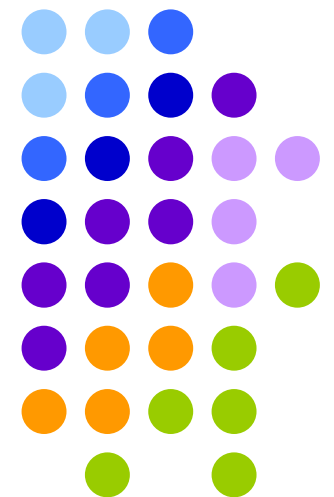
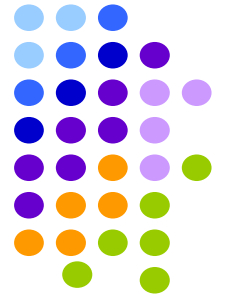


Measuring the quality of assessment using questions generated from the Semantic Web

Conejo, Barros & Bertoa
Universidad de Málaga, Spain



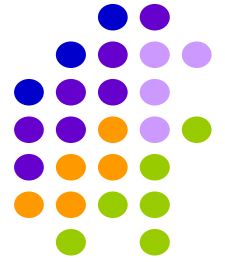
Motivation



Computer based testing make possible to generate questions automatically from databases.

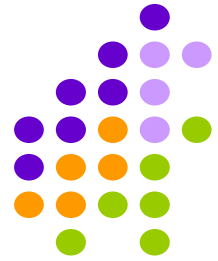
- Advantages for teachers: Generate a high number of questions with less effort
- Advantages for students: The same test can be taken several times as a learning tool.
- Problems: Moving from a hand-made to an industrial process require higher quality control.
- Limitations: The knowledge that can be assessed should be well structured.

Objectives



- Explore the techniques that can be used for automatic question generation
- Analyze the source of potential errors.
- **Measure** the effect of using automatically generated questions from incomplete or imperfect repositories in the quality of assessment.

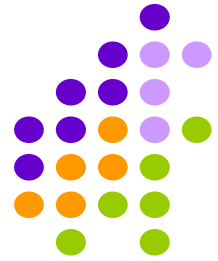
The SIETTE assessment env.



The screenshot shows the SIETTE login interface. At the top left is the SIETTE logo, with 'Siette' in red and green. To the right are links for 'Portal' (with a green checkmark icon) and 'Documentation' (with a document icon). Below these are language options: 'espanol' and 'deutsch', both with checked checkboxes. Further right are mode options: 'Individual mode' (with a person icon) and 'Collaborative Mode' (with a group icon). The main login area contains two input fields: 'User name:' and 'Password:'. A green 'Send' button with a right-pointing arrow is positioned below the password field. At the bottom, there are links for 'I forgot my password' and 'New user', both with checked checkboxes. A small envelope icon is located in the bottom right corner of the form area.

<https://www.siette.org>

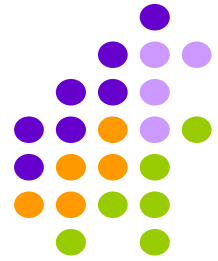
The SIETTE item models



There are three item model types. Other types of questions can be converted to one of these:

- MCQ-SA. Multiple-choice, single answer.
- MCQ-MA. Multiple-choice, multiple answer.
- SA. Short-answer questions. The answers are recognized using patterns:
 - Regular expressions: Text, numbers, etc.
 - Grammar patterns
 - Math patterns
 - Musical patterns

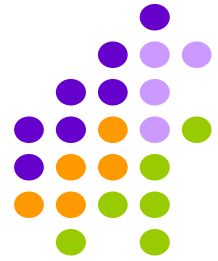
Pattern examples



Stem	<i>Who is the composer of the Moonlight Sonata?</i>
Pattern	{Ludwig {van}} Beethoven
Recognized answers	Ludwig van Beethoven Beethoven beethoven
Not recognized answers	Lewis Beethoveen

- Multiple patterns can be defined for the same question
- Upper/Lower case, accent, and misspelling variations can be controlled additionally.

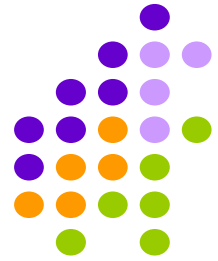
Pattern examples



Stem	<i>A car takes 30 minutes to go from a city A to a city B that distance 20 miles. What is its average speed?</i>
Pattern	[40 miles/h %1]
Recognized answers	40 miles/h 64.37 km/h 17.88 m/s
Not recognized answers	40 km/h

- Precision can be controlled.
- Physical magnitudes are automatically converted

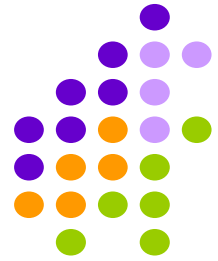
Question generation



SIETTE can dynamically generate questions from templates written in JSP (and their answer pattern):

```
<%  
    int x      = Math.abs( siette.util.Random.nextInt() % 5 ) + 1;  
    int y      = Math.abs( siette.util.Random.nextInt() % 5 ) + 1;  
    int z      = x+y;  
%>  
What is the result of this operation:<br/>  
<center> <%= x %> + <%= y %> </center>
```


(1) Questions from spreadsheets

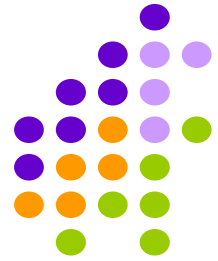


There is an API to generate questions from tables:

	A	B	C	D	E
1	ATOMIC_NUM	ATOMIC_MASS	VALENCY	SYMBOL	NAME
2	1	100.797		1 H	Hidrógeno
3	2	40.026		0 He	Helio
4	3	6.939		1 Li	Litio
5	4	90.122		2 Be	Berilio
6	5	10.811		3 B	Boro
7	6	1.201.115	2 4	C	Carbono
8	7	140.067	-3 1 2 3 4 5	N	Nitrógeno
9	8	159.994		-2 O	Oxígeno

```
<%@page import="siette.util.corpus.Table"%>
<%
    Table table = new Table("demo/periodic-table.xls");
    String[] element = table.select();
    String name = table.get( element, "NAME" );
    String symbol = table.get( element, "SYMBOL" );
%>
What is the symbol for the chemical element "<i><%= name %></i>"?
```

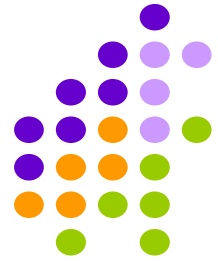
(2) Questions from databases



The same idea but getting the table data from SQL databases queries:

```
<%@page import="siette.util.corpus.DatabaseTable"%>
<%
    String query = "SELECT S.BINOMIAL, S.COMMON_NAME, P.IMG "
                  + " FROM SPECIES S JOIN PHOTOS P ON P.SPECIE = S.SPECIE "
                  + " WHERE P.FEATURE='Leaf' "
                  ;
    DatabaseTable table = new DatabaseTable("demo/tree.properties", query);
    String[] plant = table.select();
    String img = table.get( plant, "IMG" );
    String binomial = table.get( plant, "BINOMIAL" );
    String plant_name = table.get( plant, "NOMBRE" );
%>
Write the scientific name of the specie that have this leaf:
<center> <IMG SRC='<%= img %>'> </center>
```

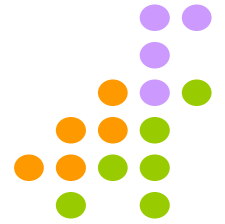
(3) Questions from the Semantic Web



The same idea again, but taken data from SPARQL queries to Semantic Web endpoints (like DBpedia):

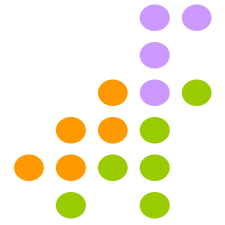
```
<%@page import="siette.util.corpus.WebTable"%>
<%
String query = "SELECT DISTINCT ?nombre ?name ?population ?flag ?img "
+"WHERE { "
+"    ?country a dbpedia-owl:Country ; rdfs:label ?name ; dbo:flag ?flag . "
+"    ?country ?hasPopulation ?population ; dbo:thumbnail ?img . "
+"    ?country dct:subject dbc:Member_states_of_the_United_Nations "
+"    FILTER (langMatches(lang(?name), \"en\")) "
+"    FILTER (?population > 1000000) "
+"} "
;
    WebTable table = new WebTable("http://dbpedia.org/sparql",query);
    String[] country = table.select();
    String name = table.get( country, "name" );
    String img = table.get( country, "img" );
%>
<center>
Which is the country of this flag? <br/>
<center><IMG src="<%= img %>" /><br/></center>
```

Problems with Generated Questions



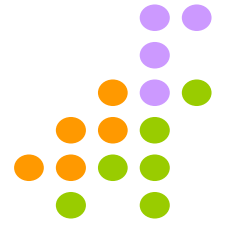
- Is it really the same?
 - In the first and (maybe) the second case, we have full control on the data available to generate the question
 - In the (second and) third case, the number of records are much higher.
 - The less control and the higher the data volume makes more difficult to guarantee 100% question quality.

Sources of errors



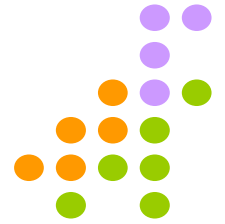
- **Incorrect or mislabeled content.**
 - i.e. An image is showing something different from expected.
- **Missing or incomplete content.**
 - i.e. There is no image available or an object is not labeled to belong a given class.
- **Undesired or unexpected content.**
 - i.e. You are looking for an image of an animal and get the image of its skeleton.
- **Incompleteness of the response pattern**
 - The automatic procedure to construct the response pattern fails.

Quality of questions



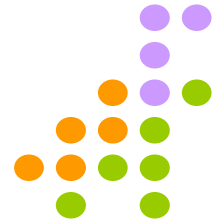
- Does 100% quality exist?
- Is it really necessary for assessment ?
 - Even with hand-made questions errors might occur. Psychometry already considers the concepts of assessment Validity and Reliability.
 - “Industrial production” introduces “Quality Control”
 - Detect and correct errors whenever possible.
 - Measure and quantify the acceptance criteria of uncorrected errors.

Measuring the queality of assessment



- Assessment purposes:
 - Low-stake assessment
 - Fun.
 - Self-assessment / Learning
 - Initial level classification.
 - High-stake assessment
 - On going assessment during the course.
 - Final exam.
 - Classification of candidates to get a job.

Assessment Validity and Reliability



Unreliable & Invalid



Unreliable, But Valid

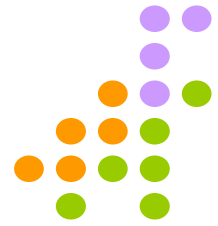


Reliable, Not Valid



Both Reliable & Valid

Reliability



- Roughly speaking, classical test theory reliability can be defined as the correlation coefficient between the scores of two “parallel” test:

- Spearman-Brown

$$\rho_{xx'} = \frac{2\rho_{12}}{1 + \rho_{12}}$$

- λ_4 Guttman-Flanagan

$$\rho_{xx'} = \frac{\sigma_V^2}{\sigma_X^2} = 2 \left(1 - \frac{\sigma_{X1}^2 + \sigma_{X2}^2}{\sigma_X^2} \right)$$

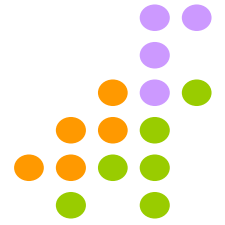
- Kuder-Ricardson KR-20

$$r = \frac{K}{K-1} \left[1 - \frac{\sum_{i=1}^K p_i q_i}{\sigma_X^2} \right]$$

- Cronbach's alpha

$$\alpha = \frac{K}{K-1} \left(1 - \frac{\sum_{i=1}^K \sigma_{Y_i}^2}{\sigma_X^2} \right)$$

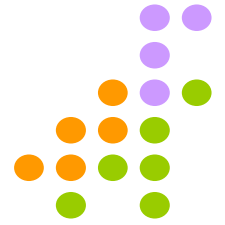
Confidence intervals



- An assessment is a measuring device, and the “true score” V depends on score obtained in the test X , and the device reliability $\rho_{XX'}$
 - i.e. The standard 95% confidence interval of the test score, assuming normal error distribution is

$$V \in \left[X \pm 1.96\sigma_X \sqrt{1 - \rho_{XX'}^2} \right]$$

Test length



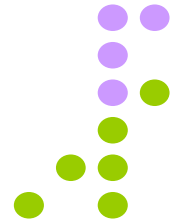
- Reliability is related to test length

$$\rho_{xx'}^* = \frac{R\rho_{xx'}}{1 + (R - 1)\rho_{xx'}},$$

where R is the ration between the two test.

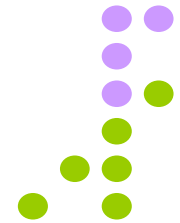
Theoretically, reliability increase when test length increase

Quality control.



- Before the test is taken
 - Detect invalid questions generated from templates (instances) and improve the template
 - Quantify the number of potential invalid questions that might be generated from a template.
 - Estimate reliability loss.
- After the test has been taken
 - Detect and remove (cancel) incorrect instances.
 - Recalculate reliability.
 - Re-assess.

Preview



Administrador · 2018/06/07 11:12:04 - 115.red-83-41-82.dynamicip.rima-tde.net

english

Change Profile

Log out

Questions

Test

Students

Documentation

New >

Search

Groups >

Preferences

List of Subjects

> Demo (English) > Art and Literature > Painting > Paintings (Imagen) SPARQL

Demo (English)

- Filosofia
- Maths
 - Simple pattern
 - GZA - Item models
 - Problems with templates
 - Algebra
 - Calculus
 - Geometry
 - Estadística
 - Arithmetics
- Physics
- Chemistry
- Computers
- Economy
 - Contabilidad
- Biology
 - Cell biology
 - Botany
 - Zoology
- Geography
 - single_answer_with_images
- External items
- Scholar
- Sistema solar
- Geografía de España
- Art and Literature
 - 1-2 Paintings
 - 1-3 Who wrote the Quixote?
 - Lloyd Puzzle
 - Sets
 - Music
 - Painting
 - Author of this painting - SI
 - Paintings (Imagen) SPAR
 - Literature
- History
- Linguistics
- Testing
- Entrettenimiento

Preview

Information

Content

Display

Selection

Evaluation

Advanced

Branches

Analyzer

Sessions

Paintings (Imagen) SPARQL

Who is the author of this painting?



This painting is entitled: "s Mother"

James McNeill Whistler

Evaluate

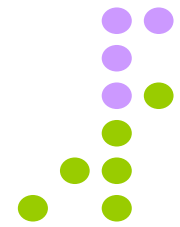
Save changes

Delete

Clone question

Create instances

Preview



Administrador · 2018/06/07 11:12:04 - 115.red-83-41-82.dynamicip.rima-tde.net

english

Change Profile

Log out

Questions

Test

Students

Documentation

New >

Search

Groups >

Preferences

List of Subjects

> Demo (English) > Art and Literature > Painting > Author of this painting - SPARQL

Demo (English)

- [-] Filosofia
- [-] Maths
 - Simple pattern
- [-] GZA - Item models
- [-] Problems with templates
- [-] Algebra
- [-] Calculus
- [-] Geometry
- [-] Estadística
- [-] Arithmetics
- [-] Physics
- [-] Chemistry
- [-] Computers
- [-] Economy
- [-] Biology
- [-] Geography
- [-] Art and Literature
 - 1-2 Paintings
 - 1-3 Who wrote the Quixote?
 - Lloyd Puzzle
 - Sets
 - Music
 - [-] Painting
 - Author of this painting - SPARQL
 - Literature
- [-] History
- [-] Linguistics
- [-] Testing
- [-] Entretenimiento

- Preview
- Information
- Content
- Display
- Selection
- Evaluation
- Advanced
- Branches
- Analyzer
- Sessions

Author of this painting - SPARQL

Who is the author of this painting?



Randomly selected from DBpedia out of 2810 images

This painting is entitled "House in Provence"

____ Cézanne (US /seɪˈzæn/ or UK /siˈzæn/; French: [pɔl sezan]; 19 January 1839 – 22 October 1906) was a French artist and Post-Impressionist painter whose work laid the foundations of the transition from the 19th-century conception of artistic endeavour to a new and radically different world of art in the 20th century. Cézanne's often repetitive, exploratory brushstrokes are highly characteristic and clearly recognizable. He used planes of colour and small brushstrokes that build up to form complex fields. The paintings convey Cézanne's intense study of his subjects.

Paul Cézanne

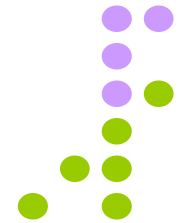
Save changes

Delete

Clone question

Create instances

Preview.



Administrador . 2018/06/07 11:12:04 - 115.red-83-41-82.dynamicip.rima-tde.net

english

Change Profile

Log out

Questions

Test

Students

Documentation

New >

Search

Groups >

Preferences

List of Subjects

> Demo (English) > Art and Literature > Painting > Author of this painting - SPARQL

Demo (English)

- [-] Filosofia
- [-] Maths
 - Simple pattern
 - GZA - Item models
 - Problems with templates
 - Algebra
 - Calculus
 - Geometry
 - Estadística
 - Arithmetics
- [-] Physics
- [-] Chemistry
- [-] Computers
- [-] Economy
- [-] Biology
- [-] Geography
- [-] Art and Literature
 - 1-2 Paintings
 - 1-3 Who wrote the Quixote?
 - Lloyd Puzzle
 - Sets
 - Music
 - [-] Painting
 - Author of this painting -
 - Literature
- [-] History
- [-] Linguistics
- [-] Testing
- [-] Entretenimiento

Preview

Information

Content

Display

Selection

Evaluation

Advanced

Branches

Analyzer

Sessions

Author of this painting - SPARQL



Who is the author of this painting?



Randomly selected from DBpedia out of 2810 images

This painting is entitled "Takka Takka "

____ Fox ____ (pronounced /'ɪktən stɑɪn/; October 27, 1923 – September 29, 1997) was an American pop artist. During the 1960s, along with Andy Warhol, Jasper Johns, and James Rosenquist among others, he became a leading figure in the new art movement. His work

Button to mark this instance as "canceled"

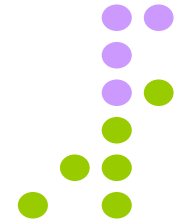
Save changes

Delete

Clone question

Create instances

Preview.




- Demo (English)
- Philosofia
- Maths
- Physics
- Chemistry
- Computers
- Economy
- Biology
- Geography
- Art and Literature
 - 1-2 Paintings
 - 1-3 Who wrote the Quixote?
 - Lloyd Puzzle
 - Sets
 - Music
 - Painting
 - Author of this painting -
 - Paintings SPARQL
 - Literature
- History
- Linguistics
- Testing
- Entretenimiento

Preview Information Content Display Selection Evaluation Advanced Branches Analyzer Sessions

Author of this painting - SPARQL

Who is the author of this painting?



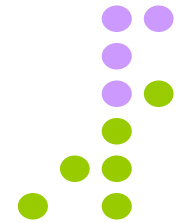
Randomly selected from DBpedia out of 2810 images

This painting is entitled "Juneau Monument"
____ (1832–1902), also known as ____ Hamilton ____, was an American sculptor. ____ was born in New York City and active in New York and Florence, Italy. Today he is best known for his Antea

Button to mark this instance as "canceled"

- Save changes
- Delete
- Clone question
- Create instances

Preview.



Administrador · 2018/06/07 11:12:04 - 115.red-83-41-82.dynanicip.rima-tde.net

english

Change Profile

Log out

Questions

Test

Students

Documentation

New >

Search

Groups >

Preferences

List of Subjects

> Demo (English) > Maths > Matematician SPARQL

Demo (English)

- Philosofía
- Maths
 - Matematician SPARQL
 - Simple pattern
 - GZA - Item models
 - Problems with templates
 - Algebra
 - Calculus
 - Geometry
 - Estadística
 - Arithmetics
- Physics
- Chemistry
- Computers
 - 3-1 Programming loops 1
 - GCD
 - GCD & LCM
 - 1. GCD
 - 2. LCM
 - 2. LCM (Copy)
 - Programming loops
 - single_answer.xml
- Informática teórica
 - C / C++
 - Java
 - Redes
- Economy
- Biology
- Geography
- Art and Literature
 - 1-2 Paintings
 - 1-3 Who wrote the Quixote?
 - Lloyd Puzzle
 - Sets
- Music
- Painting
 - Author of this painting - SI
 - Paintings (Imagen) SPAR
- Literature

Preview

Information

Content

Display

Selection

Evaluation

Advanced

Branches

Analyzer

Sessions

Matematician SPARQL

Who is this mathematician?



In particular, he is recognized for his discovery of an original method of finding the greatest and the smallest ordinates of curved lines, which is analogous to that of differential calculus, then unknown, and his research into number theory. He made notable contributions to analytic geometry, probability, and optics. He is best known for ___'s Last Theorem, which he described in a note at the margin of a copy of Diophantus' Arithmetica.

Pierre de Fermat

Evaluate

Save changes

Delete

Clone question

Create instances

Preview.



Administrador . 2018/06/07 11:12:04 - 115.red-83-41-82.dynamicip.rima-tde.net

english

Change Profile

Log out

Questions

Test

Students

Documentation

New >

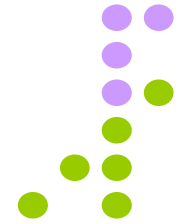
Search

Groups >

Preferences

List of Subjects

> Demo (English) > Maths > Matematician SPARQL



Demo (English)

- [-] Filosofia
- [-] Maths
 - Matematician SPARQL
 - Simple pattern
 - [-] GZA - Item models
 - [-] Problems with templates
 - [-] Algebra
 - [-] Calculus
 - [-] Geometry
 - [-] Estadística
 - [-] Arithmetics
- [-] Physics
- [-] Chemistry
- [-] Computers
 - 3-1 Programming loops 1
 - GCD
 - [-] GCD & LCM
 - 1. GCD
 - 2. LCM
 - 2. LCM (Copy)
 - Programming loops
 - single_answer.xml
- [-] Informática teórica
 - [-] C / C++
 - [-] Java
 - [-] Redes
- [-] Economy
- [-] Biology
- [-] Geography
- [-] Art and Literature
 - 1-2 Paintings
 - 1-3 Who wrote the Quixote?
 - Lloyd Puzzle
 - Sets
- [-] Music
- [-] Painting
 - Author of this painting - SI
 - Paintings (Imagen) SPAR
- [-] Literature

Preview

Information

Content

Display

Selection

Evaluation

Advanced

Branches

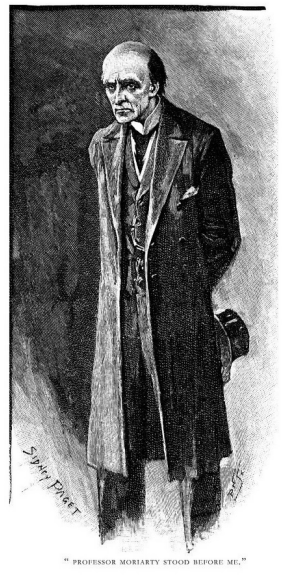
Analyzer

Sessions

Matematician SPARQL



Who is this matematician?



___ is a criminal mastermind whom Holmes describes as the "Napoleon of crime". Doyle lifted the phrase from a Scotland Yard inspector who was referring to Adam Worth, a real-life criminal mastermind and one of the individuals upon whom the character of ___ was based. The character was introduced primarily as a narrative device to enable Conan Doyle to kill Sherlock Holmes, and only featured in two of the Sherlock Holmes stories. However, in many adaptations, he has been given a greater prominence and treated as Holmes' archenemy.

Professor Moriarty

Evaluate

Button to mark this instance as "canceled"

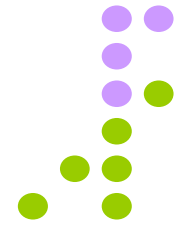
Save changes

Delete

Clone question

Create instances

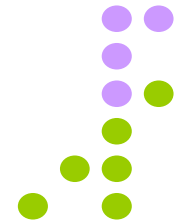
Estimate the number of invalid questions that will be posed.



- Estimate the probability of a single invalid question.
 - Sampling by browsing the preview. $\rightarrow p$
- Estimate the probability of having less than w invalid questions in a test of length n

$$p_w = Pr(X \leq w) = \sum_{i=0}^w \binom{n}{i} p^i (1-p)^{n-i}$$

Detect invalid questions already posed.




- MANUALLY: The student might report an incorrect instance, while taking the test.



Question number 7:

Which country has this flag?



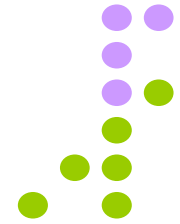
Romania / Chad

Comments

This flag is confusing, it might be from Romania or from Chad. Both flags are very similar

The image shows a screenshot of a test question interface. At the top, it says "Question number 7:". Below that is the question text "Which country has this flag?". In the center, there is a flag with three vertical stripes: blue on the left, yellow in the middle, and red on the right. Below the flag is a text input field containing the text "Romania / Chad". At the bottom, there is a "Comments" section with a text area containing the text "This flag is confusing, it might be from Romania or from Chad. Both flags are very similar".

Detect invalid questions already posed.



- MANUALLY: The teacher can review the student sessions and comments together and remove invalid instances.
- The instance is also removed from any other session it might appear.

The screenshot shows the 'Administrador' interface. On the left is a sidebar with a tree view of subjects including Filosofia, Maths, Algebra, Calculus, Geometry, Estadística, Arithmetica, Physics, Chemistry, Computers, and Entrenamiento. The 'Entrenamiento' section is expanded, showing various SPARQL question types like 'Flags SPARQL'. The main area displays a table of sessions with columns for 'FC', 'Date', 'Family name, Name', and 'Comments'. A table with one row is visible, showing a session from 2018/06/07 12:55:23 by 'Administrador' with a comment: 'This flag is confusing, it might be from Romania or from Chad. Both flags are very similar'. Below the table are buttons for 'Save changes', 'Delete', 'Clone question', and 'Create instances'.

FC	Date	Family name, Name	Comments
<input type="checkbox"/>	2018/06/07 12:55:23	Administrador .	This flag is confusing, it might be from Romania or from Chad. Both flags are very similar

Question number 7: Flags SPARQL (318194)
Which country has this flag?

Romania / Chad
 Chad

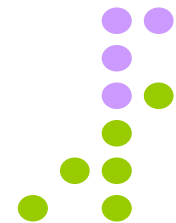
Comments :
This flag is confusing, it might be from Romania or from Chad. Both flags are very similar

Question number 8: Flags SPARQL (318194)
Which country has this flag?

Mali

Question number 9: Flags SPARQL (318194)
Which country has this flag?

Honduras

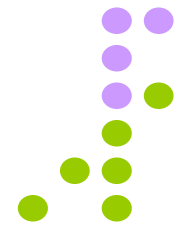


Detect invalid questions already posed.

- AUTOMATICALLY: Analyze correct response frequency, discrimination index and item-test point biserial correlation for all instances

The screenshot shows the Siette system interface. At the top, there is a navigation bar with 'Preguntas', 'Test', and 'Alumnos' tabs. Below this is a breadcrumb trail: 'Demo > Geografía > Banderas SPARQL'. A sidebar on the left lists various subjects and question types. The main area displays a table of question instances with columns for 'Instance Ids', 'Frecuencias absolutas', 'Índice de dificultad', 'Índice de discriminación', 'Correlación biserial-puntual', and 'Anulada'. A callout box labeled 'Instance quick preview' shows a question about the flag of Chile. Other callouts point to 'Re-assess button', 'Difficulty index', 'Discrimination index', and 'Frequency'.

Instance Ids	Frecuencias absolutas	Índice de dificultad	Índice de discriminación	Correlación biserial-puntual	Anulada
79E8625632428108	2	0.50	0.00		<input type="checkbox"/>
09F6D0F6560FA56403C00F62AFC8FBD2	3	0.33	0.00		<input type="checkbox"/>
FE55CBE4F8243484379DB319023329A7	1	1.00	1.00	1.00	<input type="checkbox"/>
C8000B76516FA86FC373198827110EEA	2	1.00	1.00		<input type="checkbox"/>
2D940E6AA3FD9DF863CD24595AE99A5F					<input type="checkbox"/>
734B995109FA7BD3843E3E053FE9CAD3				1.00	<input type="checkbox"/>
B8D10540B457AE2E84B70F96667D98E9				1.00	<input type="checkbox"/>
E40170BF53A3B678269DBDC381DE31D					<input type="checkbox"/>
3C6FD2C6239FAB41A8165A2A8687BA69					<input type="checkbox"/>
DFECBFE736523E87E10D51CF2E4DAC10					<input type="checkbox"/>
0E8FE1F3721B4B02FE26AB129B217B69	1	0.00	0.00	1.00	<input type="checkbox"/>
DE60777C2DE0632A060BD40C1B7A625A	3	0.33	0.00		<input type="checkbox"/>



Detect invalid questions already posed.

- Discrimination index

$$D = p_{4c} - p_{1c}$$

p_{4c} percentage of correct answer of the 25% of the student with the highest score.

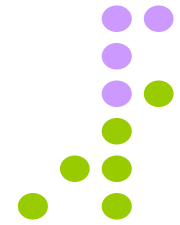
p_{1c} percentage of correct answer of the 25% of the student with the lowest score.

- Point-biserial correlation

$$r_{pb} = \frac{\bar{X}_1 - \bar{X}_T}{S_X} \sqrt{\frac{p}{q}}$$

Correlation between item and test score for all students

Estimate reliability after cancelling.



- Applying Spearman-Brown prophecy formula the new reliability can be calculated.
- The score confidence interval increases depending on the whole test reliability and the percentage of incorrect instances removed:

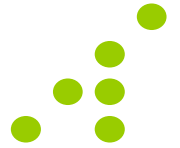
$\rho_{xx'}$	50%	40%	30%	20%	10%	5%	2.5%
0.70	1.240	1.179	1.125	1.078	1.037	1.018	1.007
0.75	1.265	1.195	1.136	1.085	1.040	1.019	1.008
0.80	1.219	1.213	1.147	1.091	1.043	1.021	1.008
0.85	1.319	1.231	1.159	1.098	1.045	1.022	1.009
0,90	1.348	1.250	1.170	1.104	1.048	1.023	1.009
0,95	1.380	1.270	1.183	1.111	1.051	1.025	1.010

Conclusions



- Generating short answer questions reduce the number of invalid questions.
- Invalid questions can be detected manually or semi-automatically
- Cancelling invalid questions and re-assess guarantees a valid assessment.
- the effect of using invalid questions can be measured and acceptance limits can be established.

Future work



- Application to a real case:
 - Generate question for plant recognition based on the sources: Dbpedia + Wikidata + EOL
- Use crawling to construct tables from highly reliable web repositories
- Assess higher order knowledge / skills (in the sense of Bloom's taxonomy)

Demo....



The image shows a web interface for Siette. At the top left is the Siette logo, with 'Siette' in red and green. To the right, there are links for 'Portal' (with a green checkmark icon) and 'Documentación' (with a document icon). Below the logo, there are language selection options: 'english' and 'deutsch', both with checked checkboxes. To the right of these are mode selection options: 'Modo individual' (with a person icon) and 'Modo colaborativo' (with a group icon). The main area contains a login form with two input fields: 'Nombre de usuario:' and 'Contraseña:'. Below the password field is a green 'Enviar' button with a right-pointing arrow. At the bottom of the form area, there are links for 'Olvidé mi contraseña' and 'Nuevo usuario', both with checked checkboxes. A small envelope icon is located at the bottom right of the form area. At the bottom of the page, there is a row of logos: the European Union flag, the Spanish Ministry of Science and Innovation, the Junta de Andalucía, the University of Málaga, and a logo for '(IA)² Investigación y Aplicaciones en Inteligencia Artificial'.

<https://www.siette.org>