



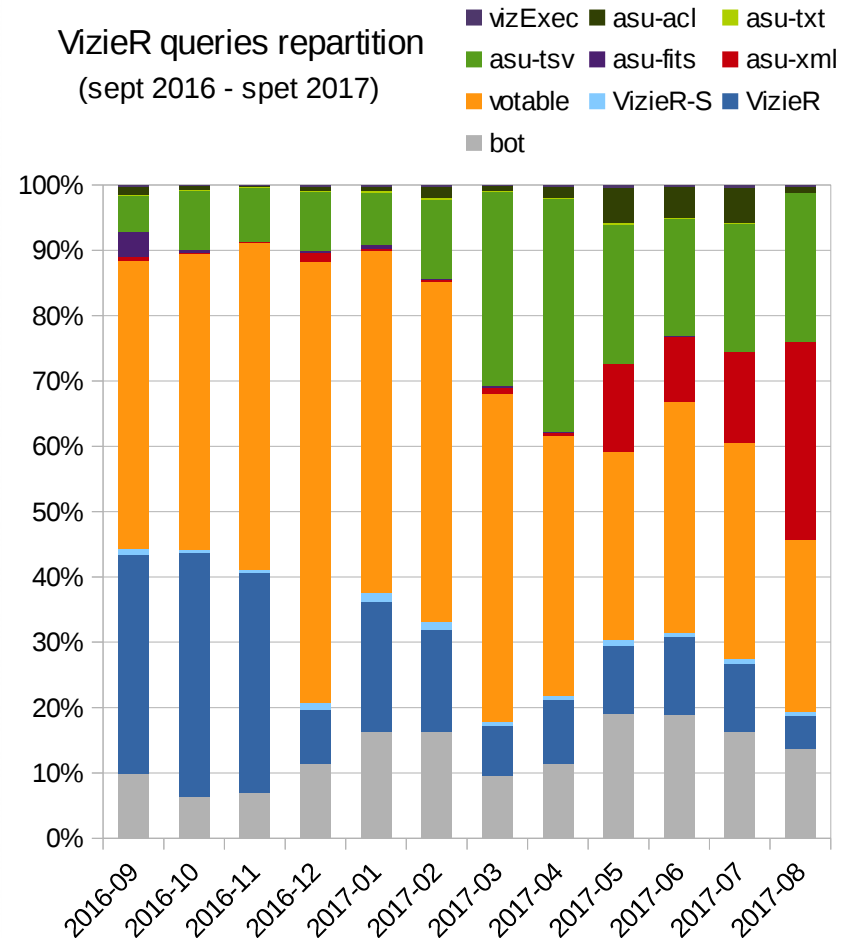
Retour sur l'utilisation des validateurs pour VizieR

G.Landais, S.Derriere



- Vizier est un service du CDS qui donne accès à des catalogues issus des publications et des centres spatiaux
- Un ensemble de données hétérogène mais avec des types de données communes: position, magnitude
- Volumétrie Vizier:
 - 17,000 catalogues (~35,000 tables)
 - ~32,000 images (110 catalogues)
 - ~6,000,000 spectres/time-series
~104,000 spectres si on enlève LAMOST/COROT
- Les services VO de Vizier et l'importance de la sortie VO
 - Conesearch
 - TAPVizieR
 - SIA/SSA/ObsTAP

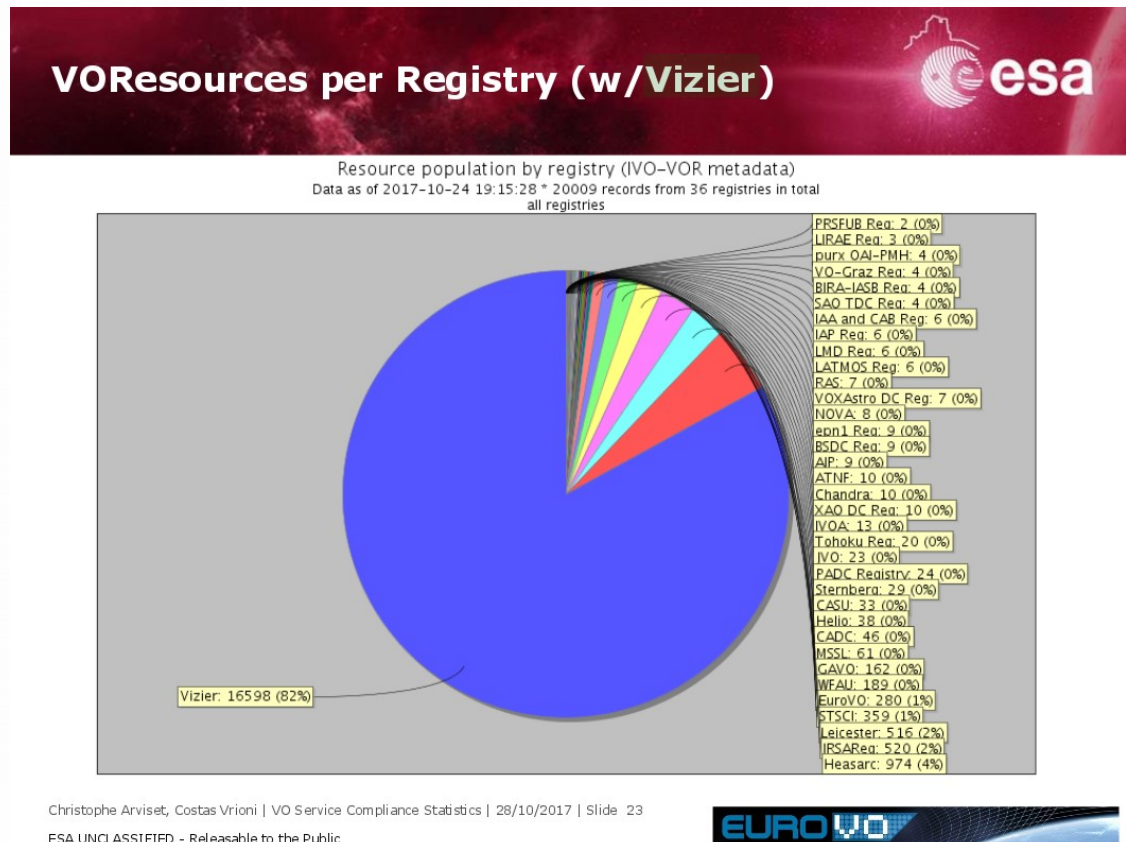
VizieR queries repartition
(sept 2016 - sept 2017)



Une prépondérance de VizieR dans le registry

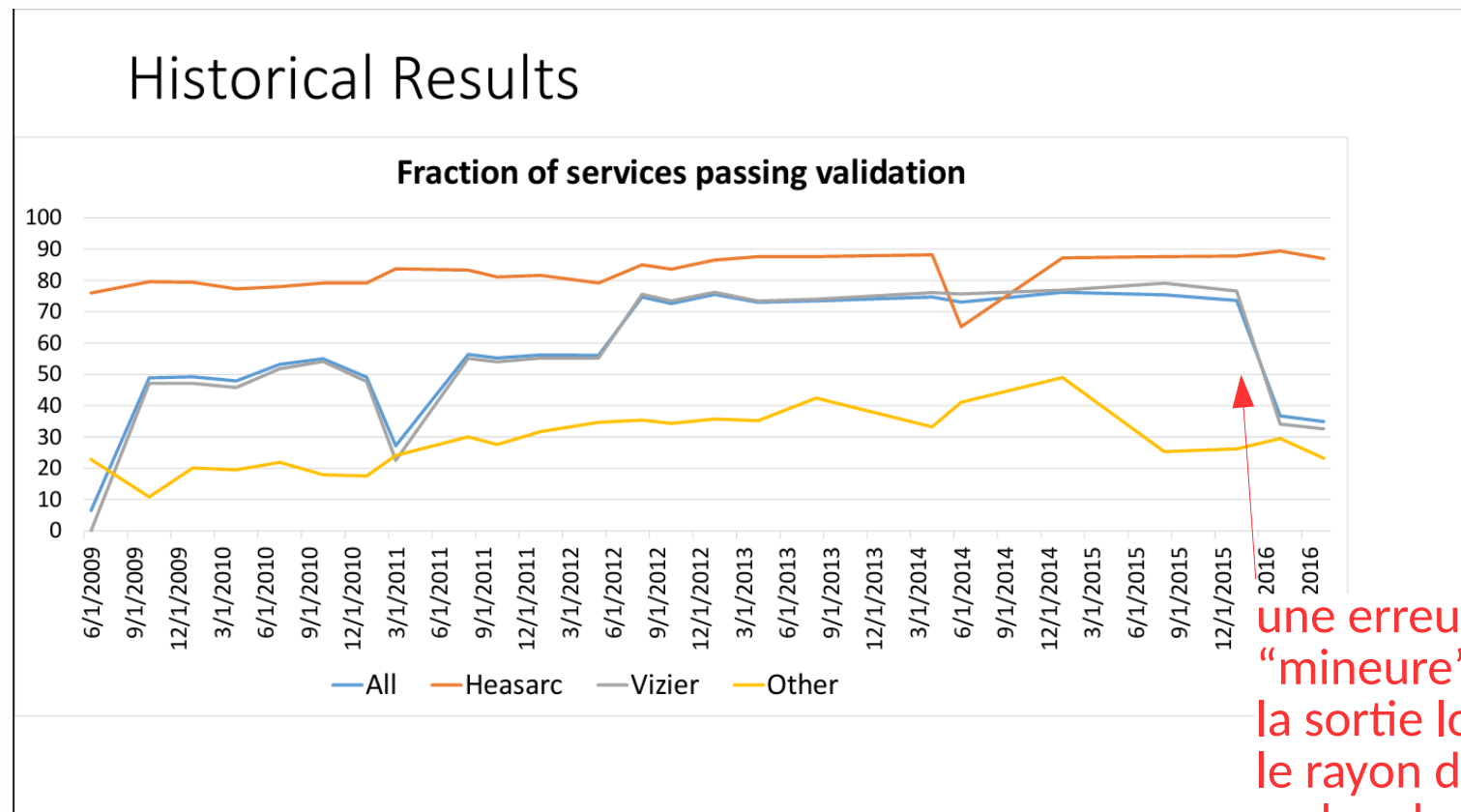


~17,000 entrées VizieR dans le registry aujourd'hui



Mise en évidence d'erreurs Vizier lors d'une conférence IVOA

Interop Trieste (2016) – validateur HEASARC



une erreur
"mineure" due à
la sortie lorsque
le rayon de
recherche = 0

□ Le conesearch Vizier



Autres corrections d'erreurs détectées

- UCD MAIN_ID absent

Explications:

- absence d'UCD (rare)
- (ou) nature des tables recues qui n'ont pas d'identifiants

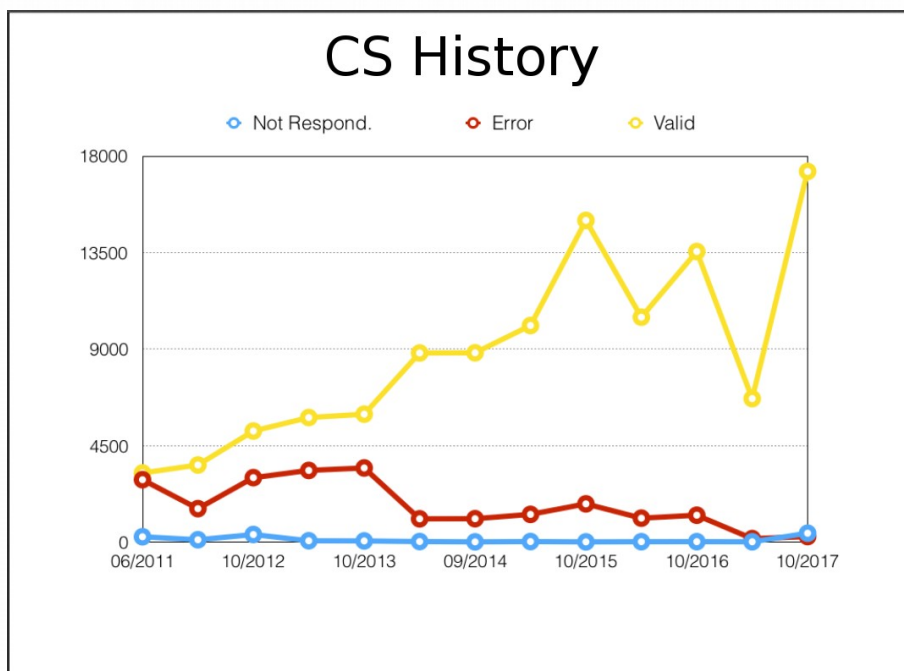
→ utilisation de la colonne “technique”
“recno”

- UCD MAIN_ID en double

- 2x assignation d'UCD principal

→ update du programme d'ingestion
Vizier pour empecher les 2x
assignations

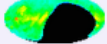
Interop Santiago (2017)
– validateur VOParis



Le conesearch VizieR



Exemple de catalogue VizieR qui n'a pas d'identifiant principal

1975AGK3.C.....0H [ReadMe+ftp](#) 

I/61B [Post annotation](#) [Similar Catalogs](#)

1.1/I/61B/agk3 Catalog Data (183145 rows)

[Simple Constraint](#) [List Of Constraints](#)

Query by [Constraints](#) applied on Columns (Output Order: + -)

Show	Sort	Column	Clear	Constraint	Explain (UCD)
<input type="checkbox"/>	<input type="radio"/>	recno	<input type="button" value="Clear"/>		Record number assigned by the VizieR team. Should Not be used for identification. (meta.record)
<input checked="" type="checkbox"/>	<input type="radio"/>	AGsign	<input type="button" value="Clear"/>	(char) [-+]Sign of AGK3 number. (meta.id.part:meta.main)	
<input checked="" type="checkbox"/>	<input type="radio"/>	AGzone	<input type="button" value="Clear"/>	AGK3 zone. (meta.id.part:meta.main)	
<input checked="" type="checkbox"/>	<input type="radio"/>	AGnum	<input type="button" value="Clear"/>	(i) AGK3 number (Note) (meta.id.part:meta.main)	
<input checked="" type="checkbox"/>	<input type="radio"/>	m_AGnum	<input type="button" value="Clear"/>	(char) [0ab]Component for duplicates (Note) (meta.code.multip)	
<input checked="" type="checkbox"/>	<input type="radio"/>	Pmag	<input type="button" value="Clear"/>	mag Photographic magnitude from AGK2 (phot.mag:em.opt)	
<input checked="" type="checkbox"/>	<input type="radio"/>	Sp	<input type="button" value="Clear"/>	(char) Spectral type (Note) (src.spType)	
<input checked="" type="checkbox"/>	<input type="radio"/>	RA1950	<input type="button" value="Clear"/>	"h:m:s"(i) Right ascension hours (1950.0) (pos.eq.ra:meta.main)	
<input checked="" type="checkbox"/>	<input type="radio"/>	DE1950	<input type="button" value="Clear"/>	"d:m:s"(i) Declination degrees (1950.0) (pos.eq.dec:meta.main)	
<input type="checkbox"/>	<input type="radio"/>	nobs	<input type="button" value="Clear"/>	Number of photographic observations (Note) (meta.number)	
<input checked="" type="checkbox"/>	<input type="radio"/>	Ep-1900	<input type="button" value="Clear"/>	yr Epoch of AGK3 position minus 1900 (time.epoch)	
<input type="checkbox"/>	<input type="radio"/>	<input type="button" value="Reset All"/>	<input type="button" value="Clear"/>		(i) indexed column <input type="button" value="Submit"/>
<input checked="" type="checkbox"/>	<input type="radio"/>	pmRA	<input type="button" value="Clear"/>	mas/yr Annual proper motion in right ascension (Note) (pos.pm:pos.eq.ra)	
<input checked="" type="checkbox"/>	<input type="radio"/>	pmDE	<input type="button" value="Clear"/>	mas/yr Annual proper motion in declination (Note) (pos.pm:pos.eq.dec)	
<input type="checkbox"/>	<input type="radio"/>	dEpoch	<input type="button" value="Clear"/>	yr Epoch difference AGK3-AGK2 (time.interval)	
<input checked="" type="checkbox"/>	<input type="radio"/>	BD	<input type="button" value="Clear"/>	(char) Full BD catalog designation (Note) (meta.id)	
<input checked="" type="checkbox"/>	<input type="radio"/>	m_BD	<input type="button" value="Clear"/>	(char) Component of multiple system (Note) (meta.code.multip)	
<input type="checkbox"/>	<input type="radio"/>	Sum	<input type="button" value="Clear"/>	Sum of discrepancy codes (Note) (meta.code)	
<input type="checkbox"/>	<input type="radio"/>	dRA	<input type="button" value="Clear"/>	arcsec The difference in right ascension in the sense AGK3 - AGK2modified (Note) (pos.eq.ra:arith.diff)	
<input type="checkbox"/>	<input type="radio"/>	dDE	<input type="button" value="Clear"/>	arcsec The difference in declination in the sense AGK3 - AGK2modified (pos.eq.dec:arith.diff)	
<input checked="" type="checkbox"/>	<input type="radio"/>	_RA.icrs	<input type="button" value="Clear"/>	"h:m:s"(n)(i) Right ascension (ICRS) at Epoch=J2000, proper motions taken into account (computed by VizieR, not part of the original data) (pos.eq.ra)	
<input checked="" type="checkbox"/>	<input type="radio"/>	_DE.icrs	<input type="button" value="Clear"/>	"d:m:s"(n)(i) Declination (ICRS) at Epoch=J2000, proper motions taken into account (computed by VizieR, not part of the original data) (pos.eq.dec)	
<input type="checkbox"/>	<input type="radio"/>	<input type="button" value="Reset All"/>	<input type="button" value="Clear"/>		(n) indicates a possible blank or NULL column (i) indexed column <input type="button" value="Submit"/>

Ret

□ Le conesearch VizieR



L'augmentation des entrées VizieR dans le registry

le conesearch ne s'applique qu'à une seule table, or la granularité des entrées VizieR dans les registries VO est le catalogue qui peut contenir plusieurs tables

→ “migration” des entrées VizieR avec une granulation au niveau “table”

17,000 entrées auquel ont été ajoutées des capacités conesearch+TAP par table

The screenshot displays the GAVO (German Astrophysical Virtual Observatory) web interface. On the left is a navigation sidebar with buttons for 'Quick Start', 'Help', and various services like 'Image Services', 'Spectral Services', 'TAP Services', 'ObsTAP Services', 'Radio Resources', and 'IR Resources'. The main content area is titled 'Web Interface to the Relational Registry'. It features a search interface with a dropdown menu set to 'Text Fields', a search type dropdown set to 'match', and a search box containing 'Bixler'. Below the search box is a '+ Add Constraint' button. At the bottom of the search area, there are controls for 'Items displayed per page' (set to 20) and 'Offset' (set to 0 items), along with a 'Run Query' button and a star icon. A 'Query History' section with left and right navigation arrows is also visible. At the bottom of the page, there is a link to 'Far-UV excess object survey (Bixler+ 1991)', a user profile icon, and navigation links for 'SCS', 'SCS', and 'WEB'.

□ Le conesearch Vizier



Ambiguïté des spécifications

- Exemple le conesearch:

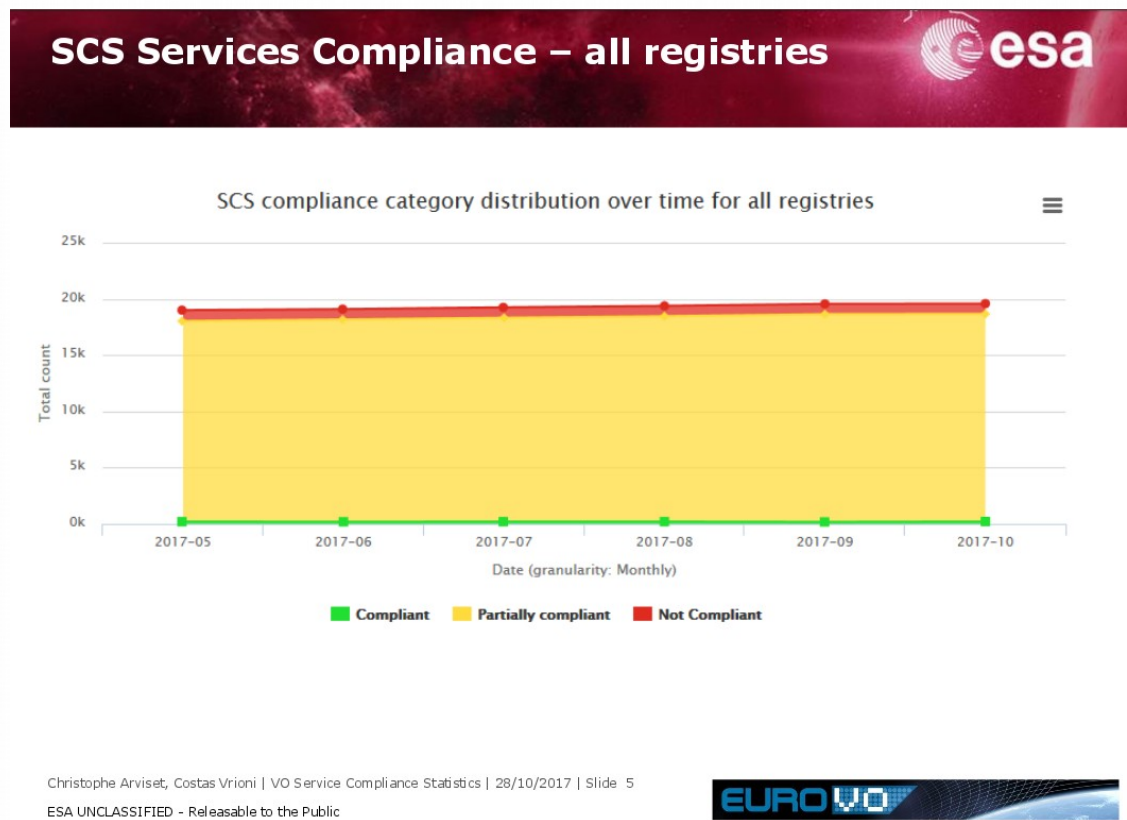
Le tag <INFO> doit être utilisé pour signaler une position non comprise.

Or certain validateurs excluent tout autre type de tags <INFO>

- Ambiguïté sur les paramètres non spécifiées par le protocole

Evolution de l'accès conesearch VizieR

- Une conséquence de l'interop Santiago (2017) suite à la presentation des statistiques du validateur de l'ESA



□ Le conesearch Vizier



Explication de l'erreur (warning) relevée par le validateur de l'ESA

- Une interprétation de l'URL racine d'un conesearch Vizier

- Exemple:

<http://vizier.u-strasbg.fr/viz-bin/votable/-A?-source=II/246/out&CS=0+0&SR=1>

URL racine de la ressource 2MASS

Paramètres SCS

- Une nouvelle URL proposée:

<http://vizier.u-strasbg.fr/viz-bin/conesearch/II/246/out?CS=0+0&SR=1>

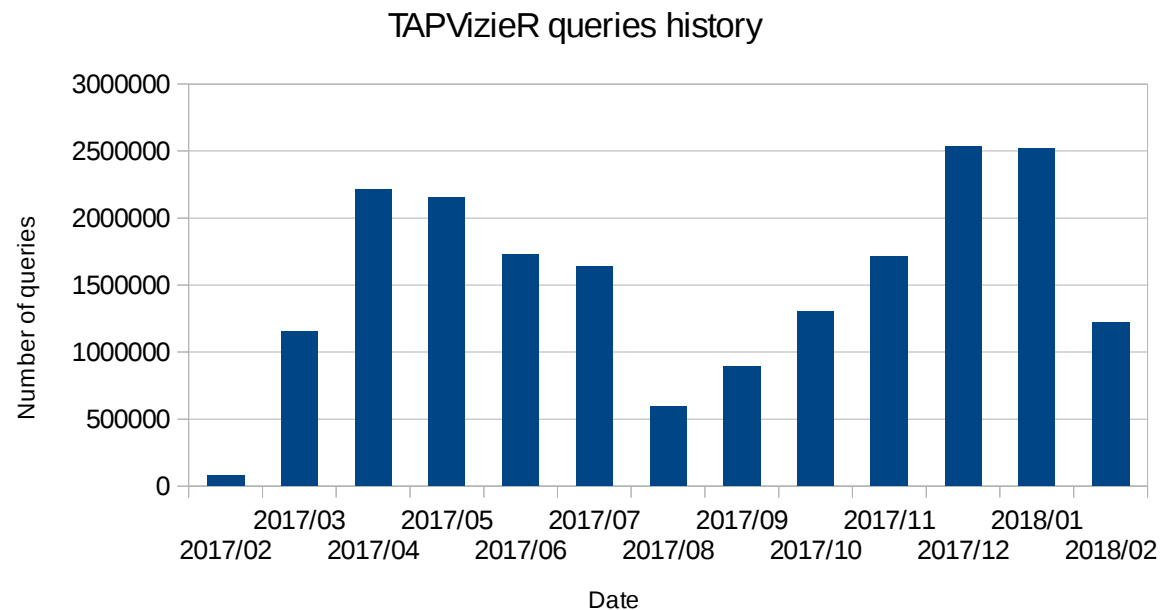
Validation du service TAPVizieR



Validation du service TAPVizieR

- Évolution des requetes dans TAPVizieR – un succès?

Moyenne sur l'année
~50,000 req sync/jour



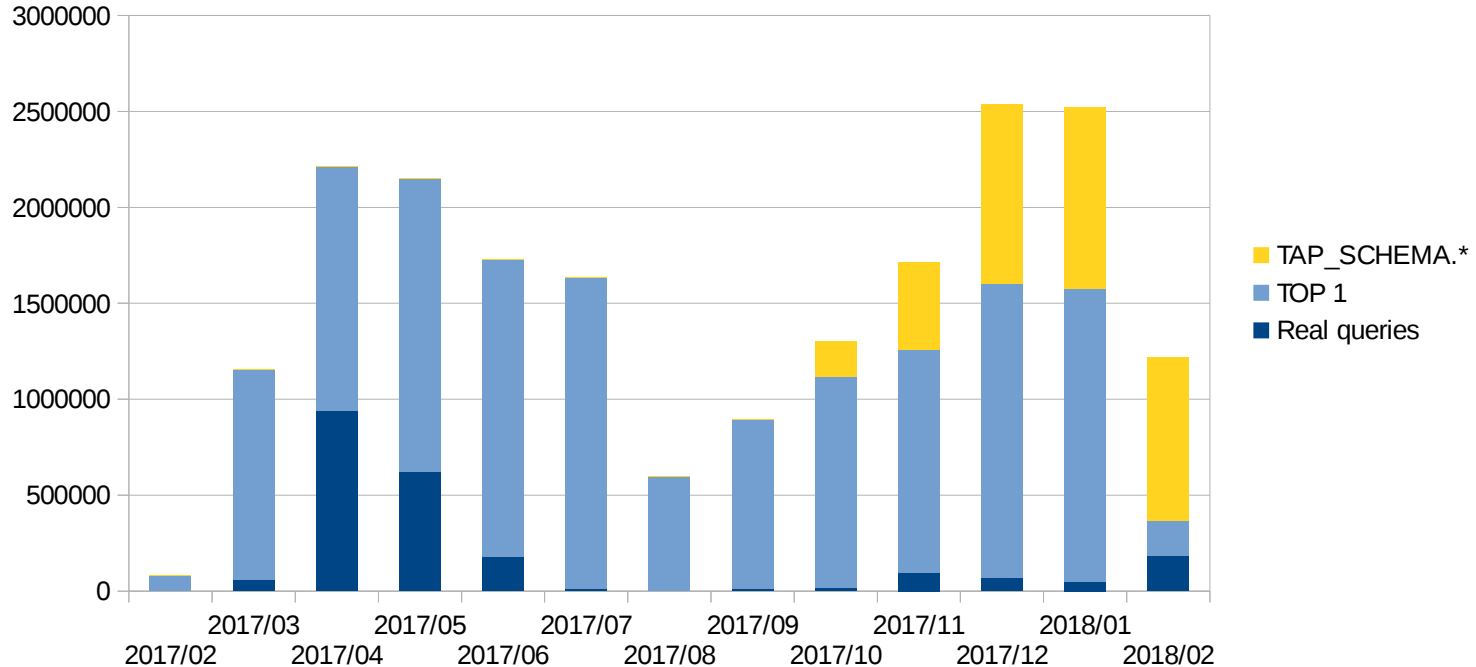
Validation du service TAPVizieR



TAPVizieR overhead queries

- ~36,000 req/jour : SELECT TOP 1 FROM {table}
- une augmentation des requêtes sur le TAP_SCHEMA
- Nombre de requetes réelles par jour :
~6,000 req/jour - soit 11.4% du nombre total de requêtes

TAPVizieR - queries



□ Validation du service TAPVizieR



Quelques explications de l'overhead

- VizieR a ~35,000 tables (cf SELECT TOP 1)
 - L'URL /table (protocole TAP) ne renvoie dans le cas de TAPVizieR que la description de la table et exige un autre moyen d'obtenir les infos de description des colonnes
- TOPCAT exécute pour chaque table VizieR sélectionnée:
- TAP_SCHEMA.tables
 - TAP_SCHEMA.columns

□ Validation du service TAPVizieR



TAPVizieR overhead queries

Exemple de Requetes journalières illustrant l'overhead TAPVizieR
(issues de voparis-vo.obspm.fr.)

```
ERROR 01/02/18 00:08:08 (TAPVizieRServlet.java) - sync | 145.238.193.18 |  
STILTS/3.0-9+ Java/1.7.0_151 | | sql=SELECT TOP 1 * FROM VII/220A/barnard
```

```
INFO 16/02/18 22:11:49 (TAPVizieRServlet.java) - sync | 133.40.215.124 |  
Java/1.8.0_66 | votable | sql=SELECT * FROM TAP_SCHEMA.columns AS t0 WHERE  
T0.table_name = 'II/220/hr1035'
```

□ Conclusions



- Les services de validations sont très utiles
- Ils n'ont pas tous le même niveau d'exigences:
- Une erreur même minime sur VizieR a de grosses incidences sur les statistiques générales montrées à l'IVOA.
- Ils relèvent tous de problèmes et se complètent (une erreur sur le standard n'est pas forcément relevé par l'ensemble des validateurs)

Conclusion sur l'overhead TAPVizieR

STILTS dispose d'une possibilité de validation qui vérifie la cohérence du schéma. Très utile pour les développeurs il est aussi utilisé par les validateurs.

Est ce un bon choix?