

31/05/2022  
Palazzo Bo - Padova

**Prof. Marco Gori**  
Università degli Studi di Siena

**Marco Ernandes**  
expert.ai  
Hybrid Lang Technology Director

**Giovanni Angelini**  
expert.ai  
Hybrid Lang Technology R&D



# WEBCROW 2.0

## The human-machine crossword challenge



INTERNAL





# Possono le macchine risolvere cruciverba come noi?

Da oltre 100 anni i **cruciverba** rappresentano una delle sfide più intriganti per l'uomo...

...per via delle diverse accezioni delle parole e sfumature del linguaggio.

**Il linguaggio** è uno degli ambiti di applicazione dell'**intelligenza artificiale** più complessi e affascinanti

1	R	2	I	3	P	4	A
2	O	D	E	R			
3	S	E	R	A			
4	A	M	E	N			

*Orizzontali:* 1. Guai se l'onda mi varca o mi spezza 2. In Germania son acqua corrente 3. Ogni di quando il sole è morente 4. Così soglion le preci finir.

*Verticali:* 1. Sono un fiore di rara bellezza 2. Il medesimo in lingua latina 3. Quali frutti noi siamo, indovina! 4. Per la messe di là da venir.



- Orizzontali**
1. Guai se l'onda mi varca o mi spezza (8)
  2. In Germania son acqua corrente (7)
  3. Ogni di quando il sole è morente (7)
  4. Così soglion le preci finir. (7)

- Verticali**
1. Sono un fiore di rara bellezza (4)
  2. Il medesimo in lingua latina (4)
  3. Quali frutti noi siamo, indovina! (4)
  4. Per la messe di là da venir. (4)





# Comprendere, sapere, ragionare.



Per arrivare alla risposta corretta occorrono diversi passaggi logici e di analisi del linguaggio che iniziano con la **comprensione degli indizi**.



Per trovare il significato giusto servono inferenze e ragionamenti che sfruttano le **conoscenze** presenti in varie fonti.



Una combinazione di **tecnologie di intelligenza artificiale** per identificare le possibili parole, valutarle e completare così la griglia del cruciverba.



## Conoscenze



### Conoscenza lessicale/ontologica

Il **Knowledge Graph** di expert.ai è il più ricco database concettuale per la comprensione del linguaggio presente sul mercato.

### Conoscenza enciclopedica e Sapere comune

Il **Web** rappresenta una raccolta vasta e continuamente aggiornata del sapere umano.



### Conoscenza enigmistica

**Cruciverba** già risolti ed interpretazione di giochi di parole.



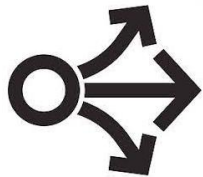
# Processo di risoluzione

## 1. “Clue Answering” → Rispondere alle definizioni



- Diversi “sistemi esperti” in collaborazione, in base alla definizione, producono una lista di possibili risposte
- Altri esperti “analizzano” la definizione e “valutano” (o arricchiscono) la lista prodotta
- Duplice obiettivo:
  - Cercare di avere il più spesso possibile la risposta corrette tra le candidate
  - Avere la risposta corretta con il più alto rank e score (probabilità) possibile

## 2. “Belief Propagation” → Propagazione delle probabilità



- Sfruttiamo lo schema (e i vincoli che impone) per ricalcolare le probabilità delle risposte candidate

## 3. “Grid filling” → Riempimento della griglia

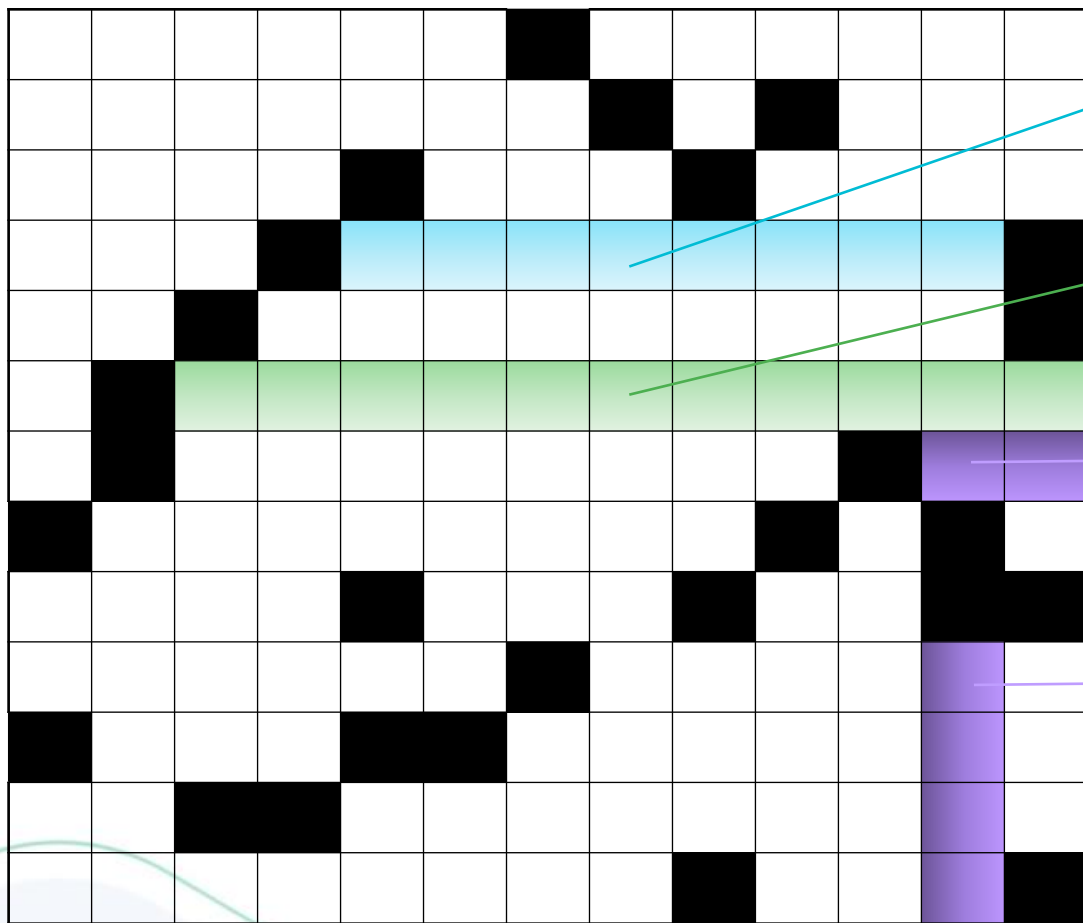
- Effettivo inserimento delle parole/lettere nello schema.





Clue Answering

# Analisi delle definizioni



[20 oriz.] Inquieti e nervosi

[24 oriz.] Gregorio, campione di nuoto

[27 oriz.] Un po' ebbro

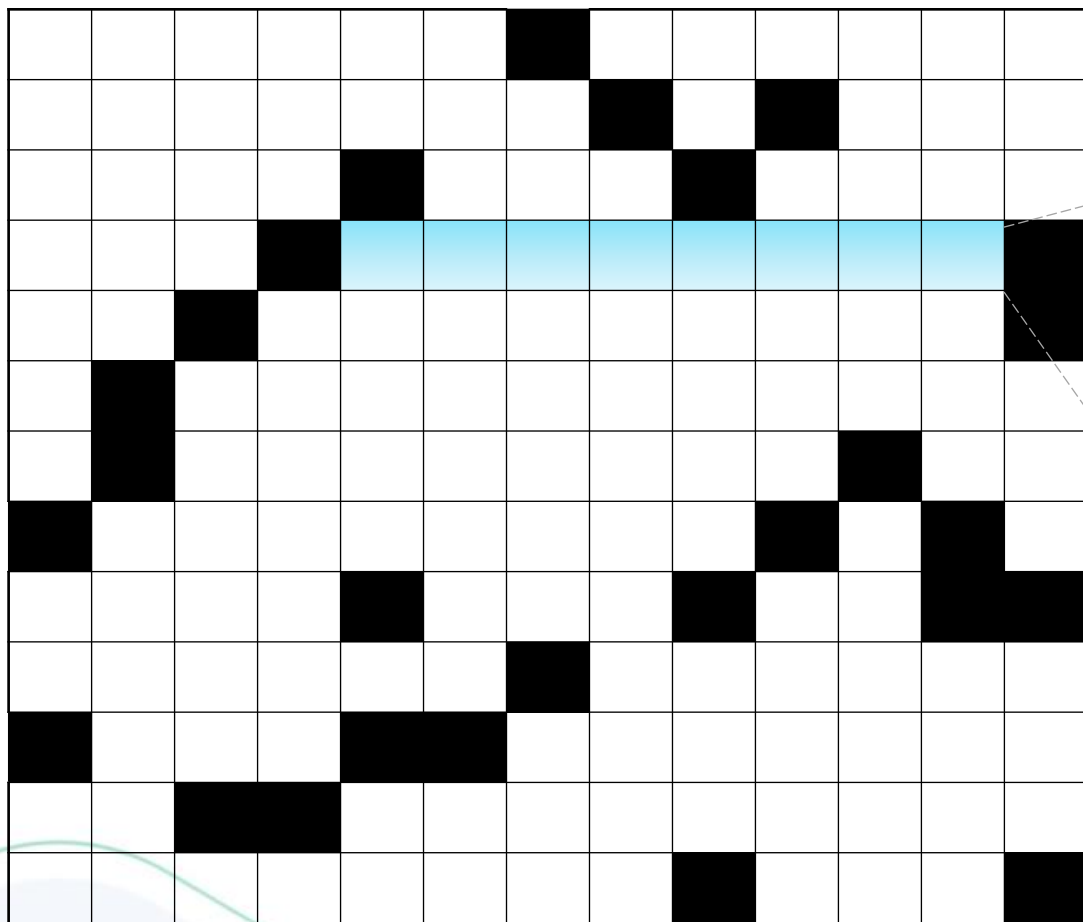
[36 vert.] Vetrinetta da museo





Clue Answering

# Conoscenza lessicale/ontologica



[20 oriz.] Inquieti e nervosi

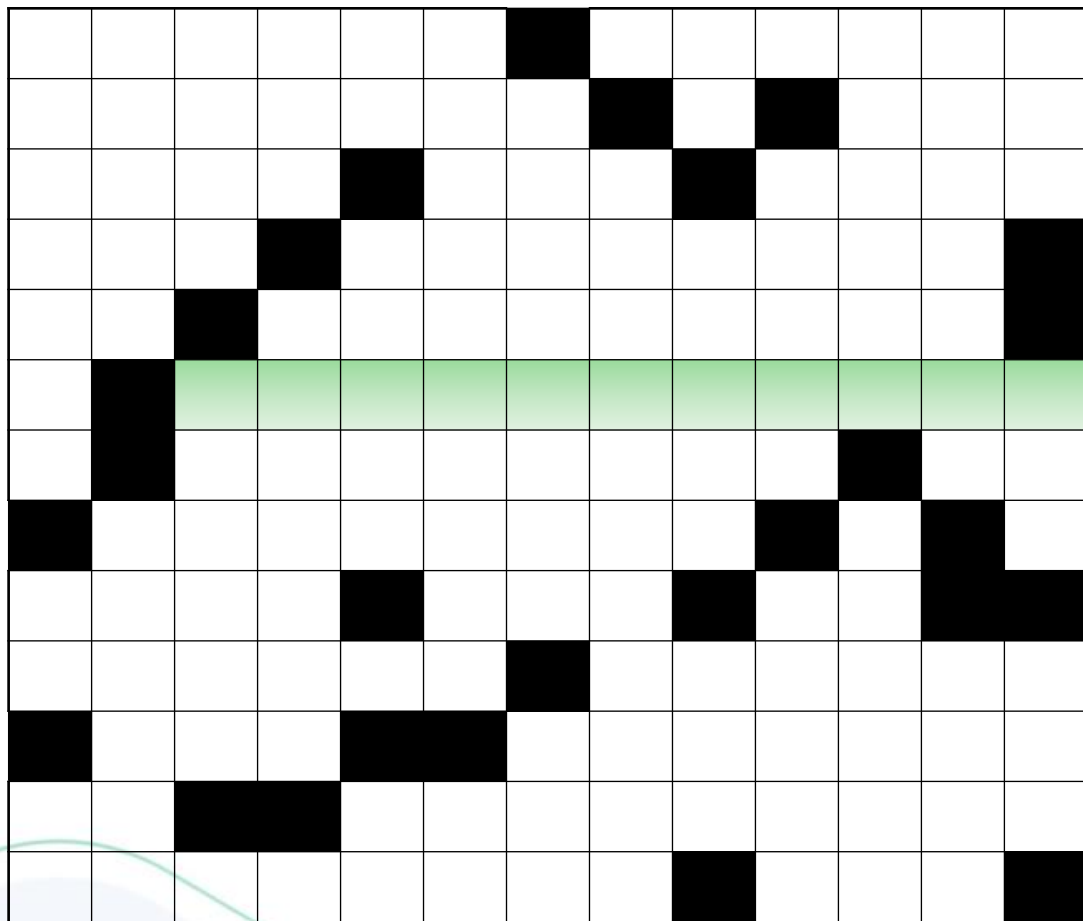
RANK	CANDIDATA	SCORE
1	IRRITATI	20,7%
2	STRANITI	14,3%
3	ALTERATI	9,16%
4	SDEGNATI	9,06%
5	ISTERICI	5,22%
...	...	...
...	...	...
...	...	...
24230	GENETICI	0,0000000006%
...	...	...
...	...	...



Clue Answering

# Conoscenza enciclopedica

## [24 oriz.] Gregorio, campione di nuoto



[Gregorio Paltrinieri - Wikipedia](#)

[https://it.wikipedia.org > wiki > Greg...](https://it.wikipedia.org/wiki/Gregorio_Paltrinieri)

Gregorio Paltrinieri, detto Greg (Carpi, 5 settembre 1994), è un nuotatore italiano specializzato nello ...

[https://www.vanityfair.it > ... > Nuoto](https://www.vanityfair.it/.../Nuoto)

Il *campione olimpico* **Gregorio Paltrinieri** ha messo insieme tutte le sue passioni: *nuoto*, ... Un circuito di gare di *nuoto* (per agonisti, ..

RANK	CANDIDATA	SCORE
1	PALTRINIERI	35%
2	PIETRANGELI	2,5%
...	...	...



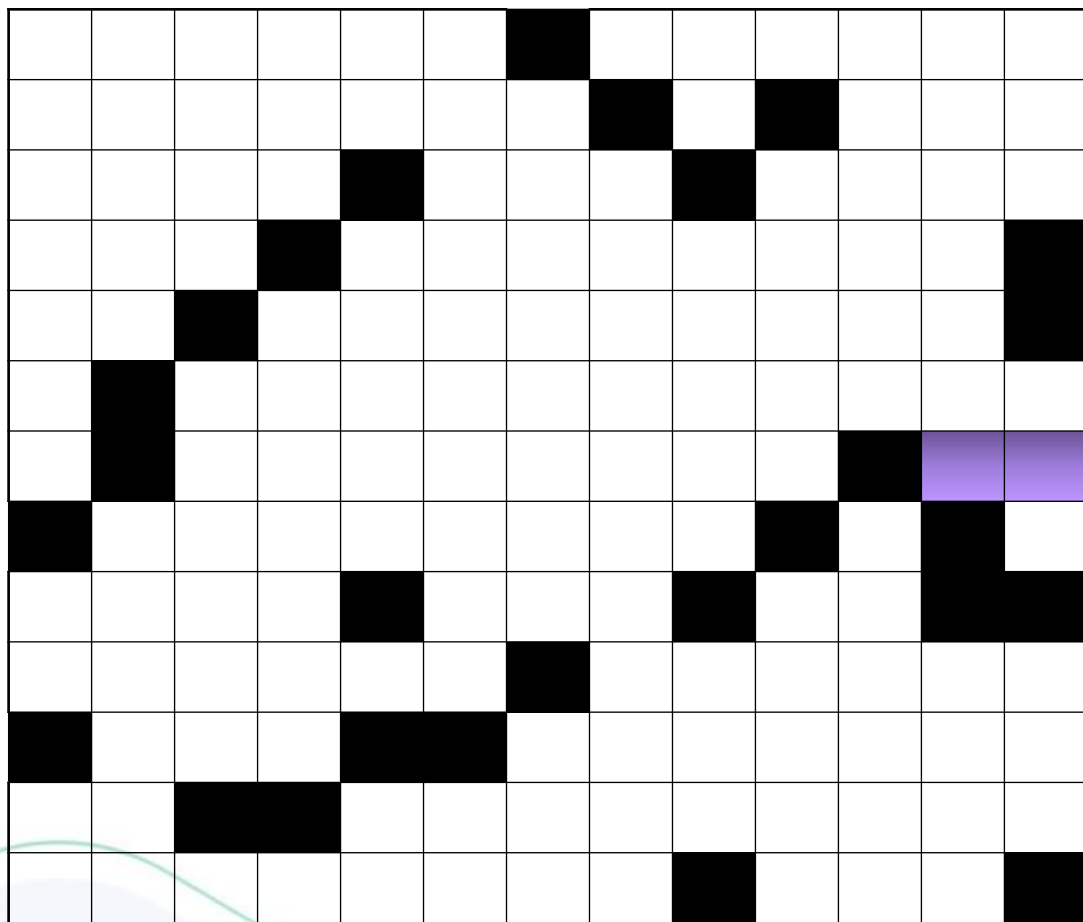




Clue Answering

# Conoscenza enigmistica

[27 oriz.] Un po' **ebbro**



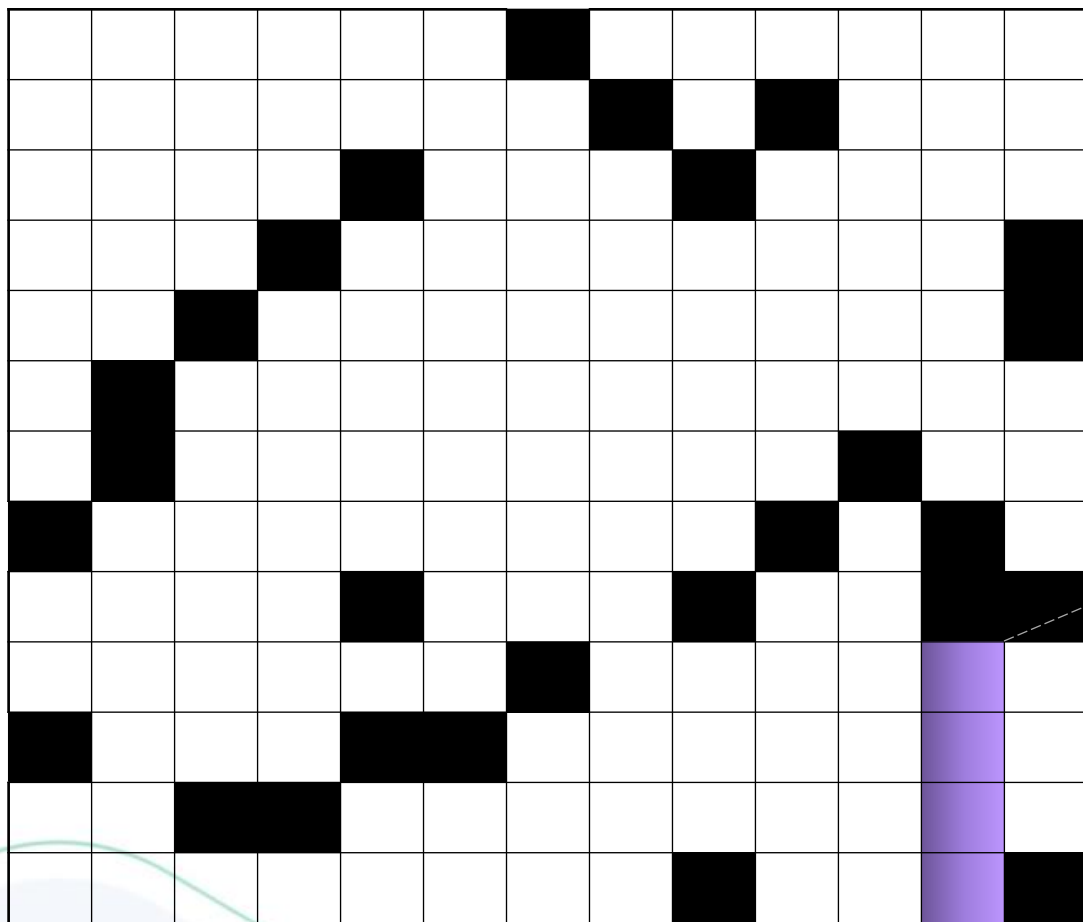
### Modulo ruled based

RANK	CANDIDATA	SCORE
1	EB	100%



Clue Answering

# Conoscenza enigmistica



[36 vert.] Vetrinetta da museo

«Una custodia per reliquie» -> TECA

«Un astuccio da museo» -> TECA

«Vetrina per reliquie» -> TECA

RANK	CANDIDATA	SCORE
1	TECA	89,99%
2	RARA	9,4%
3	ARTE	3,4%
...	...	...
2531	VIVA	0,0000000001%
...	...	...





Belief Propagation

## Propagazione Probabilità sui vincoli



### • Informazioni a disposizione:

- Liste di risposte **candidate**
- **Vincoli** dello schema
  - Le lettere negli incroci tra le parole devono essere identiche

### • Ipotesi:

- La risposta corretta sarà quella che massimizza congiuntamente la confidenza rispetto alla definizione e rispetto ai vincoli della griglia

### • Due obiettivi principali:

- Aumentare la *“confidenza”* (score) della parola corretta
- Agevolare il processo di inserimento successive, rendendolo più robusto alle *“risposte mancanti”*

C	A	R	N	A	P		B	O	H	E	M	E
R	O	A	D	M	A	P		K		D	I	N
E	S	T	E		R	E	C		A	O	N	I
S	T	O		S	T	R	A	N	I	T	I	
I	A		B	A	I	O	N	E	T	T	A	
M		P	A	L	T	R	I	N	I	E	R	I
A		A	R	I	A	A	R	I	A		E	B
	D	I	E	S	I	R	A	E		R		M
K	R	O	L		V	E	N		A	E		
F	O	L	L	I	A		D	O	S	A	T	A
	G	O	A			M	A	N	A	G	E	R
E	H			G	R	E	G	O	R	A	C	I
M	E	D	I	O	C	R	I		I	N	A	

7 Orizzontale - 6 lettere

*“Corrente artistica francese... pucciniana”*

RANK	CANDIDATA	SCORE
1	ORFICA	15,9%
2	INGRES	11,1%
3	FELICE	7,47%
4	MADAMA	5,75%
5	GRIEUX	2,73%
6	DANTON	2,51%
7	ANTICA	2,41%
...	...	...
23	BOHEME	1,62%
24	LIRICA	1,44%
...	...	...

RANK	CANDIDATA	SCORE
1	CAROVANIERI	4,7%
2	SENZAUSCITA	3,9%
3	ABBANDONATI	2,7%
4	ALTERNATIVI	2,5%
...	...	...

17 Vertical - 11 lettere *“Lo sono i quattrozampe a zonzo per le strade”*

INTERNAL



Belief Propagation

## Propagazione Probabilità sui vincoli

Risposte verticali

C	A	R	N	A	P		B	O	H	E	M	E
R	O	A	D	M	A	P		K		D	I	N
E	S	T	E		R	E	C		A	O	N	I
S	T	O		S	T	R	A	N	I	T	I	
I	A		B	A	I	O	N	E	T	T	A	
M		P	A	L	T	R	I	N	I	E	R	I
A		A	R	I	A	A	R	I	A		E	B
	D	I	E	S	I	R	A	E		R		M
K	R	O	L		V	E	N		A	E		
F	O	L	L	I	A		D	O	S	A	T	A
	G	O	A			M	A	N	A	G	E	R
E	H			G	R	E	G	O	R	A	C	I
M	E	D	I	O	C	R	I		I	N	A	

Effetto della Propagazione

Risposta orizzontale

TOP 1
TOP 5
TOP 10
TOP 50
TOP 100
TOP 500
TOP 1000
TOP 10000
PRESENT
ABSENT





Belief Propagation

## Propagazione Probabilità sui vincoli

Risposte orizzontali

Risposte verticali

TOP 1
TOP 5
TOP 10
TOP 50
TOP 100
TOP 500
TOP 1000
TOP 10000
PRESENT
ABSENT

C	A	R	N	A	P		B	O	H	C	A	R	N	A	P		B	O	H	E	M	E	N	A	P		B	O	H	E	M	E		
R	O	A	D	M	A	P		K		R	O	A	D	M	A	P		K		D	I	N	D	M	A	p		K		D	I	N		
E	S	T	E		R	E	C		A	E	S	T	E		R	E	C		A	O	N	I	E		R	E	C		A	O	N	I		
S	T	O		S	T	R	A	N	I	S	T	O		S	T	R	A	N	I	T	I			S	T	R	A	N	I	T	I			
I	A		B	A	I	O	N	E	T	I	A		B	A	I	O	N	E	T	T	A			B	A	I	O	N	E	T	T	A		
M		P	A	L	T	R	I	N	I	M		P	A	L	T	R	I	N	I	E	R	I		A	L	T	R	I	N	I	E	R	I	
A		A	R	I	A	A	R	I	A	A		A	R	I	A	A	R	I	A		E	B		R	I	A	A	R	I	A		E	B	
	D	I	E	S	I	R	A	E			D	I	E	S	I	R	A	E		R		M		E	S	I	R	A	E		R		M	
K	R	O	L		V	E	N		A	K	R	O	L		V	E	N		A	E				L		V	E	N		A	E			
F	O	L	L	I	A		D	O	S	F	O	L	L	I	A		D	O	S	A	T	A		L	I	A		D	O	S	A	T	A	
	G	O	A			M	A	N	A		G	O	A			M	A	N	A	G	E	R		A			M	A	N	A	G	E	R	
E	H			G	R	E	G	O	R	E	H			G	R	E	G	O	R	A	C	I			G	R	E	G	O	R	A	C	I	
M	E	D	I	O	C	R	I			M	E	D	I	O	C	R	I								I	O	C	R	I					



*Belief Propagation*

## Propagazione Probabilità sui vincoli

Risposte orizzontali

Risposte verticali

TOP 1
TOP 5
TOP 10
TOP 50
TOP 100
TOP 500
TOP 1000
TOP 10000
PRESENT
ABSENT

C	A	R	N	A	P		B	O	H	E	M	E
R	O	A	D	M	A	P		K		D	I	N
E	S	T	E		R	E	C		A	O	N	I
S	T	O		S	T	R	A	N	I	T	I	
I	A		B	A	I	O	N	E	T	T	A	
M		P	A	L	T	R	I	N	I	E	R	I
A		A	R	I	A	A	R	I	A		E	B
	D	I	E	S	I	R	A	E		R		M
K	R	O	L		V	E	N		A	E		
F	O	L	L	I	A		D	O	S	A	T	A
	G	O	A			M	A	N	A	G	E	R
E	H			G	R	E	G	O	R	A	C	I
M	E	D	I	O	C	R	I		I	N	A	





Grid Filling

## Riempimento dello schema

### Ricerca della migliore combinazione di parole inserite

1. Create Lists
Fill Grid

1	2		3	
			4	
5				
			6	7
8	9	10		
11				
	12			

**Across:**

1. Il primo all'opera.
4. In mezzo all'addome.
5. Lo tradisce il bugiardo.
6. Titolo di Juan Carlos.
8. La stella del cane
11. Organiz. Nazioni Unite.
12. Un capitano sotto i mari.

**Down:**

2. Lo erano i Re Magi.
3. Il naso li distingue.
5. Lo aprì Pandora.
7. Dio dei venti.
9. Locale inglese.
10. Una strada di Parigi.

### Inserimento caratteri + "validazione" per parola

C	A	R	N	A	P		B	O	H	E	M	E
R	O	A	D	M	A	P		K		D	I	N
E	S	T	E		R	E	C		A	O	N	I
S	T	O		S	T	R	A	N	I	T	I	
I	A		B	A	I	O	N	E	T	T	A	
M		P	A	L	T	R	I	N	I	E	R	I
A		A	R	I	A	A	R	I	A		E	B
	D	I	E	S	I	R	A	E		R		M
K	R	O	L		V	E	N		A	E		
F	O	L	L	I	A		D	O	S	A	T	A
	G	O	A			M	A	N	A	G	E	R
E	H			G	R	E	G	O	R	A	C	I
M	E	D	I	O	C	R	I		I	N	A	



## Grid Filling

# Inserimento caratteri e "validazione" per parola

C	A	R	N	A	P		B	O	H	E	M	E
R	O	A	D	M	A	P		K		D	I	N
E	S	T	E		R	E	C		A	O	N	I
S	T	O		S	T	R	A	N	I	T	I	
I	A		B	A	I	O	N	E	T	T	A	
M		P	A	L	T	R	I	N	I	E	R	I
A		A	R	I	A	A	R	I	A		E	B
	D	I	E	S	I	R	A	E		R		M
K	R	O	L		V	E	N		A	E		
F	O	L	L	I	A		D	O	S	A	T	A
	G	O	A			M	A	N	A	G	E	R
E	H			G	R	E	G	O	R	A	C	I
M	E	D	I	O	C	R	I		I	N	A	

C	A	R	N	A	P		B	O	H	E	M	E
R	O	A	D	M	A	P		K		D	I	N
E	S	T	E		R	E	C		A	O	N	I
S	T	O		S	T	R	A	N	I	T	I	
I	A		B	A	I	O	N	E	T	T	A	
M		P	A	L	T	R	I	N	I	E	R	I
A		A	R	I	A	A	R	I	A		E	B
	D	I	E	S	I	R	A	E		R		M
K	R	O	L		V	E	N		A	E		
F	O	L	L	I	A		D	O	S	A	T	A
	G	O	A			M	A	N	A	G	E	R
E	H			G	R	E	G	O	R	A	C	I
M	E	D	I	O	C	R	I		I	N	A	

C	A	R	N	A	P		B	O	H	E	M	E
R	O	A	D	M	A	P		K		D	I	N
E	S	T	E		R	E	C		A	O	N	I
S	T	O		S	T	R	A	N	I	T	I	
I	A		B	A	I	O	N	E	T	T	A	
M		P	A	L	T	R	I	N	I	E	R	I
A		A	R	I	A	A	R	I	A		E	B
	D	I	E	S	I	R	A	E		R		M
K	R	O	L		V	E	N		A	E		
F	O	L	L	I	A		D	O	S	A	T	A
	G	O	A			M	A	N	A	G	E	R
E	H			G	R	E	G	O	R	A	C	I
M	E	D	I	O	C	R	I		I	N	A	







# Competizione WCCI 2022 | Padova, 20/07/2022

- Due diverse **lingue**: ITALIANO e INGLESE (American Crosswords)
- **Autori/editori** degli schemi di gara:
  - SPREA SUDOKU EDIZIONI per l'Italiano
  - AVCX per le American Crosswords
- **Partecipanti:**
  - Tutti coloro che si registreranno sulla piattaforma ("Crossword Arena") e si collegheranno tramite proprio dispositivo
  - "onsite" chiunque vorrà partecipare dalla sede della conferenza (in una apposita aula)
  - "online" chiunque vorrà partecipare remotamente
- **Premi:**
  - Vi saranno premi per la gran parte dei partecipanti
  - Premi speciali andranno a vincitori e ai migliori risolutori.
  - Sono previsti abbonamenti e gadget.

Cruciverba italiani e americani sono un po' diversi



INTERNAL





## Competizione WCCI 2022 | Padova, 20/07/2022

### • **Graduatoria:**

- Per ogni lingua saranno previsti più schemi:
  - Uno schema di prova (facile) per prendere dimestichezza con la piattaforma
  - Due schemi (difficoltà media e alta) che contribuiranno al punteggio
- Il punteggio finale sarà dato dalla somma dei punti ottenuti sui due schemi di gara
- Il sistema di punteggio si baserà sulla qualità della soluzione, con ulteriore bonus per qualità (prioritaria) e velocità
- Verranno stilate graduatorie per l'Italiano e l'Inglese

### • Piattaforma di gioco **“Crossword Arena”**

- Sviluppata per permettere la risoluzione contemporanea dei partecipanti
- Pensata per essere “facile” sia da dispositivi mobili (App e Browser) che PC (Browser)
- Permetterà di monitorare in tempo reale l'andamento della gara
- Calcolerà in modo automatico la graduatoria
- Permetterà di esplorare le soluzioni dei partecipanti e di WebCrow
- Mostrerà il regolamento di gara e permetterà di ricevere le notifiche degli organizzatori

### • **Chi vincerà?**



Associazione  
Italiana per  
l'Intelligenza  
Artificiale



# WebCrow Team

**Prof. Marco Gori**

Università di Siena

**Prof. Marco Maggini**

Università di Siena

**Tommaso Iaquina**

Università di Siena

con la collaborazione di studenti del corso di laurea

**Marco Ernandes**

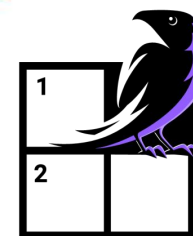
expert.ai

**Giovanni Angelini**

expert.ai

**Andrea Zugarini**

expert.ai



**WEBCROW 2.0**



Associazione  
Italiana per  
l'Intelligenza  
Artificiale



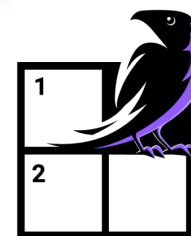
# Thank You.



<http://webcrow.diism.unisi.it>



[webcrow@unisi.it](mailto:webcrow@unisi.it)



**WEBCROW** 2.0