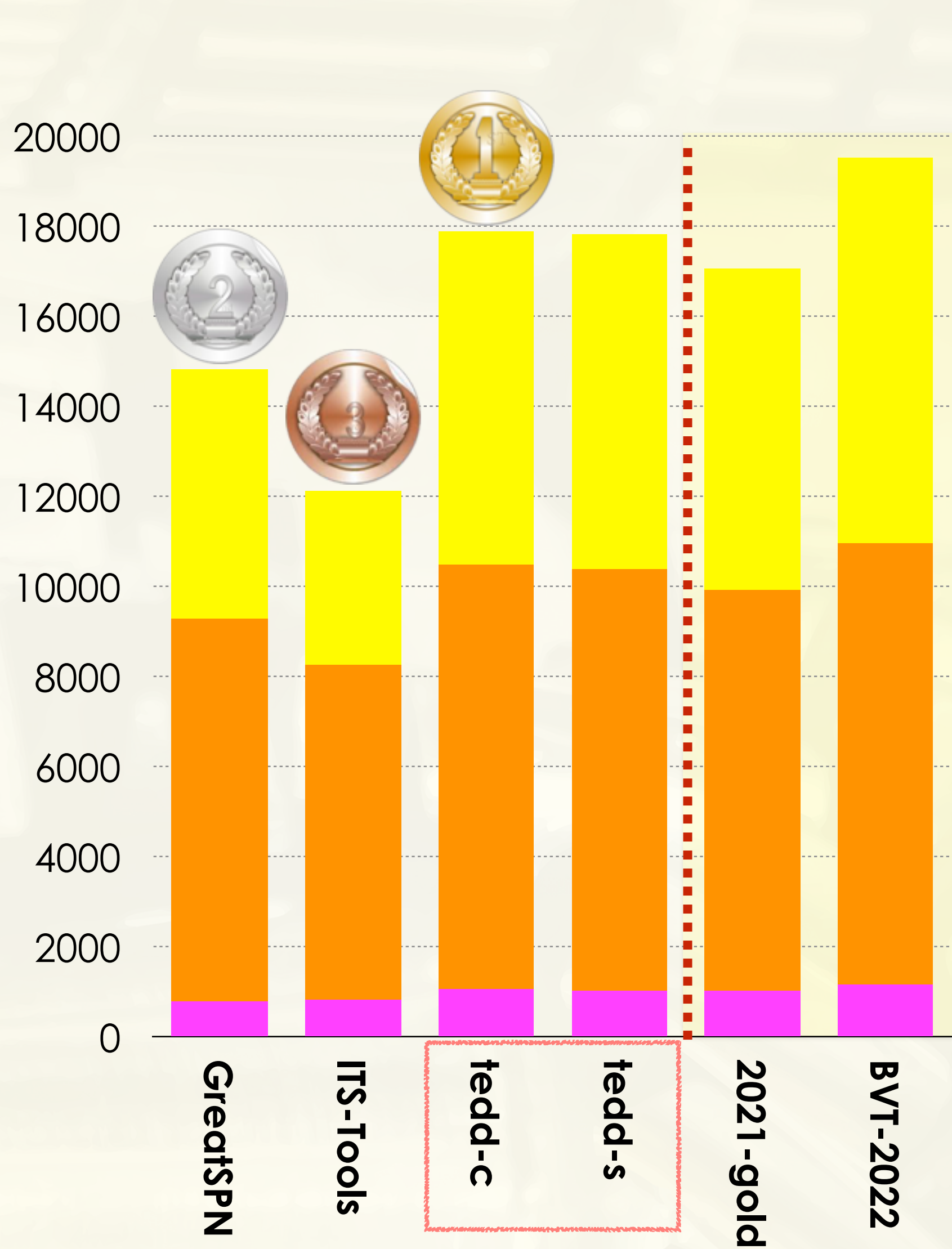


- tall
- octoginta2
- small
- tajo

**Thank you very much**  
 Rostock University  
 Université Nanterre  
 Sorbonne Université

		Small	Octoginta-2	Tajo	Total
<b>Number of runs</b>	98592	42640	9152	17784	<b>168168</b>
<b>Total CPU required</b>	1613d, 20h, 56m, 41s	548d, 17h, 38m, 29s	112d, 17h, 14m, 33s	270d, 16h, 25m, 41s	<b>2546d, 0h, 15m, 25s</b>
<b>Total CPU</b>	<b>About 6 years, 11 months and 26 days</b>				-
<b>Time spent to complete benchmarks</b>	<b>about 19 days</b>				-
<b>VM boot time of VMs + management (overhead)</b>	<b>About 10d (Included in total CPU)</b>				-

# Results



- Surprise&P/T
- Surprise&Colored
- Known&P/T
- Known&Colored



**GreatSPN**  
480 times



**GreatSPN**  
591 times

StateSpace	confidence	success	selected
<b>GreatSPN</b>	<b>99.918 %</b>	3648	3651
<b>ITS-Tools</b>	<b>100.000 %</b>	2691	2691
<b>tedd-c</b>	<b>99.762 %</b>	3774	3783
<b>tedd-s</b>	<b>99.762 %</b>	3772	3781
<b>2021-gold</b>	<b>100.000 %</b>	3669	3669
<b>BVT-2022</b>	<b>100.000 %</b>	3774	3774



## Be careful with +Inf\*\*\*\*\*

There are a few infinite models but...

Philosophers — Colored (87.31 pts max)						
	GreatSPN	ITS-Tools	tedd-c	tedd-s	2021-gold	BVT-2022
Score	60.48	58.80	73.91	73.91	73.91	73.91
fastest	5	1	6	0	0	0
smallest memory	6	4	1	0	0	1
	243 945 1 10	243 ? 1 10	243 945 1 10	243 945 1 10	243 945 1 10	243 945 1 10
000005	TTTT	T-TT	TTTT	TTTT	TTTT	TTTT
	----:M / 6.72 →	----: / 5.88 →	----:P- / 6.72 →	----: / 6.72 →	----: / 6.72 →	----: / 6.72 →
	59049 459270 1	3120	2270 1	70 1	459270 1	59049 459270 1
			----:P- / 6.72 →	----: / 6.72 →	----: / 6.72 →	----:M / 6.72 →
050000	DNF	DNF	+Inf***** ? 1	CC	CC	[undef]
	0 →	0 →	100000	0 →	0 →	? →
			????			
			----:PM / 0.00 →			----: / 0.00 →

## Be careful with +Inf\*\*\*\*\*

There are a few infinite models but...

Philosophers — Colored (87.31 pts max)						
	GreatSPN	ITS-Tools	tedd-c	tedd-s	2021-gold	BVT-2022
Score	60.48	58.80	73.91	73.91	73.91	73.91
fastest	5	1	6	0	0	0
smallest memory	6	4	1	0	0	1
000005	243 945 1 10 TTTT ----:M / 6.72 →	243 ? 1 10 T-TT ----: / 5.88 →	243 945 1 10 TTTT ----:P- / 6.72 →	243 945 1 10 TTTT ----: / 6.72 →	243 945 1 10 TTTT ----: / 6.72 →	243 945 1 10 TTTT ----: / 6.72 →
050000	59049 459270 1 DNF 0 →	31 20 DNF 0 →	2270 1 +Inf***** ? 1 100000 ???? ----:PM / 0.00 →	70 1 CC 0 →	459270 1 CC 0 →	59049 459270 1 [undef] ? ----: / 0.00 →

```

STATE_SPACE
11554096304905898911445960292007068137449769006120602694155471294340846013653585907604875397122389218463011266954858098278273990593288037371713876892
TECHNIQUES COLLATERAL_PROCESSING DECISION
STRUCTURAL_REDUCTION TOPOLOGICAL_USE_NUPN
STATE_SPACE_TRANSITIONS -1 TECHNIQUES_COLLATERAL_PROCESSING
UNFOLDING_TO_PT STRUCTURAL_REDUCTION_TOPOLOGICAL_USE_NUPN
STATE_SPACE_MAX_TOKEN_IN_PLACE 1 TECHNIQUES_COLLATERAL_PROCESSING
UNFOLDING_TO_PT STRUCTURAL_REDUCTION_TOPOLOGICAL_USE_NUPN
STATE_SPACE_MAX_TOKEN_PER_MARKING 100000
LINEAR_EQUATIONS UNFOLDING_TO_PT STRUCTURAL_REDUCTION_TOPOLOGICAL_USE_NUPN
IMPLICIT LINEAR_EQUATIONS UNFOLDING_TO_PT
STATES
BK_STOP 1654907675924
    
```

**Limit of my Ada analysis programs**  
 Size of Long\_Long\_Float  
 Max = 1.0E+4932

## There is something wrong with RERS2020

2021-gold is tedd  
they are all wrong

2 1 1 5 (expected value)  
2.00 2.00 3.00 3.00 (confidence rate)

RERS2020 — T (737.70)						
	GreatSPN	ITS-Tools	tedd-c	tedd-s	2021-gold	BVT-2022
Score	0.00	-737.63	737.63	737.63	737.63	737.63
fastest	0	0	9	0	0	0
best memory	0	0	6	2	0	1
pb101	DNF	345421 ? 15 X-TT	2 1 1 5 TTTT	2 1 1 5 TTTT	2 1 1 5 TTTT	2 1 1 5 TTTT
	0	V---: / -81.95	---:PM / 81.95	---: / 81.95	---: / 81.95	---: / 81.95
pb102	DNF	3.8037E+0008 ? 17 X-TT	2 1 1 7 TTTT	2 1 1 7 TTTT	2 1 1 7 TTTT	2 1 1 7 TTTT
	0	V---: / -81.95	---:P- / 81.95	---:M / 81.95	---: / 81.95	---: / 81.95
pb103	DNF	6.4658E+0007 ? 18 X-TT	2 1 1 8 TTTT	2 1 1 8 TTTT	2 1 1 8 TTTT	2 1 1 8 TTTT
	0	V---: / -81.95	---:P- / 81.95	---: / 81.95	---: / 81.95	---:M / 81.95
pb104	DNF	2.4269E+0010 ? 19 X-TT	2 1 1 9 TTTT	2 1 1 9 TTTT	2 1 1 9 TTTT	2 1 1 9 TTTT
	0	V---: / -81.95	---:PM / 81.95	---: / 81.95	---: / 81.95	---: / 81.95
pb105	DNF	8.0169E+0012 ? 111 X-TT	2 1 1 11 TTTT	2 1 1 11 TTTT	2 1 1 11 TTTT	2 1 1 11 TTTT
	0	V---: / -81.95	---:PM / 81.95	---: / 81.95	---: / 81.95	---: / 81.95
pb106	DNF	1.0137E+0013 ? 112 X-TT	2 1 1 12 TTTT	2 1 1 12 TTTT	2 1 1 12 TTTT	2 1 1 12 TTTT
	0	V---: / -81.95	---:PM / 81.95	---: / 81.95	---: / 81.95	---: / 81.95
pb107	DNF	2.4371E+0013 ? 113 X-TT	2 1 1 13 TTTT	2 1 1 13 TTTT	2 1 1 13 TTTT	2 1 1 13 TTTT
	0	V---: / -81.95	---:PM / 81.95	---: / 81.95	---: / 81.95	---: / 81.95
pb108	DNF	2.5304E+0016 ? 115 X-TT	2 1 1 15 TTTT	2 1 1 15 TTTT	2 1 1 15 TTTT	2 1 1 15 TTTT
	0	V---: / -81.95	---:PM / 81.95	---: / 81.95	---: / 81.95	---: / 81.95
pb109	DNF	6.3468E+0016 ? 116 X-TT	2 1 1 16 TTTT	2 1 1 16 TTTT	2 1 1 16 TTTT	2 1 1 16 TTTT
	0	V---: / -81.95	---:P- / 81.95	---:M / 81.95	---: / 81.95	---: / 81.95

## There is something wrong with RERS2020

2021-gold is tedd  
they are all wrong

2 1 1 5 (expected value)  
2.00 2.00 3.00 3.00 (confidence rate)

The model form  
contradicts (CAESAR-BDD)

Size of the marking graphs

Parameter	Number of reachable markings
problem 101	345421 <sup>(u)</sup>
problem 102	≥ 1.14736e+07 <sup>(w)</sup>
problem 103	6.46576e+07 <sup>(z)</sup>
problem 104	≥ 3.92045e+06 <sup>(ab)</sup>
problem 105	≥ 2.96522e+06 <sup>(ae)</sup>
problem 106	≥ 3.75815e+08 <sup>(ah)</sup>
problem 107	≥ 6.59343e+07 <sup>(ak)</sup>
problem 108	≥ 2.61085e+08 <sup>(an)</sup>
problem 109	≥ 4.82768e+06 <sup>(aq)</sup>

RERS2020 — T (737.70) (5 min)

	GreatSPN	ITS-Tools	tedd-c	tedd-s	2021-gold	BVT-2022
Score	0.00	-737.63	737.63	737.63	737.63	737.63
fastest	0	0	9	0	0	0
best memory	0	0	6	2	0	1
pb101	DNF	345421 ? 15 X-TT	2 1 1 5 TTTT	2 1 1 5 TTTT	2 1 1 5 TTTT	2 1 1 5 TTTT
	0 →	V---: / -81.95 →	---:PM / 81.95 →	---: / 81.95 →	---: / 81.95 →	---: / 81.95 →
pb102	DNF	3.8037E+0008 ? 17 X-TT	2 1 1 7 TTTT	2 1 1 7 TTTT	2 1 1 7 TTTT	2 1 1 7 TTTT
	0 →	V---: / -81.95 →	---:P- / 81.95 →	---:M / 81.95 →	---: / 81.95 →	---: / 81.95 →
pb103	DNF	6.4658E+0007 ? 18 X-TT	2 1 1 8 TTTT	2 1 1 8 TTTT	2 1 1 8 TTTT	2 1 1 8 TTTT
	0 →	V---: / -81.95 →	---:P- / 81.95 →	---: / 81.95 →	---: / 81.95 →	---:M / 81.95 →
pb104	DNF	2.4269E+0010 ? 19 X-TT	2 1 1 9 TTTT	2 1 1 9 TTTT	2 1 1 9 TTTT	2 1 1 9 TTTT
	0 →	V---: / -81.95 →	---:PM / 81.95 →	---: / 81.95 →	---: / 81.95 →	---: / 81.95 →
pb105	DNF	8.0169E+0012 ? 111 X-TT	2 1 1 11 TTTT	2 1 1 11 TTTT	2 1 1 11 TTTT	2 1 1 11 TTTT
	0 →	V---: / -81.95 →	---:PM / 81.95 →	---: / 81.95 →	---: / 81.95 →	---: / 81.95 →
pb106	DNF	1.0137E+0013 ? 112 X-TT	2 1 1 12 TTTT	2 1 1 12 TTTT	2 1 1 12 TTTT	2 1 1 12 TTTT
	0 →	V---: / -81.95 →	---:PM / 81.95 →	---: / 81.95 →	---: / 81.95 →	---: / 81.95 →
pb107	DNF	2.4371E+0013 ? 113 X-TT	2 1 1 13 TTTT	2 1 1 13 TTTT	2 1 1 13 TTTT	2 1 1 13 TTTT
	0 →	V---: / -81.95 →	---:PM / 81.95 →	---: / 81.95 →	---: / 81.95 →	---: / 81.95 →
pb108	DNF	2.5304E+0016 ? 115 X-TT	2 1 1 15 TTTT	2 1 1 15 TTTT	2 1 1 15 TTTT	2 1 1 15 TTTT
	0 →	V---: / -81.95 →	---:PM / 81.95 →	---: / 81.95 →	---: / 81.95 →	---: / 81.95 →
pb109	DNF	6.3468E+0016 ? 116 X-TT	2 1 1 16 TTTT	2 1 1 16 TTTT	2 1 1 16 TTTT	2 1 1 16 TTTT
	0 →	V---: / -81.95 →	---:P- / 81.95 →	---:M / 81.95 →	---: / 81.95 →	---: / 81.95 →

## There is something wrong with RERS2020

2021-gold is tedd  
they are all wrong

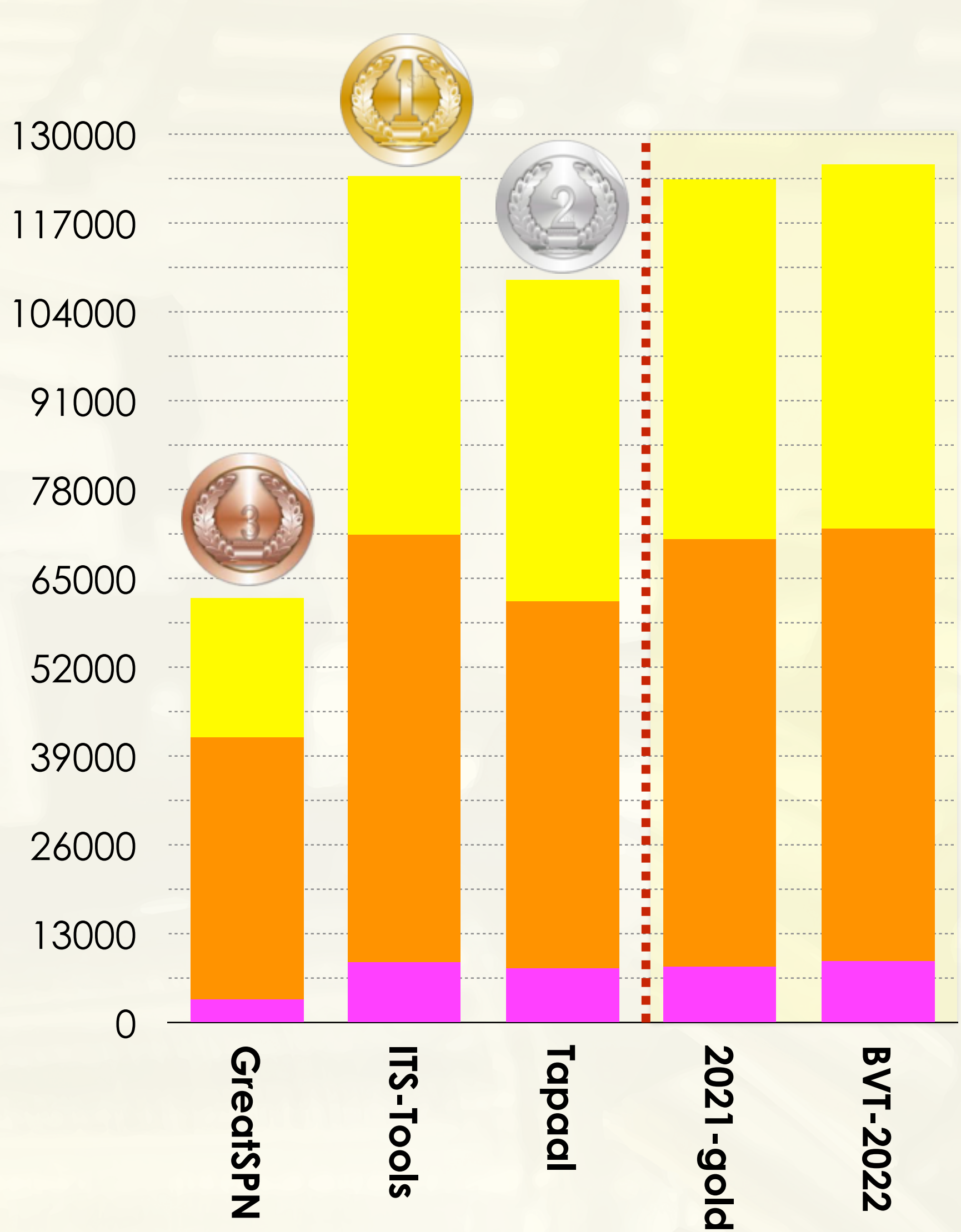


It's a limit of the qualifier!

Very difficult to handle  
No impact in the ranking

Confidence rate/core updated (manually) on the slides  
probably some 10 false negatives among 1369599 computed values < 0,0002% ≈ 2 out of 10<sup>4</sup> runs

	GreatSPN	ITS-Tools	tedd-c	tedd-s	2021-gold	BVT-2022
Score	0.00	-737.63	737.63	737.63	737.63	737.63
fastest	0	0	9	0	0	0
						1
15						2 1 1 5
						TTTT
1.95						----- / 81.95
17						2 1 1 7
						TTTT
1.95						----- / 81.95
18						2 1 1 8
						TTTT
1.95						-----M / 81.95
19						2 1 1 9
						TTTT
1.95						----- / 81.95
11						2 1 1 11
						TTTT
1.95						----- / 81.95
pb106	DNF	2.4371E+0013 ?	2 1 1 12	2 1 1 12	2 1 1 12	2 1 1 12
	0	1 12	TTTT	TTTT	TTTT	TTTT
		X-TT	-----PM / 81.95	----- / 81.95	----- / 81.95	----- / 81.95
		V----- / -81.95				
pb107	DNF	2.4371E+0013 ?	2 1 1 13	2 1 1 13	2 1 1 13	2 1 1 13
	0	1 13	TTTT	TTTT	TTTT	TTTT
		X-TT	-----PM / 81.95	----- / 81.95	----- / 81.95	----- / 81.95
		V----- / -81.95				
pb108	DNF	2.5304E+0016 ?	2 1 1 15	2 1 1 15	2 1 1 15	2 1 1 15
	0	1 15	TTTT	TTTT	TTTT	TTTT
		X-TT	-----PM / 81.95	----- / 81.95	----- / 81.95	----- / 81.95
		V----- / -81.95				
pb109	DNF	6.3468E+0016 ?	2 1 1 16	2 1 1 16	2 1 1 16	2 1 1 16
	0	1 16	TTTT	TTTT	TTTT	TTTT
		X-TT	-----P- / 81.95	-----M / 81.95	----- / 81.95	----- / 81.95
		V----- / -81.95				



- Surprise&P/T
- Surprise&Colored
- Known&P/T
- Known&Colored

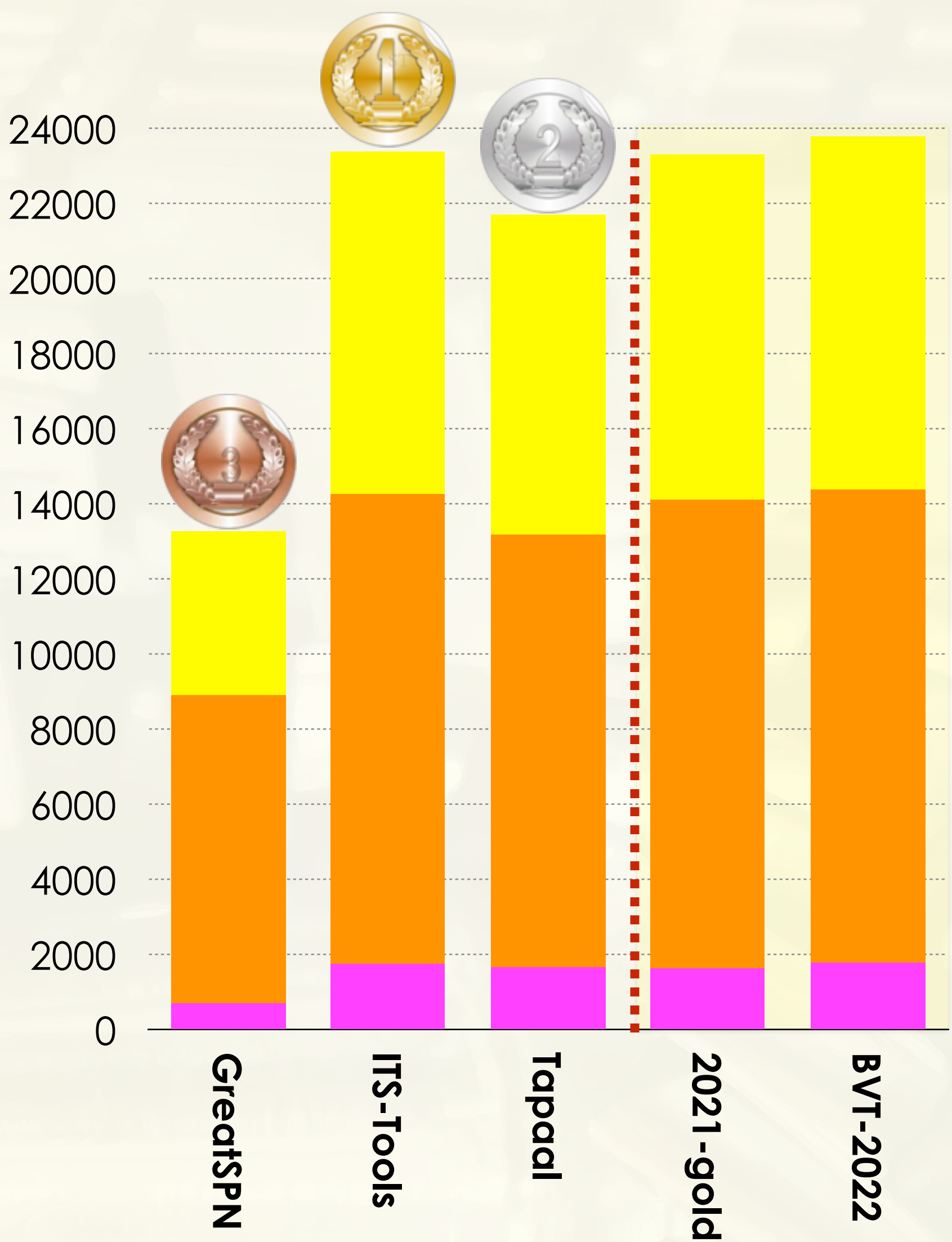


**Tapaal**  
4262 times



**Tapaal**  
5609 times

UpperBound	confidence	success	selected
<b>GreatSPN</b>	100.000 %	4160	4160
<b>ITS-Tools</b>	100.000 %	7004	7004
<b>Tapaal</b>	100.000 %	6624	6624
<b>2021-gold</b>	100.000 %	6937	6937
<b>BVT-2022</b>	100.000 %	7007	7007



- Surprise&P/T
- Surprise&Colored
- Known&P/T
- Known&Colored

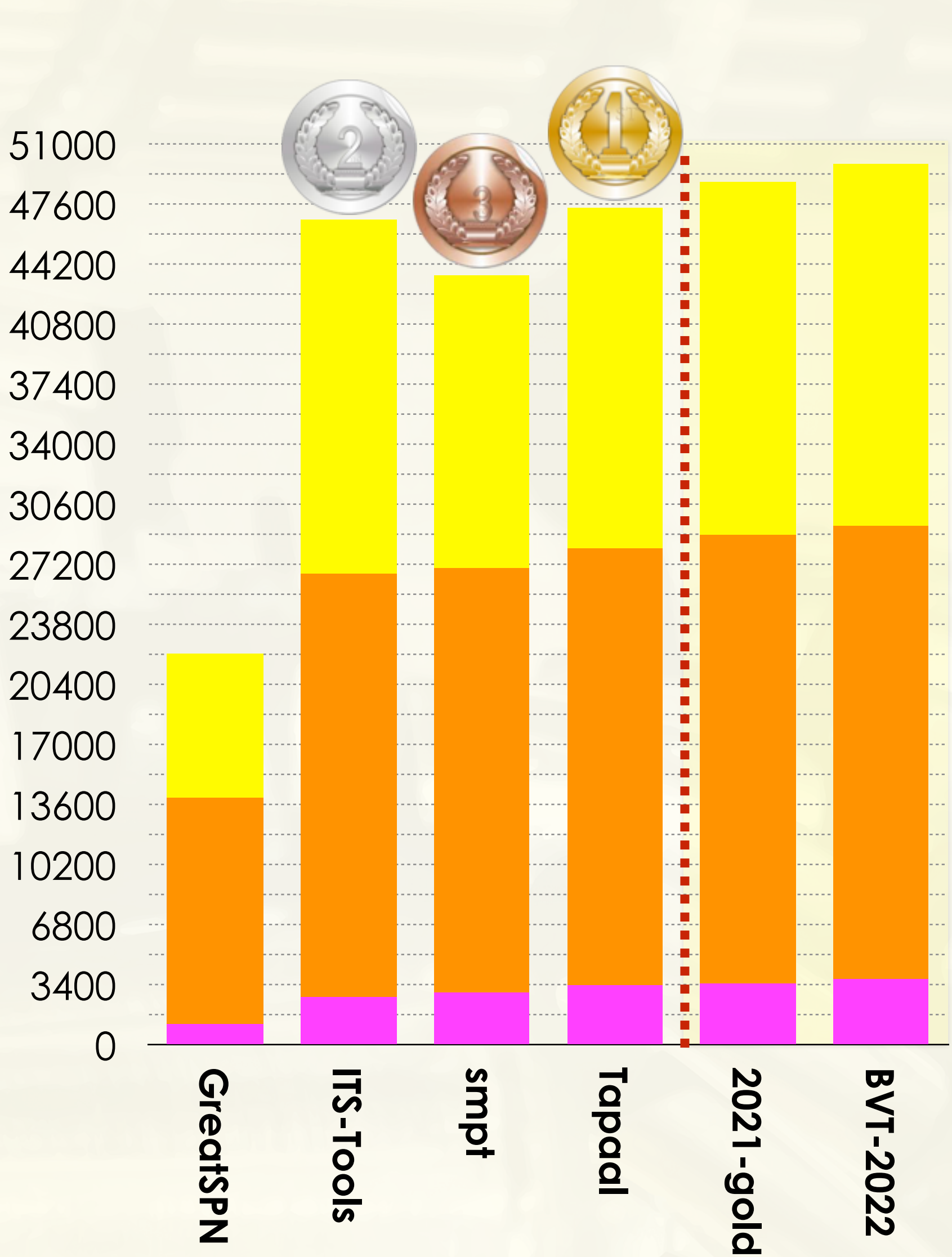


**ITS-Tools**  
562 times



**Tapaal**  
923 times

UpperBound	confidence	success	selected
<b>GreatSPN</b>	100.000 %	14343	14343
<b>ITS-Tools</b>	100.000 %	23203	23203
<b>Tapaal</b>	100.000 %	22317	22317
<b>2021-gold</b>	100.000 %	22958	22958
<b>BVT-2022</b>	100.000 %	23206	23206



- Surprise&P/T
- Surprise&Colored
- Known&P/T
- Known&Colored



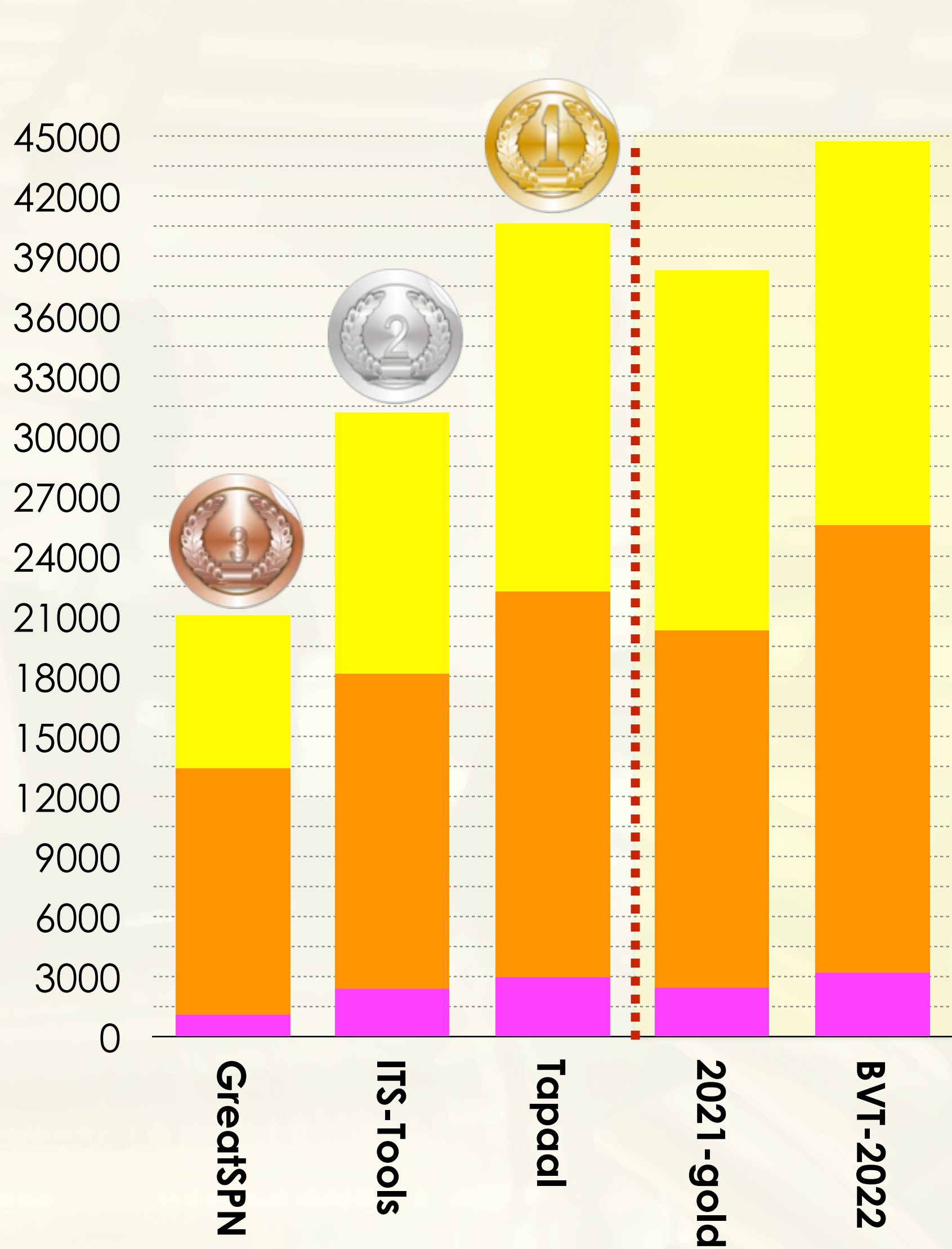
**Tapaal**  
1143 times



**Tapaal**  
1950 times

Reachability	confidence	success	selected
<b>GreatSPN</b>	<b>99.996 %</b>	22531	22532
<b>ITS-Tools</b>	<b>100.000 %</b>	45269	45269
<b>smpt</b>	<b>100.000 %</b>	45585	45585
<b>Tapaal</b>	<b>100.000 %</b>	47419	47419
<b>2021-gold</b>	<b>100.000 %</b>	47981	47981
<b>BVT-2022</b>	<b>100.000 %</b>	48295	48295





- Surprise&P/T
- Surprise&Colored
- Known&P/T
- Known&Colored

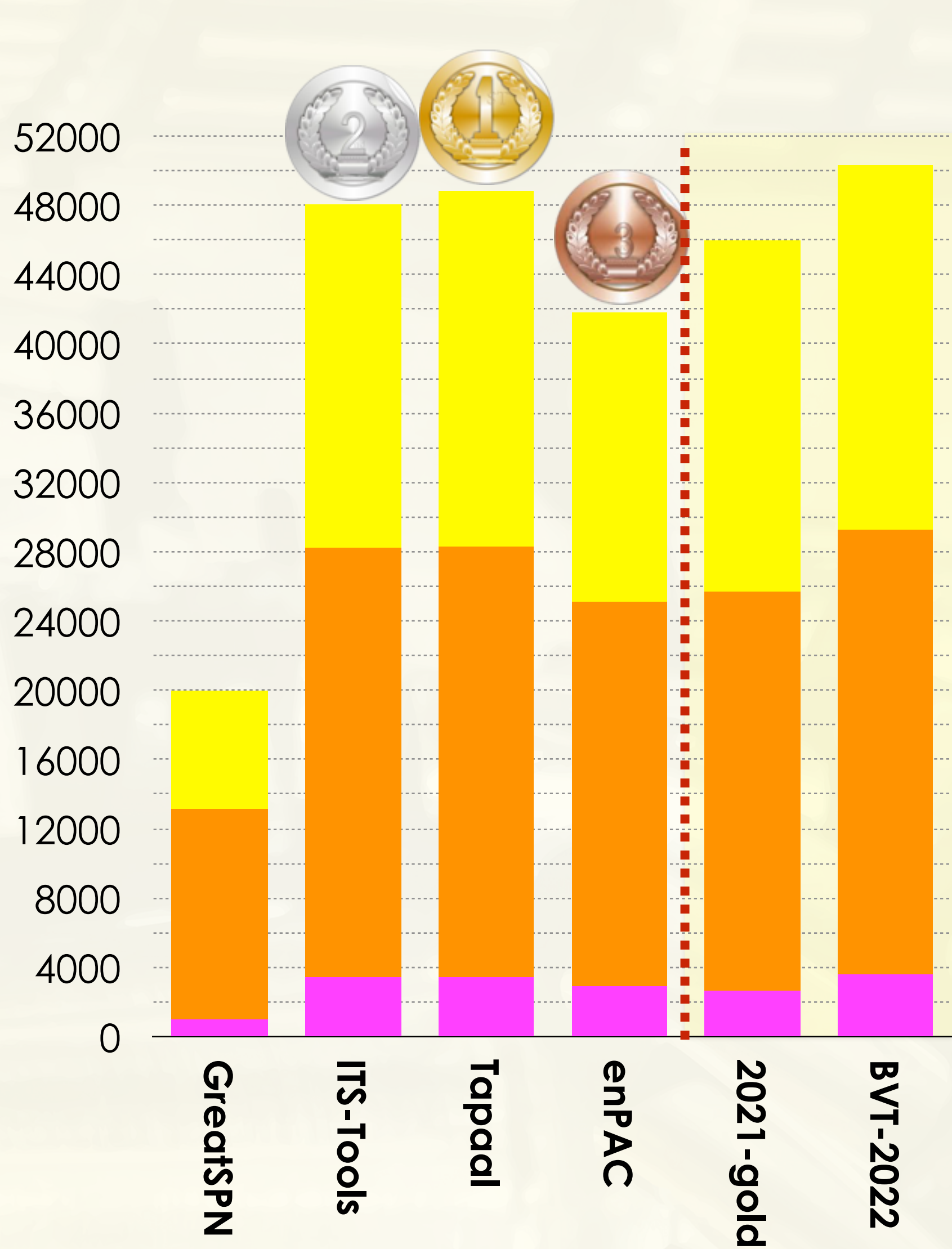


**GreatSPN**  
777 times



**GreatSPN**  
641 times

CTL	confidence	success	selected
<b>GreatSPN</b>	<b>100.000 %</b>	17856	17856
<b>ITS-Tools</b>	<b>99.996 %</b>	25297	25298
<b>Tapaal</b>	<b>99.992 %</b>	26491	26493
<b>2021-gold</b>	<b>99.996 %</b>	25157	25158
<b>BVT-2022</b>	<b>100.000 %</b>	26546	26546



- Surprise&P/T
- Surprise&Colored
- Known&P/T
- Known&Colored



**Tapaal**  
920 times



**Tapaal**  
1147 times

LTL	confidence	success	selected
<b>GreatSPN</b>	<b>98.341 %</b>	21277	21636
<b>ITS-Tools</b>	<b>100.000 %</b>	46293	46293
<b>Tapaal</b>	<b>100.000 %</b>	47017	47017
<b>enPAC</b>	<b>99.965 %</b>	42251	42266
<b>2021-gold</b>	<b>99.995 %</b>	42859	42861
<b>BVT-2022</b>	<b>100.000 %</b>	47073	47073

# Conclusive remarks

- Many generated web pages (about 112K in 2022)
- Numerous generated charts (96814 in 2022)

The screenshot shows a browser window with the URL 'localhost' and a search bar. The main content is a report titled 'MCC'2022- Results for LTLFireability'. It features several tables:

Summary of Results for LTLFireability						
	GreatSPN	ITS-Tools	Tapaal	enPAC	2021-gold	BVT-2022
Total Points	10062.54	23603.87	24046.13	20824.44	22426.34	24951.84

All «Surprise» models						
	GreatSPN	ITS-Tools	Tapaal	enPAC	2021-gold	BVT-2022
Total Points	3507.05	9654.18	10069.44	8233.62	9918.41	10402.02
fastest	27	42	30	6	26	12
less memory	15	34	40	1	28	25

AutonomousCar — P/T (906.00 pts max)						
	GreatSPN	ITS-Tools	Tapaal	enPAC	2021-gold	BVT-2022
Score	869.22	897.54	886.26	795.41	874.91	906.00
fastest	9	1	5	2	1	2
smallest memory	6	1	6	0	3	4

The bottom half of the screenshot shows a detailed performance comparison for various models (01a, 01b, 02a, 02b, 03a, 03b, 04a, 04b, 05a, 05b, 06a, 06b, 07a) across the same teams. Each model's performance is visualized using progress bars and numerical values for different metrics.

- Many generated web pages (about 112K in 2022)
- Numerous generated charts (96814 in 2022)

The screenshot displays a web browser window with a detailed HTML report for the Model Checking Contest 2022. The report is titled "Model Checking Contest 2022" and is the 12th edition, held in Bergen, Norway, on June 21, 2022. The specific report is for the execution of r294-sml-16546387290412, last updated on Jan 13, 2022.

The report includes a navigation menu with options like Home, Calls, Models, For Developers, Publications, Committees, Rules, and Pr. The main content area is titled "About the Execution of ITS-Tools for AutonomousCar-PT-01a".

Below the title, there is an "Execution Summary" table:

Max Memory Used (MB)	Time wait (ms)	CPU Usage (ms)	I/O Wait (ms)	Computed Result	Execution Status
259.264	18938.00	34706.00	397.10	TFFFFFFFFTFTFFFF	normal

Below the summary table, there is an "Execution Chart" section. The chart is titled "Resources Consumption for ITS-Tools LTLFireability on AutonomousCar-PT-01a". The chart shows the percentage of resources consumed over time (0 to 20 seconds). The Y-axis represents the percentage of resources consumed (0% to 100%). The X-axis represents Execution time (seconds). The legend indicates that the chart tracks Memory (red), I/O wait (orange), Core 1 (blue), Core 2 (purple), Core 3 (pink), and Core 4 (green). The chart shows that Core 4 (green) and Core 3 (pink) reach the highest consumption levels, peaking at approximately 80% around 5 seconds and 15 seconds respectively. Core 1 (blue) and Core 2 (purple) also show significant consumption, peaking around 60%.

At the bottom of the screenshot, there is a "Trace from the execution" section, which is partially visible and shows a list of test cases (01a, 01b, 02a, 02b, 03a, 03b, 04a, 04b, 05a, 05b, 06a, 06b, 07a) with their respective scores and execution times.

Many generated web pages (about 112K in 2022)

Numerous generated charts (96814 in 2022)

not completed).

### ITS-Tools versus GreatSPN

Some statistics are displayed below, based on **3234** runs (1617 for ITS-Tools and 1617 for GreatSPN, so there are 1617 plots on each of the two charts). Each execution was allowed **1 hour** and **16 GByte** of memory. Then performance charts comparing ITS-Tools to GreatSPN are shown (you may click on one graph to enlarge it).

Statistics on the executions			
	ITS-Tools	GreatSPN	Both tools
All computed OK	703	0	888
ITS-Tools = GreatSPN	—	—	0
ITS-Tools > GreatSPN	—	—	3
ITS-Tools < GreatSPN	—	—	11
Do not compete	0	0	0
Error detected	0	0	0
Cannot Compute + Time-out	0	703	12

	ITS-Tools	GreatSPN
Times tool wins	1047	558
Times tool wins	1108	497

On the chart below, ● denote cases where the two tools did computed all results without error, ○ denote cases where the two tool did computed the same number of values (but not al values in the examination), ▲ denote cases where ITS-Tools computed more values than GreatSPN, ▼ denote cases where ITS-Tools computed less values than GreatSPN, ◆ denote the cases where at least one tool did not compete, ◻ denote the cases where at least one tool computed a bad value and ◯ denote the cases where at least one tool stated it could not compute a result or timed-out.

ITS-Tools wins when points are below the diagonal, GreatSPN wins when points are above the diagonal.

Many generated web pages (about 112K in 2022)

Numerous generated charts (96814 in 2022)

**Summary of Results**

	GreatSPN	ITS-Tools
Total Points	10062.54	23603.87
Total Points	3507.05	9654.18
fastest	27	42
less memory	15	34

**AutonomousCar**

	GreatSPN	ITS-Tools
Score	869.22	897.54
fastest	9	1
smallest memory	6	1

**Statistics on the executions**

	ITS-Tools	GreatSPN	Both tools
All computed OK	703	0	888
ITS-Tools = GreatSPN	—	—	0
ITS-Tools > GreatSPN	—	—	3
ITS-Tools < GreatSPN	—	—	11
Do not compete	0	0	0
Error detected	0	0	0
Cannot Compute + Time-out	0	703	12

**Times tool wins**

	ITS-Tools	GreatSPN
Smallest Memory Footprint	1047	558
Shortest Execution Time	1108	497

**Execution Summary**

	Max Memory Used (MB)	Time wait	CPU Usage
005	10240	0.00	0.00
010	14336	0.00	0.00
020	10240	0.00	0.00
030	8192	0.00	0.00
040	4096	0.00	0.00
050	2048	0.00	0.00
060	1024	0.00	0.00
070	512	0.00	0.00
080	256	0.00	0.00
090	128	0.00	0.00
100	64	0.00	0.00

**Execution time, ITS-Tools versus GreatSPN for UpperBounds (All models)**

Legend: GreatSPN (blue line), ITS-Tools (red line), LoLA (green line), smpt (purple line), Tapaal (grey line), 2020-gold (pink line), BVT-2021 (brown line).

- Many generated web pages (about 112K in 2022)
- Numerous generated charts (96814 in 2022)

**Summary of Results**

	GreatSPN	ITS-Tools
Total Points	10062.54	23603.87
Total Points	3507.05	9654.18
fastest	27	42
less memory	15	34

**AutonomousCar**

	GreatSPN	ITS-Tools
Score	869.22	897.54
fastest	9	1
smallest memory	6	1

**Model Checking Competition**  
12th edition, Bergen, Norway, June 2022  
Last Updated Jan 13, 2022

**Execution Summary**

	Max Memory Used (MB)	Time wait	CPU Usage
GreatSPN			
ITS-Tools			

**CANConstruction, P/T for ReachabilityCardinality, Memory consumption**

Megabytes

model instances

I/O wait

27/06/2021, 21:28

27/06/2021, 21:30

**Time to Process Known Models (ReachabilityDeadlock)**

seconds

Number of model instances processed

GreatSPN  
ITS-Tools  
LoLA  
Tapaal  
2020-gold  
BVT-2021

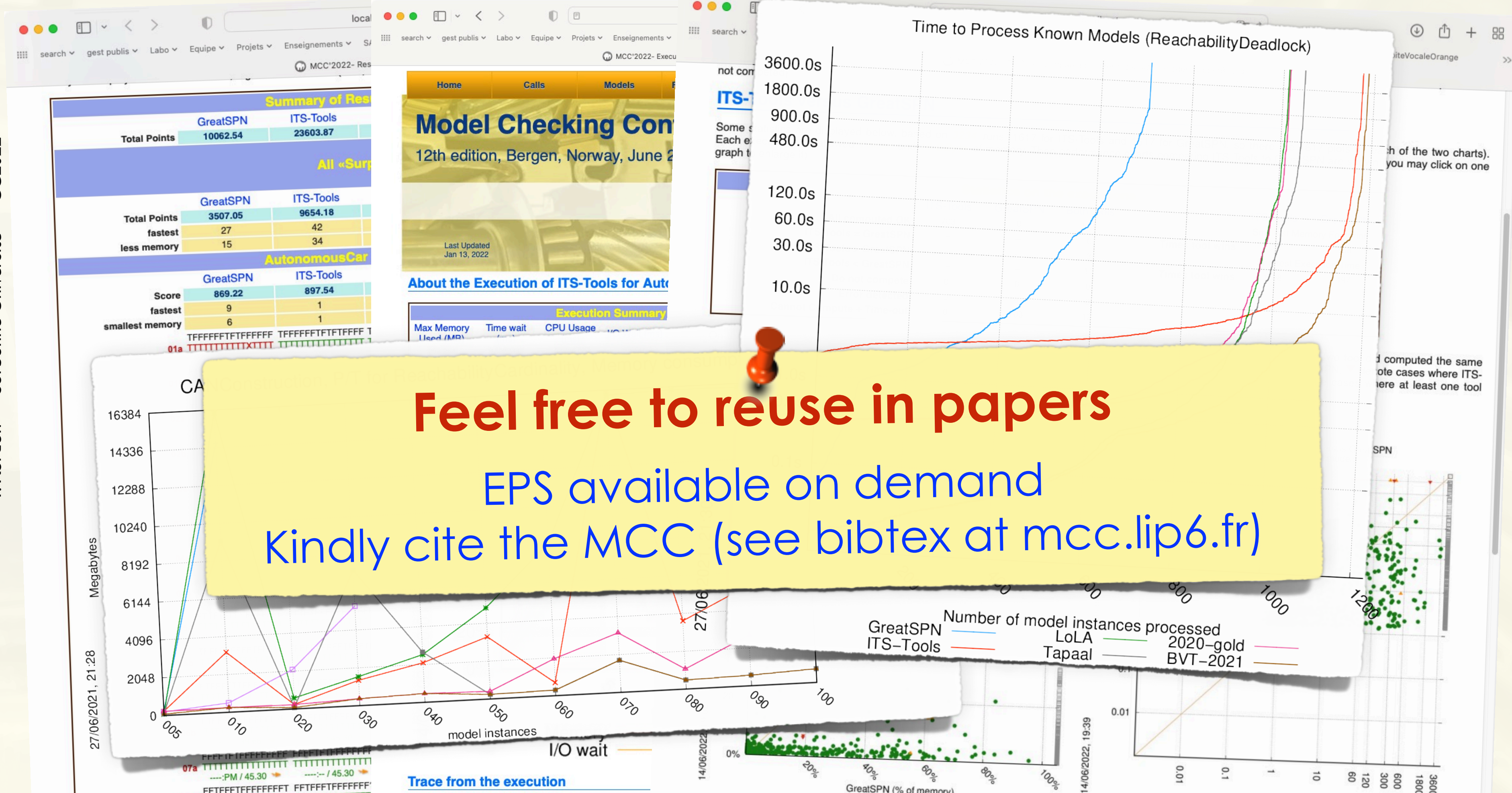
**SPN**

GreatSPN (% of memory)

Trace from the execution



- Many generated web pages (about 112K in 2022)
- Numerous generated charts (96814 in 2022)



## Handling some bugs in 2022 surprise models?

-  Maybe fine renaming of some surprise models

## Handling some bugs in 2022 surprise models?

- Maybe fine renaming of some surprise models

## The return of TOOLympics

- ETAPS '2023, Paris, April 22-27
  - Dirk Beyer (DE)
  - Arnd Hartmanns (NL)
  - Fabrice Kordon (FR)
  - Proceedings planned
- So, a shorter path to next MCC



**TOOLympics 2019**  
April 6-7, 2019  
Prague, Czech Republic  
<https://tacas.info/toolympics.php>

**Organizers**  
D. Beyer (LMU Munich, Germany)  
M. Huisman (University of Twente, Netherlands)  
F. Kordon (Sorbonne Université, France)  
B. Steffen (TU Dortmund, Germany)

**Involved Competitions**  
**CADE ATP System Competition**  
G. Sutcliffe (University of Miami, USA)  
**CHC-COMP**  
A. Gurfinkel (University of Waterloo, Canada), G. Fedyukovich (Princeton, USA)  
P. Ruemmer (Uppsala University, Sweden)  
**Confluence Competition**  
A. Middeldorp (University of Innsbruck, Austria), J. Nagele (Queen Mary University of London, UK), K. Shintani (JAIST, Japan)  
**CRV: The International Runtime Verification Competition**  
E. Bartocci (TU Wien, Austria), Y. Falcone (University of Grenoble Alps, France), G. Reger (University of Manchester, UK)  
**A Rodeo for Software Verification Tools based on Formal Methods**  
P. E. (NIST, United States), L. Badger (NIST, United States), B. Guttman (NIST, United States)  
I. Bojanova (NIST, United States), D. Flater (NIST, United States)  
**Model Checking Contest**  
F. Kordon (Sorbonne Université, France), H. Garavel (Inria, France), L. M. Messan (Université de Nanterre/LIP6, France)  
F. Hulin-Hubard (CNRS/Sorbonne Université, France), L. Jezequel (Université de Nantes, France)  
E. Paviot-Adet (Université Paris Descartes, France)  
**2019 Comparison of Tools for the Analysis of Quantitative Formal Models**  
A. Hartmanns (University of Twente, Netherlands)  
**Rewriting Engines Competition**  
H. Garavel (INRIA Grenoble, France)  
**Rigorous Examination of Reactive Systems Challenge**  
B. Steffen (TU Dortmund, Germany), F. Howar (TU Dortmund, Germany)  
F. Vaandrager (Radboud University, Netherlands), D. Hendriks (ESI (TNO), Netherlands)  
**SAT Competition 2018**  
T. Balyo (Karlsruhe Institute of Technology, Germany), M. Suda (Czech Technical University in Prague, Czech Republic)  
**Separation Logic Competition**  
M. Sighireanu (University Paris Diderot, France)  
**International Satisfiability Modulo Theory Competition**

## Handling some bugs in 2022 surprise models?

Maybe fine renaming of



### submission of legacy tools?

To see evolution of the state of the art?

▶ Arnd Hartmanns (NL)

▶ Fabrice Kordon (FR)

▶ Proceedings planned

So, a shorter path to next MCC

https://tacas.info/toolympics.php

**Involved Competitions**

**CADE ATP System Competition**  
G. Sutcliffe (University of Miami, USA)

**CHC-COMP**  
A. Gurfinkel (University of Waterloo, Canada), G. Fedyukovich (Princeton, USA)  
P. Ruemmer (Uppsala University, Sweden)

**Confluence Competition**  
A. Middeldorp (University of Innsbruck, Austria), J. Nagele (Queen Mary University of London, UK), K. Shintani (JAIST, Japan)

**CRV : The International Runtime Verification Competition**  
E. Bartocci (TU Wien, Austria), Y. Falcone (University of Grenoble Alps, France), G. Reger (University of Manchester, UK)

**A Rodeo for Software Verification Tools based on Formal Methods**  
P. E. (NIST, United States), L. Badger (NIST, United States), B. Guttman (NIST, United States)  
I. Bojanova (NIST, United States), D. Flater (NIST, United States)

**Model Checking Contest**  
F. Kordon (Sorbonne Université, France), H. Garavel (Inria, France), L. M. Messan (Université de Nanterre/LIP6, France)  
F. Hulin-Hubard (CNRS/Sorbonne Université, France), L. Jezequel (Université de Nantes, France)  
E. Paviot-Adet (Université Paris Descartes, France)

**2019 Comparison of Tools for the Analysis of Quantitative Formal Models**  
A. Hartmanns (University of Twente, Netherlands)

**Rewriting Engines Competition**  
H. Garavel (INRIA Grenoble, France)

**Rigorous Examination of Reactive Systems Challenge**  
B. Steffen (TU Dortmund, Germany), F. Howar (TU Dortmund, Germany)  
F. Vaandrager (Radboud University, Netherlands), D. Hendriks (ESI (TNO), Netherlands)

**SAT Competition 2018**  
T. Balyo (Karlsruhe Institute of Technology, Germany), M. Suda (Czech Technical University in Prague, Czech Republic)

**Separation Logic Competition**  
M. Sighireanu (University Paris Diderot, France)

**International Satisfiability Modulo Theory Competition**

## 🌐 Handling some bugs in 2022 surprise models?

📌 Maybe fine renaming of



### submission of legacy tools?

To see evolution of the state of the art?

- ▶ Arnd Hartmanns (NL)
- ▶ Fabrice Kordon (FR)



### Participate next year!!!

It is not so difficult (simple protocol)  
 A way to have «free» evaluation and comparison to the state of the art  
 An excellent way to find bugs





SORBONNE  
UNIVERSITÉ

And now...

let's have time for discussion

מנע  
בבכ