

Milestone 16 –

The White paper with best practices and recommendations about the NLP and AI use on public procurement data has been produced

Grant Agreement Nº	INEA/CEF/ICT/A2020/2373713
Action Title (Art. 1 of G.A.)	Open Harmonized and Enriched Procurement Data Platform (nextProcurement)
Action number (Art. 1 of the G.A.)	2020-EN-IA-0255

Table of Contents

1	Exec	cutive Summary3
2	Intro	oduction4
	2.1	Document-based public tendering4
	2.2	Who should read this whitepaper5
3	Next	tProcurement NLP and AI Best Practices and Recommendations7
	3.1 3.1.1 3.1.2 3.1.3	Promote complete and optimal accessibility to tender documents
	3.2 3.2.1 3.2.2 3.2.3 3.3 3.3.1	COGS calculation using NLPs and AI
	3.3.2 3.3.3	Example of good practice
	3.4.1 3.4.2 3.4.3	Example of malpractice
	3.5 3.5.1 3.5.2 3.5.3	Extraction of Special Conditions Using NLP and Al
	3.6 3.6.1 3.6.2 3.6.3	Extraction of Award Criteria using NLP and Al
	3.7 3.7.1 3.7.2 3.7.3	Extracting Tenderers Using NLP and AI21Example of malpractice21Example of good practice22Recommendations25
	3.8 3.8.1 3.8.2 3.8.3	Object search, including semantic search25Example of malpractice26Example of good practice27Recommendations28
4	Othe	er recommendations not related to NLP and Al29
	4.1 4.1.1 4.1.2	Better data integration from Connected Platforms with PLACE

Healt	th and D		CEF Telecom
Gran	t Agree 4.1.3	Recommendations	
4	.2	Improving data from contracting bodies	
	4.2.1	Example of malpractice	
	4.2.2	Example of good practice	
	4.2.3	Recommendations	
4	.3	Improve the quality of metadata in general	35
	4.3.1	Example of malpractice	
	4.3.2	Example of good practice Recommendations	
4	.4	Statistics/Indicators at the contracting agency level	
	4.4.1	Example of malpractice	
	4.4.2	Example of good practice	
	4.4.3	Recommendations	
4	.5	Tenders and statistics at SME level	41
	4.5.1	Example of malpractice	
	4.5.2	Recommendations	
4	.6	Implementing APIs for Data Access	43
	4.6.1	Example of malpractice	
	4.6.2	Example of good practice	
_	4.6.3	Recommendations	
5	Othe	er recommendations to make the most of NLP and AI technologies	47
5	.1	Repository of processed tender documents for AI training	47
5	.2	Repository of pre-trained AI models for public bidding	47
5	.3	Generation of tender models based on previous tenders	47
5	.4	Risk analysis, detection of irregularities and discriminatory clauses .	47
5	.5	Automatic generation of bid summaries	47
5	.6	Development of virtual assistants for the tender	48
5	.7	Predictive analysis of concurrency and trends in hiring	48
5	.8	Development of recommendation systems for SMEs	48
5	.9	Development of recommendation systems for contracting bodies	48
5	.10	Automatic Frequently Asked Questions (FAQs) Generation for Tend	ers48
6	Con	clusions	49
6	.1	Main findings	49
6	.2	Summary of key recommendations for the future.	50

1 Executive Summary

Public procurement is a key pillar of the European economy, accounting for between 10% and 20% of the GDP of EU countries. Despite its importance, the public procurement market continues to face significant challenges, such as barriers to access for new bidders, lack of competition, opacity in award processes, and deficiencies in data management and publication.

In recent years, the advent of advanced technologies such as Natural Language Processing (NLP) and Artificial Intelligence (AI) has opened up new opportunities to transform public procurement. These technologies can automate document analysis, improve accessibility to information, optimise the search for opportunities, and detect irregularities in bidding processes. However, their adoption in the public sector remains limited because until recently these technologies were not within global reach, which has been the case for just two years.

This document aims to provide a reasoned set of **best practices and recommendations on the use of NLP and AI in public procurement based on the experiences of the implementation of NextProcurement**, which has focused on the extraction and generation of complementary information from tender documents (specifications and minutes).

In addition, best practices and recommendations not directly related to NLP and AI are also included, but which, due to the developments made in NextProcurement, can provide a complement to the improvement of public procurement.

Finally, a series of other recommendations on the use of NLP and AI are included, which have not been developed in NextProcurement because they are not within the scope of action, but which have been detected as possible improvements and future applications.

2 Introduction

Despite efforts to modernise and digitise tendering processes, significant challenges remain in terms of accessibility, interoperability, transparency and efficiency.

When an administration wants to know the characteristics of tenders similar to one that is about to be published, it is not optimal to review one by one, opening the administrative specifications and manually looking for the relevant information. The same happens with companies that analyse their market in public tenders or look for opportunities aligned with their activity. In both cases, having structured and quality data is essential for agile and accurate analysis.

The use of advanced technologies, such as Natural Language Processing (NLP) and Artificial Intelligence (AI), can transform the way public procurement data is managed and analysed. These tools allow you to automate the extraction of information, improve the search and classification of tenders and facilitate the analysis of market trends. Through its application, it is possible to facilitate the access of companies, especially SMEs, to procurement opportunities, facilitate the preparation and drafting of new tenders to contracting bodies, increase competition and ensure a more efficient use of public resources.

However, there are still challenges to be addressed related to the very mechanics of public tenders that are based on documents and the application of NLP and AI technologies.

2.1 Document-based public tendering

NextProcurement focuses on generating rich metadata from tender documents, in order to extend the quantity and quality of existing metadata.

As they are legally valid documents (in reality it is the documents that have legal validity and not the metadata with which publications are made on the tender portals), the format used is usually PDF, which was created to maintain the appearance of the documents, ensuring that they look the same on any device.

But that same rigidity makes it complicated for machines, since extracting data from files in PDF format is still a great challenge, even more so when the documents are text images, which forces the use of Optical Character Recognition (OCR) programs to convert them into editable data.

OCR has been around since the 70s, and while it has improved over time, it still has limitations. Their approach is based on recognising pixel patterns, which work well with clear, well-scanned text. But when documents have complex layouts or poor image quality, errors are common.

The difficulty increases with old documents or with complex content such as tables or formulas. And with the application of AI to tender documents, the problem has also been seen.

In the most current context of AI, startups are also trying to make progress in this field. For example, the French company Mistral recently launched its Mistral OCR model, designed to deal with complex documents. However, the first tests have not been entirely positive. The model still makes mistakes when processing tables, repeating data, or confusing numbers.

Therefore, the current context is that, although progress is evident, LLMs still have certain risks. Errors in the interpretation of tables or the invention of data can go unnoticed, which is a threat in important documents such as those of public tenders. Automated tools can speed up the work, but they need to be carefully reviewed to avoid mistakes that could have serious consequences.

2.2 Who should read this whitepaper?

This whitepaper is aimed at all actors involved in public procurement and in the modernisation of its processes through advanced technologies. In particular, its content is relevant to:

1. Public administrations and contracting bodies

- Responsible for designing, managing and publishing tenders that wish to improve the efficiency and transparency of their processes.
- Government entities looking to implement AI and NLP-based solutions to optimise recruitment data management.

2. Supplier companies and bidders

- SMEs and large companies interested in participating in public tenders and improving their access to procurement opportunities.
- Technology companies and consultancies specialising in automation and data analysis solutions applied to public procurement.

3. Regulators and watchdogs

- Entities in charge of supervising transparency and regulatory compliance in public procurement.
- Bodies that want to adopt AI-based methodologies for the detection of fraud or bad practices in awards.
- 4. Data Researchers and Analysts

- Academics and research centres studying the digitalisation and modernisation of public procurement.
- Data analysts interested in exploring new applications of NLP and Al in the public sector.

5. Civil society and transparency organizations

- NGOs and citizen groups that promote open access to public information and accountability in government contracting.
- Journalists and media outlets investigating the use of public funds and the dynamics of the recruitment market.

3 NextProcurement NLP and AI Best Practices and Recommendations

This section presents best practices and general recommendations based on the context of NextProcurement project implementation related to NLP and AI.

3.1 Promote complete and optimal accessibility to tender documents

Public procurement is based on documents such as specifications, minutes and other reports. Therefore, access to the information contained in these documents is essential to ensure transparency and equal opportunities in bidding processes.

However, having information only in documents is not optimal for automated searches, comparisons and analysis. For this reason, public procurement platforms are based on metadata, which allows tenders to be searched and filtered by different criteria. However, not all relevant data is structured in metadata, as is the case with the complete list of bidders in a tender.

In addition, in most contracting platforms, metadata is entered manually, which can lead to errors, omissions, or variability in the information. In many cases, only a reference to the document is introduced instead of extracting and structuring the information in a complete way. The lack of automated mechanisms to validate this data or their poor implementation causes inconsistencies and reduces the quality of the records.

An additional problem is that contracting authorities, drafters and document managers are not always aware of these limitations, since, from their perspective, the legal validity of the tender lies in the document, not in the metadata. However, if you want to advance in intelligent systems that use NLP (Natural Language Processing) and AI (Artificial Intelligence), it is essential to raise awareness of these issues and improve the production of documents and metadata.

3.1.1 Example of malpractice

Some common document issues include:

• Scanned PDF documents, making them difficult to process automatically.



• Documents that require unnecessary authentication, controls, or permissions, making it difficult to access information.

DESCAR	GAR ARCHIVO					
Si solo de:	ea descargar el archivo solic	itado pulse Descar	rgar archivo			
Descarg	er archivo					
Bienvenie	o al servicio de alertas de	el Perfil de Contr	atante de Euska	li.		
DAR DE	ALTA SUSCRIPCIÓN					
Si desea r	cibir información sobre mod	lificaciones que se	produzcan en el e:	pediente de contrata	ción actual, introduzca	ı su email.
Email:						
Introduz	a las letras siguientes:		_64a	71		
Ao	ptar					

• Erroneous URLs or broken links, which prevent the download of documents.

ک 🙃 contractacio	publica.cat/ca/inici			■ ☆
C Plataform	na de Serveis	de Contractació Púb	lica	
		1	\odot	۵
		3		
Inici	Presentació	Perfils de contractant	Empreses licitadores	Subscripc
La Nova El dia 13 de no actuacio Departam	Plataforma d ovembre de 2024 l'Es ons innovadores i les ient d'Economia i Fin Qua	e serveis de contract scola d'Administració Pública de s bones pràctiques en gestió pút ances ha estat guardonat amb e litat pel projecte de la Nova Plat	ació pública guanya (Catalunya ha lliurat els VII Premi llica que duen a terme les admin el 1r Premi de la categoria de Tra aforma de serveis de contractaci	el 1r Premi a s Alfons Ortuño e istracions públiqu nsparència, Bon C ó pública
		Cer	cadors	
c	U		Ú141	DCCD O

• **Relevant information only accessible within documents** and not structured in metadata (e.g. the list of all bidders).

3.1.2 Example of good practice

At NextProcurement, a system has been developed that allows documents to be converted into actionable texts, which can be analysed using NLP and AI to extract rich metadata that complements those currently available on procurement platforms.

This system includes:

- **Conversion of documents to text**, allowing their processing by analysis and information extraction tools.
- **Generation of rich metadata** that complements and enhances current metadata, facilitating analysis without the need to open each document individually.

3.1.3 Recommendations

- Incorporate NLP and AI from document creation, assessing what changes in format and structure can improve your automated processing.
- Automate metadata completion using AI, reducing errors and improving the quality of structured information.

3.2 COGS calculation using NPLs and AI

In the case of CPV (Common Procurement Vocabulary), there is a specific metadata that allows tenders to be classified uniformly. However, in some cases, this metadata is incomplete, misassigned, or not representative enough, limiting its usefulness.

It should be noted that the CPV classification has more than 9,000 codes, which makes it difficult to select both for contracting bodies, which must assign the correct code, and for companies, which are looking for opportunities based on these codes. An erroneous or inaccurate classification can make a tender go unnoticed by those who might really be interested in it.

For example, for the concept of "traffic lights" there are 3 CPVs that contain the word traffic light:

- 34996100 Traffic lights
- 45316212 Installation of traffic lights
- 50232000 Maintenance services for street lighting installations and traffic lights

and in the tenders there are CPVs as varied as:

- 50232000 Maintenance services for street lighting installations and traffic lights
- 34996100 Traffic lights

- 34923000 Road traffic control equipment
- 34996000 Control, safety or road marking equipment
- 50232100 Street lighting maintenance services
- 45316212 Installation of traffic lights
- 45233141 Road maintenance work
- 71356000 Technical Services

Another key aspect is that many platforms do not allow text searches on the object of the tender, so the only data available to identify its content is the CPV. This makes the quality and accuracy of this metadata critical to ensure an effective search and prevent relevant tenders from being excluded from the analysis.

3.2.1 Example of malpractice

Some common problems related to CPV metadata are as follows:

• Empty CPV: No code has been assigned, making it difficult to find.

Subject 🗳	CPVs
Contrato de obras de emergencia de instalación de semáforos dañados por la DANA.	-
Servicio de reparación de dos semáforos dañados por actos vandálicos.	-
Reparació semàfor policia	-
Instalación de semáforos para paso de peatones en la Calle Bajada de la Encina, intersección con Calle Los Infantes	-
Lloguer de semàfor senyalització	-
Asistencia fase implantación medidas de mejora de eficiencia energética para la gestión integral del alumbrado público, semáforos y dispositivos Smart City municipal de Santa Pola.	-
Contrato menor de suministros para la reparación y modernización de los semáforos de Ronda Sur en el cruce con carretera de Denia.	1
Contrato menor de servicios de redacción de proyecto de instalación de semáforo y pasos de peatones inteligentes dentro del plan v 2024, Boecillo	-

• Limiting single CPV: when the tender covers multiple concepts, but only one code has been assigned instead of several suitable ones.



• **Too many CPVs**: multiple codes have been assigned so it is not clear what the main concept of the tender is.

Suministros energéticos y servicios de mantenimiento con garantía total con inversión de las instalaciones en el centro consumidor de energía integrado por los edficios y equipamientos incluidos en los inmuebles municipales, alumbrado público, pasos de peatones, señales semafóricas reguladoras de la circulación y fuentes del Ayuntamiento d

Original	Generado
CPVs	CPVs
50232000,71000000,9300000,45300000,45331000,50232100,51100000,9000000,5053100	71314000
CPVS detalle 50232000 - Servicios de mantenimiento de instalaciones de alumbrado público y	CPVS detalle 71314000 - Servicios de energía y servicios conexos
semáforos	
71000000 - Servicios de arquitectura, construcción, ingeniería e inspección 9300000 - Electricidad, calefacción, energías solar y nuclear	Buscar similar por CPV
45300000 - Trabajos de instalación en edificios	
45331000 - Trabajos de instalación de calefacción, ventilación y aire acondicionado	
51100000 - Servicios de instalación de equipos eléctrico y mecánico	
9000000 - Derivados del petróleo, combustibles, electricidad y otras fuentes de	
energía	
50531000 - Servicios de reparación y mantenimiento de maquinaria no electrica 50700000 - Servicios de reparación y mantenimiento de equinos de edificios	
50710000 - Servicios de reparación y mantenimiento de equipos eléctricos y	
mecànicos de edificios	
50500000 - Servicios de reparacion y mantenimiento de bombas, valvulas, gritos, contenedores metálicos y maquinaria	
50720000 - Servicios de reparación y mantenimiento de calefacción central	
71314000 - Servicios de energía y servicios conexos	
45310000 - Trabajos de instalación eléctrica	
50232110 - Puesta a punto de instalaciones de iluminación pública	
45316100 - Instalación de equipo de alumbrado exterior 15316212 - Instalación de seméforos	
50200000 - Servicios de reparación, mantenimiento y servicios asociados	
elacionados con aeronaves, vías férreas, carreteras y la marina	
50730000 - Servicios de reparación y mantenimiento de grupos refrigeradores	
51300000 - Servicios de instalación de equipos de comunicaciones	
63712700 - Servicios de control del tráfico	
50532000 - Servicios de reparación y mantenimiento de maquinaria eléctrica,	
aparatos y equipo asociado	

3.2.2 Example of good practice

At NextProcurement, an NLP and AI-based system has been developed to determine the main CPV based on the subject matter of the tender.

The use of NLP and AI is especially valuable in tenders that present the aforementioned problems, as it allows the CPV to be calculated from the information contained in the tender.

Note: In cases where the description of the object is too brief (one or two words) or ambiguous, the results may not be optimal.

Examples of the improvements achieved are shown below:

• CPV allocation in tenders where the metadata was empty

servicios de carácter técnico, jurídico y de ges voluntaria y ejecutiva de las sanciones en mai embarcado en vehículo, un dispositivo de con	tión, complementarios, auxiliares y de recaudación en vía teria de tráfico así como la cesión de los dispositivos de radar trol de infracciones por salto de semáforo en rojo y etilómetro.
Datos generales	
CPVs	
Original	Generado
CPVs CPVS detalle	CPVs 79410000
Buscar similar por CPV	CPVS detalle 79410000 - Servicios de consultoría comercial y en gestión

• Suggestion of different CPV based on the content of the sheet.

Contratación de las obras e instalaciones del servicio d mantenimiento preventivo/correctivo de la reforma de supercomputador cuántico, y adaptación y adecuación	e redacción del proyecto técnico, construcción y l Marenostrum 4 para la implementación del ı de las instalaciones existentes.
Datos generales	
CPVs	
Original	Generado
CPVs	CPVs
45200000	71200000
CPVS detalle	CPVS detaile
45200000 - Trabajos generales de construcción de inmuebles y obras de ingeniería civil	71200000 - Servicios de arquitectura y servicios conexos

In addition, in NextProcurement, a specialised visualization has been developed that allows analysing the distribution of CPV in a set of tenders. This tool provides both contracting bodies and companies with a clear view of the variety and frequency of each type of code used, improving decision-making and strategic planning.

Licitaciones	s CPVs Solvencia técnica	Solvencia económica	Condicion	es especiales	Criterios de adjudicación	Empresas
Driginal			Generado			
Búsqueda Añadir co	personalizada ndición		Búsqueda p Añadir con	ersonalizada dición		
Filtros Act	ivos - 0 Colapsar Todo Mostrar	Todo Borrar todo	Filtros Activ	vos - 0	Colapsar Todo Mostrar Todo	Borrar todo
Nº ocurrer	ncias Q ×	5a; #‡ ×	Nº ocurrenc	ias	Q, × .	¥a; #‡ v
Buscar:	Mo	strar 25 🖌 registros	Buscar:		Mostrar	25 🗸 registros
	🛄 Ver las licitaciones 📑	📥 Descargar 🔹			🛙 Ver las licitaciones 🔹	🛓 Descargar 🔹
Mostrando :	1 a 25 de 361 registros		Mostrando 1	a 25 de 64 registri	os Anterior 1	2 3 Siguiente
	Anterior 1 2 3 4 5	15 Siguiente	сру	Descripción		N° A ocurrencias
СРУ	• • Descripción	▼ ocurrencias ▼	45233000	Trabajos de cons	trucción, cimentación y	39
50232000	Servicios de mantenimiento de instalaciones d	le 317		pavimentación d	e autopistas y carreteras	
	alumbrado público y semáforos		50232000	Servicios de man	tenimiento de instalaciones de	23
34996100	Semáforos	127		atumbrado publi	co y semaioros	
45316212	Instalación de semáforos	119	50230000	Servicios de repa	ración, mantenimiento y servicios	17
50000000	Servicios de reparación y mantenimiento	24		equipos	mados con carreteras y otros	
45000000	Trabajos de construcción	22	45316000	Trabajos de insta	lación de sistemas de alumbrado y	16
45233140	Obras viales	19		señalización		
50232100	Servicios de mantenimiento de alumbrado pút	olico 19	50330000	Servicios de man	tenimiento de equipo de	6

3.2.3 Recommendations

- Validate the accuracy of the assigned CPV using AI techniques, minimizing errors and improving bid classification.
- Automate the completion of metadata by uploading a new tender to the system, allowing the system itself to help the person responsible to enter this information accurately. This would reduce recording time and ensure that the information is machine-readable, improving its usefulness in automated analysis.

3.3 Technical Solvency Extraction through NLP and AI

In the case of technical solvency, there are specific metadata that would allow this analysis to be carried out quickly and efficiently. However, in many cases, this metadata is incomplete or poorly completed, making it difficult to use.

3.3.1 Example of malpractice

Some common problems related to the technical solvency metadata are the following:

- **Empty metadata**: although the information is present in the administrative specifications, it has not been included in the metadata.
 - *Example*: Tender ntp00862043

		Solvencia técnica-profesional
Pliego Cláusulas Administrativas	Pliego clausulas adminsitrativas particulares.pdf	Deberá acreditarse a través del siguiente medio:
Solvencia técnica		 Relación de contratos ejecutados de igual o similar naturaleza que los que constituyen el objeto del contrato en el curso de los TRES (3) últimos años avalada por centificados de buena ejecución; tesos centificados de importe, las fechas y el lugar de ejecución del contrato y se precisará sis se realizanon según las reglas por las que se rige la portesión y se llevaror
	Origir	normalmente a buen término.
Del Procedimiento		E1 importe anual acumulado en el año de mayor ejecución de contratos de igual o similar naturaleza ha de ser igual o superior al valor estimado del presente contrato (287.653,01 €).
De los Lotes		Para el cumplimiento de este apartado se tendrá en cuenta lo siguiente:
		o Se entenderá por contrato de igual o similar naturaleza al suministro o

- **Inaccurate metadata**: instead of providing details, it simply refers to "view the administrative specifications", which prevents its use in automated analyses.
 - *Example*: Tender ntp01097843

lvencia técnica	
	Original
Del Procedimiento	
Trabajos realizados - El l	citador deberá indicar que SÍ cumple con la solvencia técnica o profesional requerida, conforme a lo indicado en el apartado 8 del Anexo I del
PCAP:	

3.3.2 Example of good practice

At NextProcurement, a system based on NLP and AI has been implemented for the extraction of the technical solvency of the administrative document from the administrative specifications.

The use of NLP and AI is particularly valuable in tenders that present the problems mentioned above, as it allows technical solvency to be automatically extracted from the administrative specifications.

Examples of progress made include:

• **Correction of empty metadata**: the system detects, extracts the relevant information from the sheet and generates the rich metadata.

	Original	
el Procedimiento		
e los Lotes		
	Generado	

• **Greater precision in the metadata**: in this case, the specific figure has been detected, which was not indicated in the metadata.

	Original
Procedimiento	
abajos realizados - El licitador deberá ir CAP:	dicar que Sí cumple con la solvencia técnica o profesional requerida, conforme a lo indicado en el apartado 8 del Anexo I del
e los Lotes	
	Generado
ción de los principales servicios realizado or al 70 del valor <mark>estimado del lote.</mark>	os de igual o similar naturaleza en los últimos tres años, con un importe anual acumulado en el año de mayor ejecución <mark>gual o</mark>

In addition, in NextProcurement, a visualization has been developed that groups the technical solvency data of the tenders analysed. This functionality facilitates a global and rapid analysis for both contracting bodies and companies, allowing the variety and frequency of each type of technical solvency used to be observed.

Licitaciones	CPVs	Solvencia técnica	Solvencia económica	Condiciones especiales	Criterios de adjudicación	Empresas
Buscar:	-	M	ostrar 25 🗸 registros		🖽 Ver las licitaciones 🔹 🎍	Descargar 🔻
Mostrando 1 a 2	1 de 21 regist	ros			Anterio	or 1 Siguiente
Descripción			¢	Nº ocurrencias	🗧 Nº ocurrencias enriqueci	das 🛔
Trabajos realizad	os			159	52	
Sin datos orig	nales			38	38	
Otros				106	25	
Maquinaria, mat	erial y equipo te	écnico para la ejecución del co	ontrato	20	6	
Técnicos o unida	des técnicas			16	3	
Certificados de c	ontrol de calida	ad expedidos por los instituto:	s o servicios oficiales	14	2	
Títulos académio	os y profesiona	les del empresario y directivo	os, o responsables de la ejecución	19	2	
Control por la en	tidad contratar	ite sobre la capacidad del em	presario	3	1	

3.3.3 Recommendations

- **Promote the automatic extraction** of technical solvency from administrative specifications using AI techniques, with the aim of increasing transparency and facilitating comparability between procedures.
- Automate the completion of metadata by uploading a new tender to the system, allowing the system itself to help the person responsible to enter this information accurately. This would reduce recording time and ensure that the information is machine-readable, improving its usefulness in automated analysis.

3.4 Extracting Economic Solvency through NLP and AI

In the case of economic solvency, there are specific metadata that would allow this analysis to be carried out quickly and efficiently. However, in many cases, this metadata is incomplete or poorly completed, making it difficult to use.

3.4.1 Example of malpractice

Some common problems related to the metadata of economic solvency are the following:

• **Empty metadata**: although the information is present in the administrative specifications, it has not been included in the metadata.



• **Inaccurate metadata**: instead of providing details, it simply refers to "view the administrative specifications", which prevents its use in automated analyses.

Solvencia económica

	Original
Del Procedimiento	
Seguro de indemnización - Ver Anexo I del Pliego de Clausu	ulas Administrativas
Cifra anual de negocio - Ver Anexo I del Pliego de Cláusulas	Administrativas

3.4.2 Example of good practice

At NextProcurement, a system based on NLP and AI has been implemented for the extraction of the economic solvency of the administrative document of the administrative specifications.

The use of NLP and AI is especially valuable in tenders that present the problems mentioned above, as it allows the economic solvency to be automatically extracted from the administrative specifications.

Examples of progress made include:

• **Correction of empty metadata**: the system detects, extracts the relevant information from the sheet and generates the rich metadata.

	Original
el Procedimiento	
e los Lotes	
	Generado

• **Greater precision in the metadata**: in this case, the detail has been detected, which in the original referred to see the administrative specifications document.

lve	encia económica
	Original
De	Procedimiento
Se	guro de indemnización - Rellenar conforme anexo III PCAP
De	los Lotes
	Generado
istifi	cante de existencia de seguro de responsabilidad civil que cubra hasta un importe mínimo de 2.000.000,00 que deberá mantener durante el plazo de vigencia del contra

In addition, in NextProcurement, a visualization has been developed that groups the economic solvency data of the tenders analysed. This functionality facilitates a global and rapid analysis for both contracting bodies and companies, allowing the variety and frequency of each type of economic solvency used to be observed.

Licitaciones	CPVs	Solvencia técnica	Solvencia económica	Condiciones especiales	Criterios de adjudicación	Empresas
Buscar:			Mostrar 25 🗸 registros		🛄 Ver las licitaciones 🔹 🕹 De	scargar 💌
Mostrando 1 a 1	1 de 11 regist	ros			Anterior	1 Siguiente
Descripción			¢	Nº ocurrencias	🔹 Nº ocurrencias enriquecidas	÷
Sin datos orig	inales ni enriqu	ecidos		1022	0	
Cifra anual de ne	gocio			180	82	
Sin datos orig	inales			96	96	
Otros				59	28	
Seguro de indem	nización			34	14	
Patrimonio neto				8	6	

3.4.3 Recommendations

- Promote the automatic extraction of economic solvency from administrative specifications using AI techniques, with the aim of increasing transparency and facilitating comparability between procedures.
- Automate the completion of metadata by uploading a new tender to the system, allowing the system itself to help the person responsible to enter this information accurately. This would reduce recording time and ensure that the information is machine-readable, improving its usefulness in automated analysis.

3.5 Extraction of Special Conditions Using NLP and AI

In the case of special conditions, there are specific metadata that would allow this analysis to be carried out quickly and efficiently. However, in many cases, this metadata is incomplete or poorly completed, making it difficult to use.

3.5.1 Example of malpractice

Some common problems related to special condition metadata are as follows:

• **Empty metadata**: although the information is present in the administrative specifications, it has not been included in the metadata.



• **Inaccurate metadata**: instead of providing details, it simply refers to "view the administrative specifications", which prevents its use in automated analyses.

	Original
Some	timiento del contratista a la normativa de protección de datos - Según punto 22 del cuadro de características del pliego de clausulas administrativas particulares
Consid	deraciones tipo social - Según punto 22 del cuadro de características del pliego de clausulas administrativas particulares,

3.5.2 Example of good practice

At NextProcurement, a system based on NLP and AI has been implemented for the extraction of the special conditions of the document from the administrative specifications.

The use of NLP and AI is particularly valuable in tenders that present the problems mentioned above, as it allows special conditions to be automatically extracted from the administrative specifications.

Examples of progress made include:

• **Correction of empty metadata**: the system detects, extracts the relevant information from the sheet and generates the rich metadata.

Condiciones especiales	
	Original
	Generado
Condiciones especiales de ejecución	
- Social	
- Se podrán imponer condiciones especiales de ejecu	ción de carácter social y ético, en función de los objetivos que se estime son más factibles de alcanzar según el tipo de
suministro de que se trate en cada caso. Estas condicio	ones se especificarán en el cuadro de características de cada contrato.
- Medidas de fomento del empleo con personas de co	vlectivos desfavorecidos no podrán llevarse a cabo a costa de los trabajadores con los que ya cuente la empresa
adjudicataria.	
- Ética	
- Asegurar el respeto a los derechos laborales básicos	a lo largo de la cadena de producción, el comercio justo y la mayor transparencia y trazabilidad de toda la cadena
comercial.	
- Medioambiental	
- Se podrán imponer condiciones especiales de ejecu	ción de carácter medioambiental, en función de los objetivos que se estime son más factibles de alcanzar según el tipo de
suministros de que se trate en cada caso. Estas condici	iones se especificarán en el cuadro de características de cada contrato.

• **Greater precision in the metadata**: in this case, the detail has been detected, which in the original referred to see the administrative specifications document.

Condiciones especiales	
	Original
Sometimiento del contratista a la normativa de protección de datos Consideraciones tipo social - Según punto 22 del cuadro de característ	- Según punto 22 del cuadro de características del pliego de clausulas administrativas particulares. icas del pliego de clausulas administrativas particulares.
	Generado
ondiciones especiales de ejecución	
Social Se garantizará que al menos una de las nuevas contrataciones se refiera a p	ersonas en situación de desempleo y que al menos una sea mujer.
Ética	
Se establece la obligación de cumplir con la normativa vigente en materia d	e protección de datos.
Medioambiental	
Se compromete a respetar los principios de economía circular y evitar impa	ctos negativos significativos en el medio ambiente.
Otro	
Se deberá presentar un plan de prevención de riesgos laborales y cumplir o	on las condiciones de subrogación de trabajadores.

In addition, in NextProcurement, a visualization has been developed that groups the data of the special conditions of the tenders analysed. This functionality facilitates a global and rapid analysis for both contracting bodies and companies, allowing the variety and frequency of each type of special condition used to be observed.

Licitaciones	CPVs	Solvencia técnica	Solvencia económica	Condiciones especiales	Criterios de adjudicación	Empresas
Buscar:			Mostrar 25 🗸 registros		🗍 Ver las licitaciones 🔹 🎿	Desca <mark>r</mark> gar 🔹
Mostrando 1 a 1	1 de 11 registi	ros			Anterio	or 1 Siguiente
Descripción			\$	Nº ocurrencias	🗘 Nº ocurrencias enriquecio	las 🛔
Sin datos orig	inales ni enriqu	ecidos		1022	0	
Cifra anual de ne	gocio			180	82	
Sin datos orig	inales			96	96	
Otros				59	28	
Seguro de indem	nización			34	14	
Patrimonio neto				8	6	

3.5.3 Recommendations

- To promote the automatic extraction by AI techniques of special conditions from administrative specifications, with the aim of increasing transparency and facilitating comparability between procedures.
- Automate the completion of metadata by uploading a new tender to the system, allowing the system itself to help the person responsible to enter this information accurately. This would reduce recording time and ensure that the information is machine-readable, improving its usefulness in automated analysis.

3.6 Extraction of Award Criteria using NLP and AI

In the case of award criteria, there are specific metadata that would allow this analysis to be carried out quickly and efficiently. However, in many cases, this metadata is incomplete or poorly completed, making it difficult to use.

3.6.1 Example of malpractice

Some common problems related to the metadata of award criteria are the following:

• **Empty metadata**: although the information is present in the administrative specifications, it has not been included in the metadata.

		300V30 IV		
		Original		
		a (210) - 10000 - 1 (21 - 10		
S PARTIC	ULARES QUE	HAD 80 / 319 - 100% + ⊡ ♦)		
		Departamento	de Tecnologias del	Tráfico
		EX	EDIENTE: 300/2022	2/00590
	Nº criterio	Descripción	Puntuación	
	1	Calidad de la propuesta	23	
	1.a.	Metodología de los procedimientos, organización y planificación de los trabajos de operación y mantenimiento	7	
	1.b.	Metodología para la operación y mantenimiento de los sistemas informáticos y de comunicaciones	7	
	1.b. 1.c	Metodología para la operación y mantenimiento de los sistemas informáticos y de comunicaciones Características del regulador de tráfico ofertado	7	
	1.b. 1.c 1.d.	Metodología para la operación y mantenimiento de los sistemas informáticos y de comunicaciones Características del regulador de tráfico ofertado Propuesta de plan de innovación	7 7 2	
	1.b. 1.c 1.d. 2	Metodología para la operación y mantenimiento de los sistemas informáticos y de comunicaciones Características del regulador de tráfico ofertado Propuesta de plan de innovación Plan de ahorro energético	7 7 2 2	
	1.b. 1.c 1.d. 2	Metodología para la operación y mantenimiento de los sistemas informáticos y de comunicaciones Características del regulador de tráfico ofertado Propuesta de plan de innovación Plan de ahorro energético Total	7 7 2 2 2 25	
	1.b. 1.c 1.d. 2	Metodogia para la operación y mantenimiento de los sistemas informácios y de comunicaciones Características del regulador de tráfico ofertado Propuesta de plan de innovación Plan de ahorro energético Total	7 7 2 2 2 25	
10.	1.b. 1.c 1.d. 2	Metodología para la operación y mantenimiento de los sistemas informáticos y de comunicaciones Características del regulador de tráfico ofertado Propuesta de plan de innovación Plan de ahorro energético Total vropuesta: 23 puntos.	7 7 2 2 25	
<u>1G</u>	1.b. 1.c 1.d. 2	Metodogia para la operación y mantenimiento de los sistemas informánicos y de comunicaciones Características del regulador de tráfico ofertado Propuesta de plan de innovación Plan de ahorro energético Total ropuesta: 23 puntos.	7 7 2 2 25	
<u>1C</u>	1.b. 1.c 1.d. 2 alidad de la p	Metodogia para la operación y mantenimiento de los sistemas informáticos y de comunicaciones Características del regulador de tráfico ofertado Propuesta de plan de innovación Plan de ahorro energético Total ropuesta: 23 puntos. ará una memoría donde se relacionen todos los apartados	7 2 2 25 aue se descrit	ten a

• **Inaccurate metadata**: instead of providing details, it simply refers to "view the administrative specifications", which prevents its use in automated analyses.

3.6.2 Example of good practice

At NextProcurement, a system based on NLP and AI has been implemented for the extraction of the award criteria from the administrative specifications document.

The use of NLP and AI is particularly valuable in tenders that present the problems mentioned above, as it allows the award criteria to be automatically extracted from the administrative specifications.

Examples of progress made include:

• **Correction of empty metadata**: the system detects, extracts the relevant information from the sheet and generates the rich metadata.

Criterios de adjudicación	
Original	
Generado	
Criterios de adjudicación	
- Criterios no sujetos a juicios de valor Hasta 100 puntos	
- Precio Máximo 70 puntos	
- Se valorará con la máxima puntuación 70 puntos la oferta más baja presentada que cumpla todas las condiciones de los Pliegos proporcionalmente.	i, valorándose las demás ofertas
- Mejora relativa al compromiso de aportar sin coste para el Ayuntamiento un inventario de señalización municipal Máximo 20 pur	ntos
- Se asignará la máxima puntuación 20 puntos a la oferta que ofrezca el compromiso de entregar el inventario en el plazo menor	de tiempo mínimo de tres meses.
- Mejoras respecto al aumento de los medios mínimos exigidos en el PPT Máximo 7 puntos	
- Mejora de la garantía de la señalización horizontal Máximo 3 puntos	

 Greater precision in the metadata: in this case, the detail of detailed scores of an aspect has been detected. Criterios de adjudicación

	Original
Sin tipo - Sin subtij	o - 40 - 1.1 Criterio: Reducción del precio.
Sin tipo - Sin subti	o - 15 - 1.2 Criterio: Posesión de Certificados de Gestión
Sin tipo - Sin subti	o - 2 - 1.4 Criterio: Sustitución inicial de equipos lúminicos halógenos o incandescentes de 100mm de diámetro por equipos led de conformidad con e
Pliego de Prescripci	ones Técnicas.
Sin tipo - Sin subti	io - 15 - 1.5 Criterio: Reparación y puesta en marcha de los equipos de CTV existentes.
Sin tipo - Sin subti	😡 - 28 - 1.3 Criterio: Sustitución inicial de equipos lumínicos halógenos o incandescente de 200mm de diámetro por equipos LED de conformidad con
Pliego de Prescripci	ones Técnicas.
	Generado
riterios de adiudicación	
. Criterio Reducción del	precio máximo 40 puntos
- Se valorará con un má	ximo de 40 puntos la oferta que contenga la mayor reducción de precio con respecto al máximo de licitación.
- Fórmula de puntuació	nage demonstration of an and and
- PUNTUACIÓN OV 40	<pre>xPL OV PL-OMW</pre>
- Donde	
- OV Oferta a valorar.	
- PL Precio de licitaci	śn.
- OMW Oferta más ve	ntajosa de los licitadores admitidos.
. Criterio Posesión de ce	rtificados de gestión máximo 15 puntos
- Certificados considera	dos ISO-14001 gestión medioambiental e ISO-9001 gestión de calidad.
- Puntuación	
- Ninguno 0 puntos.	
- Uno 5 puntos.	
- Ambos 15 puntos.	
. Criterio Sustitución ini	cial de equipos lumínicos halógenos o incandescentes de 200mm de diámetro por equipos LED máximo 28 puntos
- Se valorará la oferta d	e al menos 10 equipos.
- Puntuación	
- Ninguna 0 puntos.	
- 10 equipos 2 puntos.	
- 20 equipos 5 puntos.	
- 30 equipos 9 puntos.	
- 40 equipos 14 punto:	
- 50 equipos 20 punto:	
- 58 equipos 28 punto:	
. Criterio Sustitución ini	cial de equipos lumínicos halógenos o incandescentes de 100mm de diámetro por equipos LED máximo 2 puntos
- Se valorará la oferta d	e al menos 6 equipos.
- Puntuación	
- Ninguna 0 puntos.	
- 6 equipos 1 punto.	
- 8 equipos 2 puntos.	
. Criterio Reparación y p	uesta en marcha de los equipos de CTV existentes máximo 15 puntos
- Puntuación	
- Ninguna 0 puntos.	
- 1 grupo 5 puntos.	
- 2 grupos 15 puntos.	
6. Criterios específicos p	ara resolver empates
- Mayor porcentaje de t	rabajadores con discapacidad.
- Mayor porcentaje de t	rabajadores en situación de exclusión social.
- Menor porcentaje de o	ontratos temporales.
- Mayor porcentaje de r	nujeres empleadas fijas.
- Mayor porcentaje de r	nujeres empleadas temporales.
0 1	los anteriores criterios no resuelvan el emnate

In addition, in NextProcurement, a visualization has been developed that groups the data of the award criteria of the tenders analysed. This functionality facilitates a global and rapid analysis for both contracting bodies and companies, allowing the variety and frequency of each type of award criterion used to be observed.

Licitaciones	CPVs	Solvencia técnica	Solvencia económica	Condiciones especiales	Criterios de adjudicación	n Empresas
Buscar:			Mostrar 25 🗸 registros		🖽 Ver las licitaciones 🔹	🛓 Descargar 🔹
Mostrando 1 a 1	1 de 11 registi	ros			Ante	erior 1 Siguiente
Descripción			÷	Nº ocurrencias	🛔 Nº ocurrencias enrique	cidas 🍦
Sin datos orig	inales ni enriqu	ecidos		1022	0	
Cifra anual de ne	gocio			180	82	
Sin datos orig	inales			96	96	
Otros				59	28	
Seguro de indem	nización			34	14	
Patrimonio neto				8	6	

3.6.3 Recommendations

- Promote the automatic extraction of award criteria from administrative specifications using AI techniques, with the aim of increasing transparency and facilitating comparability between procedures.
- Automate the completion of metadata by uploading a new tender to the system, allowing the system itself to help the person responsible to enter this information accurately. This would reduce recording time and ensure that the information is machine-readable, improving its usefulness in automated analysis.

3.7 Extracting Tenderers Using NLP and AI

One of the fundamental principles of public procurement is transparency. And a key aspect of this is the publication of detailed information about bidders. However, currently procurement metadata only includes information about successful bidders, omitting essential data about all participants in the bidding process.

In addition, published information about bidders is often incomplete or inconsistent, making it difficult to analyse, audit, and reuse. Issues such as missing unique identifiers or duplicate names create ambiguity and reduce the effectiveness of access to this data. Abnormalities occur:

3.7.1 Example of malpractice

Some common problems related to bidders are the following:

• Publication as metadata of all bidders is missing

The Public Sector Procurement Platform (PLACE) only publishes as metadata the successful bidder and, in some cases, the total number of bidders, but without detailing who they were.

For example, in a real tender we find:

- Awardee: B39640263 SITELEC GLOBAL DE SERVICIOS Y OBRAS, S.L.
- Number of bids received: 7

The problem is that the other 4 bidders are not specified, which makes it impossible to know what the competition has been without going tender by tender looking for the information in the minutes.

• Lack of quality of bidders

Another common problem is inconsistency in company names.

In the examples below you can see cases of how bidders' information is stored:

• The identifier field includes multiple identifiers and the company name: ntp01393258



• In the names, the information of the same company is put in multiple different ways. For example, in different tenders the same company may appear as:

ALUMBRADOS VIARIOS, SA	
ALUMBRADOS E S VIARIOS, S.A	
ALUMBRADOS VARIOS SA	
ALUMBRADOS VARIOS, S.A.	
ALUMBRADOS VIARIOS O5 S.A.	
ALUMBRADOS VIARIOS S.A	
ALUMBRADOS VIARIOS S.A. ALUVISA .	
ALUMBRADOS VIARIOS SA	
ALUMBRADOS VIARIOS, S.A	
ALUMBRADOS VIARIOS, S.A.	
ALUMBRADOS — ÉÉ \$ VIARIOS, S.A.	AERONAVAL DE CONSTRUCCIONES E INSTALACIONES, S.A
ALUVISA	AERONAVAL DE CONSTRUCCIONES E INSTALACIONES S.A
ALUVISA ALUMBRADOS VIARIOS S.A	AERONAVAL DE CONSTRUCCIONES E INSTALACIONES S.A.U.
ALUVISA, ARTIC	AERONAVAL DE CONSTRUCCIONES E INSTALACIONES, S.A
ALUVISA, S.A.	AERONAVAL DE CONSTRUCCIONES E INSTALACIONES, S.A.
ALUVISSA	AERONAVAL DE COSNTRUCCIONES E INSTALACIONES, SA

This creates confusion, as it is not possible to determine with certainty whether it is the same entity.

3.7.2 Example of good practice

The use of NLP and AI is especially valuable in tenders that present the issues mentioned above, as it allows bidders to be automatically extracted from the minutes documents.

• Extract all bidders as enriched metadata

In NextProcurement, a process of automatic extraction of bidders from the tender documents has been developed. This system automatically extracts the names and identifiers (if any) of all participants in the process.

The process is not always perfect and identifies bidders well. This is because, when extracting data from documents, they are not always well defined, clear, etc. In addition, an identifier (NIF/VAT) is rarely included in the bidders submitted, making it more difficult to correctly identify bidders.

For example, in a tender analysed in NextProcurement, it was possible to extract this structured information from all bidders:

- Awardee: B39640263 SITELEC GLOBAL DE SERVICIOS Y OBRAS, S.L.
- Number of bids received: 7
- Extracted bidders:

Licitadores

Original	Generado				
Adjudicatario: B39640263 - SITELEC GLOBAL DE SERVICIOS Y OBRAS, S.L 13077313.7 - Es PYME Ver Licitador	Presentado - OMNI COMERCIAL S.A. - GRUPO FERSON - 03 S.L. - AN ENERGETICOS S.L. - IEEEDEOLA GUENTES - ARU OBRAS Y SERVICIOS S.L. - ICELIGHT - SIME S.E.N.C. - ALDRO ENERGIA Y SOLUCIONES S.L.U. - SACYR SOCIAL S.L. - LE BUDIGS				

• Unification of companies

At NextProcurement, a process of cleaning and unifying company names has been developed in those cases in which the only difference was the way of naming the type of company (SA, S.A. SA., etc.), blank spaces, etc. There has been a limit in this unification when there was some additional difference in the names (some more words, other additional characters), where out of prudence not being able to verify whether or not it was the same company, the original data have been left. As an example, the unification carried out with two examples is shown:

ALUMBRADOS VIARIOS, SA						
ALUMBRADOS E S VIARIOS, S.A	ALUMBRADOS _GA VIARIOS S.A.					
ALUMBRADOS VARIOS SA						
ALUMBRADOS VARIOS, S.A.	ALUMBRADOS E S VIARIOS, S.A					
ALUMBRADOS VIARIOS O5 S.A.						
ALUMBRADOS VIARIOS S.A	ALUMBRADOS VIARIOS S.A					
ALUMBRADOS VIARIOS S.A. ALUVISA .						
ALUMBRADOS VIARIOS SA	ALUMBRADOS VIARIOS S.A. ALUVISA .					
ALUMBRADOS VIARIOS, S.A						
ALUMBRADOS VIARIOS, S.A.	ALUMBRADOS VIARIOS, S.A.					
ALUMBRADOS - ÉÉ \$ VIARIOS, S.A.						
ALUVISA	ALUMBRADUS - EE \$ VIARIUS, S.A.					
ALUVISA ALUMBRADOS VIARIOS S.A	411.0.404					
ALUVISA, ARTIC	ALUVISA					
ALUVISA, S.A.	41104000					
ALUVISSA	ALUVISSA					
ERONAVAL DE CONSTRUCCIONES E INSTALACIONES, S.A						
ERONAVAL DE CONSTRUCCIONES E INSTALACIONES S.A	AERONAVAL DE CONSTRUCCIONES E INSTALACIONES, S.A.					
ERONAVAL DE CONSTRUCCIONES E INSTALACIONES S.A.U.	AERONAVAL DE AS					
ERONAVAL DE CONSTRUCCIONES E INSTALACIONES, S.A	AERONAVAL DE CONSTRUCCIONES E INSTALACIONES S.A.U ACISA					
ERONAVAL DE CONSTRUCCIONES E INSTALACIONES, S.A.						
ERONAVAL DE COSNTRUCCIONES E INSTALACIONES, SA	AERONAVAL DE CONSTRUCCIONES E INSTALACIONES, S.A.U.					

In addition, in NextProcurement, a visualization has been developed that groups the data of the bidders of the analysed tenders. This functionality facilitates a global and rapid analysis for both contracting bodies and companies, allowing the variety and frequency of each type of bidder to be observed.

citaciones CPVs Solvencia técnica	Solvencia económica	Condiciones especia	les Criterios d	le adjudicación	Empre
judicatarias					
Búsqueda personalizada Añadir condición					
Filtros Activos - 0			Colapsar Todo	Mostrar Todo	Borrar todo
Nombre Q, × ÅA, #↓ v	Identificador Q	× 4A; #‡ v	ES PYME	Q. × 1	A; #‡ ×
Es UTE Q	× 44; #‡ v	Número licitaciones/lotes		Q, × 4	A; #‡ v
uscar: N	ostrar 10 🗸 registros		🛙 Ver las licitad	iones 👻 🔒	L Descargar
lostrando 1 a 10 de 401 registros			Anterior 1 2	3 4 5	. 41 Siguie
lombre	▲ ▼ Identificador	⊾ Es ⊾ ▼ PYME ▼	▲ Número Es UTE ▼ licitacio	nes/lotes [▼] in	ima iportes
OCIEDAD IBERICA DE CONSTRUCCIONES ELECTRICAS, S.A.	A28002335	No	No 235	26	.973.979,98€
LECTRONIC TRAFIC, S.A.	A46138921	No	No 160	7.9	04.682,82€
URCIANA DE TRAFICO, S.A.	A30035927	sí	No 113	1.5	515.539,66€
LUMBRADOS VIARIOS, S.A.	A08523094	Sí	No 61	12	.589.057,62€
apsch TrafficCom Transportation, S.A.U.	A78107349	No	No 52	8.8	864.503,10€
LECTRICA DE ALZIRA -CEAL, S.C.V.	F46161576	No	No 44	37	.701,82€
umican SA	A35038900	No	No 34	74	9.600,80€
ERONAVAL DE CONSTRUCCIONES E INSTALACIONES, S.A. (A	CISA) A28526275	No	No 29	7.2	234.381,63€
TRA BONAL, S.A.	A08522955	No	No 16	4.1	127.108,93€
TRALUX SA	A46066791	No	No 16	18	4.663,13€
Itadoras Júsqueda personalizada Añadir condición					
iltros Activos - 0			Colapsar Todo	Mostrar Todo	Borrar tod
Nombre Q × AA, #‡ v	Identificador Q	× 44, #‡ v	Número licitaciones	Q × 1	A; #‡ v
uscar:	ostrar 10 🗸 registros		🖽 Ver las licitad	iones 🝷	L Descargar
ostrando 1 a 10 de 471 registros			Anterior 1 2	3 4 5	. 48 Siguie
lombre		‡ Id	entificador	🗘 Número li	citaciones
LECTRONIC TRAFIC S.A.				11	
MEVISA S.A SUMINIS.CLINICOS				9	
OCIEDAD IBÉRICA DE CONSTRUCCIONES ELÉCTRICAS S.A. S	I. C E			9	
T.E. IMD SEVILLA 2019				7	

3.7.3 Recommendations

- Mandatory publication of all participating bidders, including in the open procurement metadata the complete list of companies that have submitted bids, not only the successful bidder.
- Use of standardised unique identifiers, associating each company with its VAT, NIF or identifier number in a public business register to avoid duplication and ambiguity.
- Automatic disambiguation of companies using AI techniques, implementing AI algorithms that normalise company names and correctly link them with their unique identifiers.
- Integration with business databases, connecting contracting platforms with open business registers (OpenCorporates, VAT Information Exchange System, national commercial registers, etc.).

3.8 Object search, including semantic search

Object search is the most intuitive and natural way for users to find tenders of interest to them. In the context of public procurement, the object of the contract is the

description of the goods, services or works to be acquired. However, limitations in the way this data is structured and published often make it difficult to search effectively.

This presents the following difficulties:

- Difficulty in finding tenders due to natural terminology: Users may use different expressions to refer to the same concept. For example, a tender on "open data" may contain terms such as open data, open data, reusable information, etc.
- No semantic search: Many platforms do not offer a search that interprets the meaning of words and their relationships, but instead rely on exact term matches, which reduces the accuracy and usefulness of the results.

Given that the subject matter of the contract is one of the main search criteria for interested companies and entities, improving the accessibility and accuracy of this information is key to more efficient and transparent public procurement.

3.8.1 Example of malpractice

One of the main current problems is that, on some platforms such as Plataforma de Contratación del Sector Público (hereinafter PLACE), there is no free text search system for the object of the contract.

Expediente			Pais	Todos			~
Tipo de Contrato	Todos	~	Lugar de Ejecución				-
Código CPV			Añadir Selección CPV				
CPV Seleccionados					*	Quitar	
					×		
			Presentación		▼ y		
lombre O. Contratación			Presentación Forma presentación	Todas	y y		~
lombre O. Contratación listema de Contratación	Todos	v	Presentación Forma presentación Procedimiento	Todas Todos	¥ У	•	*
lombre O. Contratación listema de Contratación īpo de tramitación	Todos Todos	~	Presentación Forma presentación Procedimiento Fecha publicación entre	Todas Todos	 y y	~	~
Nombre O. Contratación Sistema de Contratación Tipo de tramitación Estado	Todos Todas Todas	* * *	Presentación Forma presentación Procedimiento Fecha publicación entre Financiación UE	Todas Todas Todas	 y y	¥	· · ·
Nombre O. Contratación Sistema de Contratación Tipo de tramitación Estado Adjudicatario	Todos Todos Todos	* * *	Presentación Forma presentación Procedimiento Fecha publicación entre Financiación UE Fuente	Todas Todas Todas Todas	 y y	v	· · · · · ·

This means that users must know the exact CPV codes or manually review multiple documents to find relevant opportunities.

Returning to the example of a tender on "open data" it can contain terms such as open data, open data, reusable information, etc.

Due to the lack of a search by contract object, each term may generate different results or even not return results, despite the fact that there are relevant tenders. This fragmentation in the search makes it difficult for companies to participate, especially SMEs, which do not have teams dedicated to tracking tenders intensively.

Another frequent problem is the inconsistency in the CPV classification. For example, a tender for data analytics might be labelled under "IT consulting services" or "data processing services," causing some potential vendors to miss it.

3.8.2 Example of good practice

Within the framework of the NextProcurement project, an advanced solution has been developed that solves these limitations:

• Free text search:

A system has been implemented that allows users to search for any keyword within the contract object, eliminating the exclusive reliance on CPV codes.

IENT																				
DIDIOMA -																				
Ruscador de l	icitacio	nes	de re	ford	encia															
buscudor de l	rentacio	ine a	acre		eneru	ter edd	ncacion													
Texto:		s	amaforo																	
		BÚ	SQUEDA	APLIC	ADA: Po	r texto									-	_				
			TBL	ISQUE	DA POR 1	OTXEN			TBL	ISQUE	DA S	EMÁNI	ICA		× Limp	lar				
Licitaciones	CPVs		Solveni	ila téc	inica	Solv	rencla econi	ómica		Cor	dic	ones	spec	lales	Criterios d	ie adju	dicaci	ón	E	
Búsqueda pe	rsonalizada																			
Añadir condi	:kón																			
Filtros Activo	s - 0													Col	lapsar Todo	Mos	trar To	do	Borra	e
Organismo	9,		- 54	#1	×	Organ	iismo ID	0		~ 1	R.	#ļ	×	URL Pla	taforma	Q,		44	#]	
Tipo de contra	ato a		N.	#1	¥.	Tipe	le procedimie	ento <		× 3	И,	#1	*	Tipo de	tramitación	Ο,		N.	#1	
Estado	0.		- 50	81	×	Adjud	icatarios	9		6.3	W.	H‡	×	Adjudic	atario PYME	۵,		30	#1	
Buscan						Mostr	ar 25 👻 reg	istros					1] Ver mas ca	mpos dispor	nibles	*	4	Descar	
Mostrando 1 a	25 de 1.379 i	registr	os											Ani	erior 1 2	3	4 5		56 5	ii j
Objeto	• Obj • lote		Exper	liente					esup	iesto		Fecha	÷.	rganismo	Organism ID		PCAD			
Contrato de obr	as de		10942/	2024			2	19	614,38	ж		024-12	м	essenesse	L01461653					j
emergencia de Instalación de												31			P46167000 306490603	C LOB666				
semáforos daña la DANA.	dos por																			
Suministro 02			205091	16D			34095100 -	1.3	88,40	ł.		2024-12	в	cairent	L01460728	8				
Senaroros neu s	our .						semploros								300001100	000027				
Instalación de semáforos LP-2			OBR-	IFRAES	STRUCT20	24000353	45000000 - Trabalos de	39	712,00	×		2024-12	0	ibildo Insular la Palma	L03380010 P38000028	B				
"Circumalación TM de Los Llas	Sur"		100000000000000000000000000000000000000			construcción		a. Destrona			316620200	02062								
Aridane																				
Suministro Siste para Invidentes	en		14969)	2024			34595100 - Semáforos	1.3	56,70	2		2024-12	М	ogán	L01350120 P35013008	8				
semáforo ber La	Parada														310494003	48700				

• Semantic search:

Artificial intelligence has been incorporated to interpret the context and meaning of the terms used, so that, even if users type different words, the search can identify similar concepts. For example, if a user searches for open data, the system is able to retrieve tenders that contain related terms such as open data, reuse of public information, or data transparency.

CEF Telecom

0:		P	ortal de	datos	abiert	ios												Z E
		ВÚ	SOUEDA	APLIC	ADA:	Semántica												
		T BU	SQUEDA	PORT	EXTO		т ви	squ	EDA SE	MÁNT	ICA		× Limpian					
Licitaciones	CPVs		Solveno	ia téc	nica	Solver	ncia económi	ca	0	ondi	ciones e	special	es Criterios	de adji	dicaci	ón	Em	npres
Búsqueda persor Añadir condición	alizada																	
Filtros Activos - 0													Colapsar Tod	Mos	trar To	do	Borrar t	todo
Organismo	Q,		44.	#1	v	Organisi	no ID	ο,	×	41.	#1	w.	URL Plataforma	Q,		44	#1	×
Tipo de contrato	Q,		44.	ИŢ	×	Tipo de j	procedimiento	0,		44.	#1	w.	Tipo de tramitació	۹.		44.	#1	×
Estado	Q,		Å4.'	#1	×	Adjudica	itarios	ο,	×	44.	#1	v.	Adjudicatario PYM	a,		54.	#1	×
Buscart						Mostrar 25	 registros 					ШW	er mas campos disp	mibles	-	*	Descarga	r •
Mostrando 1 a 25 d	e 98 regi	stros												Anterio	r 1 :	2 3	4 Sig	quient
	hinin Ini	. :		Inches	÷ .		Description of the	\$	-		ò	4	Organismo	DCAD	÷	. :		
Suministro de una plataforma en CLOUD de Open Data Software	ojeto tot		2024.74	4	41 Pi 55 89	3218000 - aquetes de aftware de astión de cencias	40.000,006		2024-1	0-06	Grupo E Irrati Tel	uskal ebista	171371721817222.0	•	*		*	
Portal corporativo de publicación, visualización y analítica de datos abiertos			16133	/2019	44 Pr 55 51 51 51 51 51 51 51 51 51 51 51 51	aquetes de eftware y stemas de formación 2288000 - arvícios de iministro de eftware	34.814,64E		2019-1	1-25	Conceja Haciend Organiz Municip Empleo Ayuntan Gijón	lía de la, aclón al y del siento de	L01330241 P3302400A	•	•		*	
Servicios de Implantación de un ecosistema tecnológico para la gestión de Open Data			CONTR 5123	2023	73 Si m y si	1267000 - ervícios de antenímiento reparación de eftware	224.000,00€		2023-0	5-16	Comuni Autónor Balears	dad na Illes	A04003003 50711001H 21387900182550	*	*		*	
Trabajos de mantanimiento y actualización del SIG de Cáceres y portales de datos abiertos			CON-58 202200	ERV- 0055	7: Si m y si	2267000 - arvicios de antenimiento reparación de aftware	79.338,84€		2023-0	2-21	Cáceres		L01100377 P1003800H 30981550141915	÷	•		÷	
Servcio de asistencia tecnica para el desarrollo e			2018-0	1699	7: 51 61 52	1230000 - ervicios de esarrollo de oftware	230.000,000		2019-0	5-06	Direcció Económ Financie Entidad	n General Ico- tra de la Pública	EA0003337 Q2801659J	+	•		+	

3.8.3 Recommendations

- Allow free text search in all key fields of the tender, especially in the object of the contract
- **Incorporate semantic search as a** standard in public procurement platforms to improve the accuracy and accessibility of information.

4 Other recommendations not related to NLP and AI

In the execution of NextProcurement, in addition to carrying out NLP and AI processes, other functionalities have been developed that are considered best practices and recommendations applicable to public procurement and are those presented in this section.

4.1 Better data integration from Connected Platforms with PLACE

The Public Sector Procurement Platform (PLACE) is the central point for the publication of tenders in Spain. However, there are numerous regional and sectoral platforms that are connected to PLACE but do not fully integrate its data, which generates inconsistencies and loss of relevant information.

The Platforms connected to PLACE are as follows:

- Procurement Platform of the Generalitat of Catalonia
- Basque Government Procurement Platform
- La Rioja Contracting Platform
- Contracting Platform of the Community of Madrid
- Galician Public Procurement Platform
- Procurement Platform of the Regional Government of Andalusia
- Government of Navarre Procurement Platform

Plataformas de Comunidades y Ciudades Autónomas conectadas con la Plataforma de Contratación del Sector Público									
	piataforma KPEUSK880.02	Comunidad de La Rioja	****	CONTRATOS PÚBLICOS DE GALICIA	Α	navarra.es			

The following map shows the territorial distribution of the autonomous communities that have their own contracting platform and those that must necessarily make use of the PLACSP. As a result of all the above, it is concluded that the contracting bodies corresponding to the state, regional and local levels of the areas marked in dark colour will necessarily have their contractor profile housed in the PLACSP.



This means that they do not publish directly to PLACE but have their own applications and through a synchronization system, pass the data to PLACE.

However, a lot of data in this synchronization process is not transferred, which generates loss of very important information, making it difficult to do global searches.

The lack of a complete and standardised integration between PLACE and the regional platforms prevents a global vision of public procurement and hinders the interoperability of data. This situation affects companies looking for opportunities in different regions and organizations that try to analyse the recruitment market effectively.

4.1.1 Example of malpractice

Currently, many regional platforms send data to PLACE through synchronization processes that do not guarantee the integrity and quality of the information. Common problems include:

- Lack of key data: Some tenders published on regional platforms do not include all the information when they are synchronized with PLACE, such as award criteria or specifications.
- **Inconsistent coding:** Different platforms use different nomenclatures for the same concepts, making it difficult to normalise data.
- **Duplication or disappearance of tenders:** In some cases, tenders published on regional platforms do not appear correctly in PLACE or present incomplete information.

Example:

• Technical solvency, economic solvency, empty award criteria

Solvencia técnica		De los Lotes	
	Original	Del Procedimiento	
Del Procedimiento			Original
De los Lotes		Solvencia económica	
	Criterios de adjudicación		
		Original	

Different encodings

In PLACE, the Procedure Types have an official coding that is as follows:

https://contrataciondelestado.es/codice/cl/2.07/SyndicationTenderingProc essCode-2.07.gc

But, for example, the Procurement Platform of the Generalitat of Catalonia uses other codings, which do not coincide with either the codes or the number of codings.

- [1, "Open"], 401 Obert
- [2, "Restricted"], 402 Restricted
- [3, "Negotiated without advertising"], 421 Negociat sense publicitat
- [4, "Negotiated with advertising"], 11085 Negociat amb publicitat
- [5, "Competitive Dialogue"], 405 Competitive Dialogue
- [6, "Minor contract"], 403 Minor contract
- [7, "Derivative of framework agreement"],
- [8, "Project competition"],1000010 Concurs de projectes
- [9, "Open Simplified"], 419 Obert simplificat
- [10, "Association for Innovation"]1000011 Association for Innovation
- [11, "Derivative of partnership for innovation"],
- [12, "Based on a dynamic acquisition system"], 1000003
 Specification of Dynamic Acquisition System
- [13, "Bidding with negotiation"], 404 Bidding with negotiation
- [100, "Internal Rules"], 1000012 Other Procedures According to Internal Instructions
- [999, "Other"],
- [], 1008211 Non-minor direct awards
- [], 1000008 Abbreviated simplified open procedure

4.1.2 Example of good practice

• Technical solvency, economic solvency, enriched award criteria

In NextProcurement, the specifications have been read, and the information has been extracted, generating it as enriched metadata.

• Technical solvency

Descripción	🛔 Nº ocurrencias	Nº ocurrencias enriquecidas	*
Sin datos originales ni enriquecidos	238	0	
Sin datos originales	98	98	

• Economic solvency

Descripción	🛔 Nº ocurrencias	Nº ocurrencias enriquecidas
Sin datos originales ni enriquecidos	173	0
Sin datos originales	163	163

• Award criteria

Descripción	▲ ▼ N° ocurrencias	▲ Nº ocurrencias▼ enriquecidas
Sin datos originales ni enriquecidos	173	0
Sin datos originales	163	163

Solvencia técnica

	Original
Del Procedimiento	
De los Lotes	
	Generado
Relación de las principales obras ejecutadas e	en los últimos cinco años que incluya importe, fechas y el destinatario, público o privado.
No se menciona información específica sobre	licencias o certificaciones necesarias.

Solvencia económica

	Original
el Procedimiento	
e los Lotes	
	Generado

Criterios de adjudicación

	Original
	Generado
Criterios de adjudicación	
- Criterios sujetos a juicio de valor Máx. 25 puntos	
- A Estudio de la obra 25 puntos.	
- A.1. Forma de ejecución de la obra 15 puntos.	
- A.2. Análisis y minimización de afecciones 10 puntos.	
- Criterios automáticos Máx. 75 puntos	
- A Precio 75 puntos.	

• Unified Encodings

In the case of the Procurement Platform of the Generalitat de Catalunya, which is one of the members of the consortium, an additional data upload has been made from its own open data that has this information encoded. In addition, the codifications of the Generalitat of Catalonia have been converted to the common codification of PLACE.

4.1.3 Recommendations

- **Standardise the formats for data exchange** between PLACE and the regional platforms to avoid inconsistencies.
- Implement automatic validation mechanisms to ensure data integrity during synchronization.
- Ensure that all key metadata (award criteria, bidders, attached documents) are included in the synchronization processes.
- **Regularly monitor and audit the quality** of data integration, correcting problems of duplicity or loss of information in the transfer.

4.2 Improving data from contracting bodies

Information on contracting bodies is essential to ensure the transparency and traceability of public procurement processes.

However, in many cases, data from contracting bodies presents problems such as missing identifiers, inconsistent identifiers, lack of standardisation and little detailed information. This makes it difficult to analyse tenders, evaluate agency performance, and interoperability of data across platforms.

An improvement in the quality and accessibility of data from contracting bodies would allow greater control over procurement processes, facilitate the analysis of trends in procurement and contribute to the detection of possible irregularities.

4.2.1 Example of malpractice

Some common problems with contracting body metadata include:

- **Inconsistency in the names of the contracting bodies:** The same body may appear under different names in different tenders.
- **Lack of unique identifiers:** A standardised code (such as NIF or an official identifier) that allows the correct linking of data is not assigned.
- **Incomplete or incorrect information:** Key data such as location, type of administration or organisational structure of the contracting body are not included.

Examples of original data from contracting authorities:

• Empty Contracting Body Identifier

This happens very often in tenders that are synchronized from other platforms:

Expediente 🛔	Presupuesto	🖡 Fecha	🗘 Organismo 🛔 Organismo ID	÷
555/2021 ACTUACIONES 2023-1	49.581,15€	2023-12-19	Ayuntamiento de	
			Galdakao	
2018/010/073/08	842.975,22€	2023-12-12	Diputación Foral de	
			Bizkaia	
3051_Cont_suministro 2023/14	7.834.022,31€	2023-09-29	Ayuntamiento de	
			Pamplona	
3051_CONT_SUMINISTRO/2023/11	8.057.851,24€	2023-07-10	Ayuntamiento de	
			Pamplona	
CONTR/2023/347182	152.365,15€	2023-05-26	Junta de Andalucía	

• Identifier with decimals

Organismo	🛔 Organismo ID
Ayuntamiento de	170341703517036.0
Portugalete	

• Variable number of organism identifiers

There are organizations that have 3 identifiers, others 2 and others one.

Organismo	Organismo ID
Massanassa	L01461653
	P4616700C
	30649060108666
Organismo 🛔	Organismo ID
Junta de Extremadura	S0611001I
	21270240170784
Organismo 🛔	Organismo ID
Consejería de Vivienda, Transportes e Infraestructuras	A13003093

4.2.2 Example of good practice

NextProcurement has made a process that extracts the information from the contracting bodies and makes it available in an API. This is done with all tenders, whether they come from PLACE or from the Platforms that are synchronized. This in itself is an advance over what is now available in PLACE, since there you can only obtain the identifiers of the bodies that publish directly in PLACE.

In NextProcurement, based on this API, a search of contracting bodies by name or identifier is allowed.

Or Ver explicació	n	Buscar Entidad zaragoza			
		A50017250 - COMISIÓN EJECUTIVA DE MERCAZARAGOZA S.A.			
		A50377787 - SOCIEDAD MUNICIPAL ZARAGOZA CULTURAL, S.A.U.			
rganismo (NIF/DIR3/Id Platafc		A50381433 - GERENCIA DE ZARAGOZA DEPORTE MUNICIPAL S.A.U. (MEDIO PROPIO)			
	TE	A50907666 - GERENCIA DE ECOCIUDAD ZARAGOZA S.A.U. (MEDIO PROPIO)			
		A50928845 - CONSEJO DE ADMINISTRACIÓN DE ZARAGOZA ALTA VELOCIDAD 2002, S.A.			
_		A99056442 - EXPO ZARAGOZA EMPRESARIAL, S.A.			
11	12	B50852763 - CONSEJO DE ADMINISTRACIÓN DE TERMINAL MARÍTIMA DE ZARAGOZA, S.L.			

This is also an improvement over what is in PLACE, since it does not allow searching by text, but browsing through a drop-down menu of all the contracting bodies it has.

4.2.3 Recommendations

- Assign standardised unique identifiers to each contracting body to avoid ambiguities.
- **Standardise agency names** across all procurement platforms to ensure data consistency.
- **Publish complete and structured information about each contracting body**, including contact details, organisational structure and administrative level.
- Implement validation tools and detection of inconsistencies in the data of the contracting bodies to improve their quality.
- Facilitate access to information through APIs and open formats, ensuring interoperability with other government databases and data analysis platforms.

4.3 Improve the quality of metadata in general

In addition to the quality issues related above to platform integration, there are metadata that have a lot of lack of quality such as incomplete data, errors in transcription, or lack of consistency in the coding of information.

The lack of automatic validation mechanisms causes problems such as difficulty in classifying contracts, the impossibility of comparative analysis and the reduction in the quality of digital services that depend on this data.

4.3.1 Example of malpractice

Currently, many public procurement platforms allow manual entry of metadata without validation, leading to errors and inconsistency in published information. Some of the most common problems include:

- Lack of standardisation in the coding of procedures: Different platforms use different nomenclatures to describe procurement procedures, which prevents their automated comparison.
- Incomplete metadata: Tenders are often published without including key information, such as CPV code, award date, or evaluation criteria.
- Typographical and formatting errors: Some fields contain misentered data, such as wrong identifiers, incorrect contract amounts, or inconsistent dates.

Examples:

• No procedure type



• No type of processing

Tip <mark>o d</mark> e tramitación	Q	×	Aa‡	#‡	٨
No data					162
Ordinaria					1192
Urgente					19
Emergencia					6

• No CPV and with 1-word name

This is of particular interest to SMEs, as these are usually minor contracts (although there are also open tenders), which are the lowest in value and are generally more eligible for by SMEs. If you can't search by CSV (because it doesn't exist), you can't search by object text (currently in PLACE), and the object is just a word (which isn't always meaningful), you limit competition a lot or leave them deserted.

Vestuario		Contrato menor	593,64€
Vestuario	-	Contrato menor	67,73€
Vestuario	-	Contrato menor	64,53€
Vestuario	-	Contrato menor	9.994,29€
Vestuario	-	Contrato menor	1.309,79€
Vestuario	-	Contrato menor	832,20€
Vestuario	-	Contrato menor	324,98€
Vestuario		Contrato menor	1.416,76€
Vibratomo		Contrato menor	18.000,00€
Visado	2	Contrato menor	789,03€
Visado	-	Contrato menor	3.888,85€
fiambre	15100000 - Productos de origen animal, carne y productos cárnicos	Contrato menor	6.919,35€
fiambre	15100000 - Productos de origen animal, carne y productos cárnicos	Contrato menor	3.833,66€
fiambre	15100000 - Productos de origen animal, carne y productos cárnicos	Contrato menor	2.922,16€
Vestuario	18000000 - Prendas de vestir, calzado, artículos de viaje y accesorios	Contrato menor	6.133,50€
sustrato	24000000 - Productos químicos	Contrato menor	354,45€
sustrato	24440000 - Abonos diversos	Contrato menor	819,80€

No amount

Objeto +	Presupuesto	🛔 Tipo de procedi
Servicio de mantenimiento de semáforos.		
Sustitución de báculo de semáforo corroído en c/ Agustín Oliver	0,00€	Contrato menor
Ampliación semáforo	0,00€	Contrato menor
Servicios de asistencia técnica para redacción de los proyectos de regulación semafórica en zonas de alta ocupación, y de pasos de peatones inteligentes.	0,00€	Contrato menor
Instalación de grupo semafórico en la carretera N-550 en el entorno del colegio Outeiro das Penas, Cesantes, Redondela	0,00€	Contrato menor
suministro de adquisición de materiales para el montaje de semáforos	0,00€	Contrato menor
6 tubos metálicos enchufable de 32mm.+ 8 abrazadera -L (instalación cámaras	19,50€	Contrato menor

4.3.2 Example of good practice

In NextProcurement, the quality improvement of the enriched metadata (Solvencies, criteria, etc.) has been carried out, which has been explained in previous sections, since it was the object of the project itself.

4.3.3 Recommendations

- Implement automatic validations on recruitment platforms to detect errors in metadata entry.
- **Standardise procedure coding formats,** ensuring consistency in contract classification.
- **Require the inclusion of key metadata** in all published tenders, avoiding incomplete dossiers being published.
- Use artificial intelligence for data validation and enrichment, allowing the quality of metadata to be improved in an automated way.
- **Conduct regular audits on the quality of metadata** on public procurement platforms and take corrective action where necessary.

Contracting authorities must publish their tenders on a platform (that of the autonomous community or the national one).

Since it is mandatory, they should at least be provided with statistics/indicators services that allow them to see the data at their level.

And also, that the complete data of the tenders (not only those of the indicators already calculated) could be downloaded so that each agency can additionally make its own analyses.

4.4.1 Example of malpractice

Some common problems with agency-level indicators include:

• Information at the level of the contracting body

There is currently no service that allows you to view a performance analysis of procurement services at the contracting authority level. There are only a few statistical services, but the information is grouped at the national level.



• Downloading tender data

At the data download level, only the download of data for each graph is allowed and at the national level, which prevents each agency from independently evaluating itself using these services or doing its own analyses.

4.4.2 Example of good practice

NextProcurement has developed the Agency Viewer, where the EU procurement indicators are calculated (<u>https://single-market-scoreboard.ec.europa.eu/business-framework-conditions/public-procurement en</u>). This allows you to see the information not only of the contracting body itself without being able to compare it with other similar ones.

• Viewer at contracting body level

It measures key indicators in public procurement performance, in a transparent way that is easy to understand and compare. Like all indicators, they simplify reality and are affected by factors specific to each agency or by the type of tender itself. Therefore, indicators should be interpreted carefully, ideally in the light of additional quantitative and qualitative information. When moving the EU indicators from the country level to the contracting authority level, 3 indicators are not applicable (numbers 3, 4 and 12). Instead, 3 extra indicators have been defined, which in the application are named as IE1, IE2 and IE3

- Indicator 1: A single bidder (in open tenders)
- Indicator 2: traded or restricted
- Indicator 5: Award criterion only the lowest economic offer
- Indicator 6: Average decision period
- Indicator 7: SMEs awarded
- Indicator 8: SME tenderers
- Indicator 9. Batch Procedures
- Indicator 10: Call for tenders missing
- Indicator 11: Missing Successful Bidder's Registration Number
- Extra Indicator1: Deserted tenders
- Extra Indicator2: tenders withdrawn or renounced
- Extra Indicator3: Similar CPVs

O IDIOMA -										
Or Ver explicaci	ón									
	LICITADORES	5					ORGAN	ISMOS		
			-				_			
ganismo (NIF/DI	R3/Id Plataforma) A5037778	7 🗦 Da el nos	nbre del I	icitador	SOCIEDAD	MUNICIPAL Z	ARAGOZA	CULTURAL,	S.A.U.	
-	T Buscar por nombre		Mind	T BUSC	ar i		x umpi	41	_	
			verta	ucitation						
11	12 15 16	17 18	19	6 - C	110	111	*1	E1	**IE2	*1E3
Indicador 1	: un solo licitador (er	n licitaciones abierta	s)							
Verde 🖻	10% Muestra la propon	ción de licitaciones en los que l	ubo un s	olo licitad	or, excluye	ndo aq <mark>uello</mark> s	cuyo tipo	de procedi	miento no es a	abierto
Rojo	20% (Restringido, Nego	ciado sin publicidad, Derivado	de acuer	do marco)						
Situación o	del indicador									
		ndicador 1: un solo li	citado	r (en lic	itacion	es abierta	is)			=
Consejo d	e Administración de la Sociedad Mur	icipal Zaragoza Cultural, S.A.U.			0	0	0	50	0	v
	Vicepresidencia de la Sociedad Mur	iicipal Zaragoza Cultural, S.A.U.	2018	9	2020	2021	2022	46	2024	
			2010	2015	2020	2021	2022	2023	2024	1/1 🐨
						_			High	charts.com
Buscar:		Mostrar 10 🗸 regi	stros			🖽 Ver	las licita	ciones 💌	📩 Descar	gar 🔹
Mostrando	1 a 10 de 12 registros							Ant	erior 1 2 s	Siguiente
Año inicio					N° ‡ cu	licitaciones mplen el		N° licitaciones		\$
licitación	Organismo ID	Organismo			In	licador		total	Indicado	e) (
2018	A50377787, 51158690159629	Vicepresidencia de la Sociedad S.A.U.	Municipa	Zaragoza C	ultural,	0		2	05	16
2019	A50377787, 51158690159629	Vicepresidencia de la Sociedad S.A.U.	Municipa	l Zaragoza (ultural,	1		11	99	16
2020	A50377787	Consejo de Administración de l Zaragoza Cultural, S.A.U.	a Socieda	d Municipal		0		1	09	16
2020	A50377787, 51158690159629	Vicepresidencia de la Sociedad S.A.U.	Municipa	Zaragoza C	ultural,	1		10	10	96
2021	A50377787	Consejo de Administración de l Zaragoza Cultural, S.A.U.	a Socieda	d Municipal		0		2	05	16
	A50377787, 51158690159629	Vicepresidencia de la Sociedad S.A.U.	Municipa	l Zaragoza C	ultural,	28		56	50	96
2021						0		1	05	16
2021	A50377787	Consejo de Administración de l Zaragoza Cultural, S.A.U.	a Socieda	d Municipal						
2021 2022 2022	A50377787 A50377787, 51158690159629	Consejo de Administración de l Zaragoza Cultural, S.A.U. Vicepresidencia de la Sociedad S.A.U.	a Socieda Municipa	d Municipal I Zaragoza G	ultural,	37		64	58	96
2021 2022 2022 2023	A50377787 A50377787,51158690159629 A50377787	Consejo de Administración de l Zaragoza Cultural, S.A.U. Vicepresidencia de la Sociedad S.A.U. Consejo de Administración de l Zaragoza Cultural, S.A.U.	a Socieda Municipa a Socieda	d Municipal I Zaragoza C d Municipal	ultural,	37		4 64 2	58	96
2021 2022 2022 2023 2023	A50377787 A50377787, 51158690159629 A50377787, 51158690159629 A50377787, 51158690159629	Consejo de Administración de l Zaragoza Cultural, S.A.U. Vicepresidencia de la Sociedad S.A.U. Consejo de Administración de l Zaragoza Cultural, S.A.U. Vicepresidencia de la Sociedad S.A.U.	a Socieda Municipa a Socieda Municipa	d Municipal I Zaragoza C d Municipal I Zaragoza C	cultural,	37 1 26		64 2 57	58 50 46	96
2021 2022 2022 2023 2023 Recomend	A50377787 A50377787, 51158690159629 A50377787, 51158690159629 A50377787, 51158690159629 ador del indicador	Consejo de Administración de l Zangoza Cultural, S.A.U. Vicepresidencia de la Sociedad S.A.U. Consejo de Administración de l Zangoza Cultural, S.A.U. Vicepresidencia de la Sociedad S.A.U.	a Socieda Municipa a Socieda Municipa	d Municipal I Zaragoza C d Municipal I Zaragoza C	cultural, cultural,	37 1 26		64 2 57	58 50 46	96 96
2021 2022 2022 2023 2023 2023 Recomend 1. Mejora d	A50377787 A50377787, 51158690159629 A50377787, 51158690159629 ador del indicador e la publicidad y difusión: Ampl	Consejo de Administración de l Zaragoza Cultural, S.A.U. Vicepresidencia de la Sociedad S.A.U. Consejo de Administración de l Zaragoza Cultural, S.A.U. Vicepresidencia de la Sociedad S.A.U.	a Socieda Municipa a Socieda Municipa dar may	d Municipal I Zaragoza C d Municipal I Zaragoza C or visibilid	cultural, cultural, ad a las lici	37 1 26 itaciones abie	ertas, utili	64 2 57 izando tanto	58 50 46 medios digit	% % ales como

• Download tender data and indicators

•

At the data download level, the download of the following is allowed:

• All the tenders of the agency

		ORGANISMOS			
icitador	SOCIEDAD MI	UNICIPAL ZARAGOZA CULTURA	L, S.A.U.		
T BUS	CAR	× Limpiar			
las licita	iones				
		🛙 Ver mas campos disponit	oles 🔻	🛓 Descargar	•
KPI da	ata				



4.4.3 Recommendations

- Implement data visualization systems for each contracting agency.
- **Standardise key indicators** on competition, award times and participation of SMEs.
- **Publish this data in open formats** to encourage transparency and independent analysis.

4.5 Tenders and statistics at SME level

Small and medium-sized enterprises (SMEs) represent a fundamental pillar of the European economy.

However, their access to public procurement remains limited due to structural barriers and difficulties in identifying business opportunities. The lack of accessible information and detailed statistics on their participation in public tenders prevents them from being able to assess their competitiveness and design effective strategies to improve their access to the market.

Companies must consult the available tenders, including submitting their proposals on a platform (that of the autonomous community or the national one).

Since it is mandatory, they should at least be provided with statistics/indicators services that allow them to see the data at their level.

And also, that the complete data of the tenders (not only those of the indicators already calculated) could be downloaded so that each company can additionally make its own analyses.

Specifically, in the case of SMEs, there is a lot of emphasis on public policies, on facilitating their access to public procurement.

4.5.1 Example of malpractice

There is currently no public service that allows you to see an analysis of the performance of the bidding companies.

4.5.2 Example of good practice

NextProcurement has developed the Bidder Viewer, where the context of tenders for a company (history, competition, opportunities, etc.) can be analysed.

• Bidder-level viewer

A series of graphs and list of tenders have been generated, based on the activity of a company. The objective is to analyse the performance of the company in the tenders in which it has participated and facilitate access to others that could be of interest to it. This information is of particular interest to SMEs, for whom access to this type of information is costly and too complex to obtain.

- Graphs of won bids and amount
- List of tenders won
- List of competitors
- Generation of company profile based on your history



	LICITADORES		ORGANISMOS	
mpresa (CIF/NIF):	B99247553 Da el nombre del licita	ador		
	T Buscar por nombre	T BUSCAR	🗶 Limplar	
	Situación		Recomendador	
	Perfil de la en	npresa		
PVs:	72000000,72222300,			
Palabras clave:	SIUA			
	🎢 Autocalcular por historial	T BUSCAR		
Perfil. En plazo				
Nota. Los datos pueden te	iner 1 mes de retraso			
Buscar:	Mostrar 10 🛩	registros		🕹 Descargar 🔹
Objeto		Presupuesto	🛊 Estado 🍦 Limite	🗘 Plataforma 🍦
edjeto				
Cconceptualització dels p Comunicació Digital de l'	irojectes webs i Progressive web apps que determini la Direcció de Vjuntament de Barcelona	863.374,12€	Anuncio Previo	*
Cconceptualització dels p Comunicació Digital de l' Servicios TI: consultoría,	vrojectes webs i Progressive web apps que determini la Direcció de Njuntament de Barcelona desarrollo de software, internet y apoyo, 2021	863.374,12€ 4.000.000,00€	Anuncio Previo Anuncio Previo	*
Cconceptualització dels p Comunicació Digital de l' Servicios TI: consultoría, Soporte técnico, almacer	projectes webs i Progressive web apps que determini la Direcció de Njuntament de Barcelona desarrollo de software, internet y apoyo, 2021 amiento y uso de Blackboard Learn	863.374,12€ 4.000.000,00€ 110.078,65€	Anuncio Previo Anuncio Previo Anuncio Previo	•

Thus, SMEs are allowed to have access to this information at the same level as larger companies usually have, but free of charge, so that they can better understand the context of their tenders of interest, know who their competition is, etc., to be eligible to apply for future tenders.

Download tender and competition data •

At the data download level, the download of the following is allowed:



All bids from the bidder

4.5.3 Recommendations

Buscar:

Perfil. Competencia

Provide analytical tools that allow SMEs to assess their competitiveness in the • public sector.

T Ver las licitaciones

🕹 Descarga

Publish detailed statistics on the participation of SMEs in tenders.

Mostrar 10 v registros

• Encourage transparency in the awarding of contracts, ensuring that small businesses have equal opportunities.

4.6 Implementing APIs for Data Access

Access to public procurement data in an efficient and structured manner is key to fostering transparency and innovation in the sector. However, many recruitment platforms do not yet offer open APIs that allow companies, researchers, and citizens to interact with information programmatically. Open APIs can improve data accessibility, facilitate the development of advanced analysis tools, and enable the creation of new services based on the reuse of public information.

Currently, the lack of open and well-documented APIs represents a significant barrier to process automation, data integration, and the development of technological solutions based on artificial intelligence and data analytics.

4.6.1 Example of malpractice

On most public procurement platforms, data is only available through bulk downloads in inflexible formats, such as compressed XML or CSV files. These formats present several problems:

- They do not allow real-time queries.
- They do not facilitate integration with other databases or analysis systems.
- They require users to download and process large volumes of data, even if they only need a specific piece of information.

In PLACE, open data is only offered in the form of a download (atom or zip). But the format is very complex, so much so that there is a 250-page explanatory document explaining the structure.



In any case, this way of publishing the data requires downloading all the national tenders, when perhaps you only want an analysis, for example, of a City Council.

	O 2025 - Enero 🖾
	0 2025 - Febrero 🗵
	o 2025 - Marzo 🗷
•	Año 2024 - Archivo anual 🖾
	Año 2023 - Archivo anual 🗵
0	Año 2022 - Archivo anual 🗵
	Año 2021- Archivo anual 🖾
•	Año 2020 - Archivo anual 🛛
	Año 2019 - Archivo anual 🗵
0	Año 2018 - Archivo anual 🗵
	Año 2017 - Archivo anual 🗵
•	Año 2016 - Archivo anual 🖾
	Año 2015 - Archivo anual 🗵
0	Año 2014 - Archivo anual 🗵
	Año 2013 - Archivo anual 🗵
•	Año 2012 - Archivo anual 🗵
Car pat	da mes de año en curso se publicará un nuevo fichero comprimido con las actualizaciones del mes anterior siguiendo el rón:
http	ss://contrataciondelsectorpublico.gob.es/sindicacion/sindicacion_643/licitacionesPerfilesContratanteCompleto3_AAAAMM.z

4.6.2 Example of good practice

In NextProcurement, an API has been implemented that allows access to public procurement data in a structured and efficient way.

For example, a company interested in tenders for technology services can perform a specific query to the API to obtain only the relevant opportunities, without the need to manually process large volumes of data.

In NextProcurement, all the data displayed in the Search and Viewer services is APIbased.

Swagger. sum_series	Explore		
Nexprocurement API Control Con		NP's Solr Service API	fur illuia information
lanımı. Tatha Armanever əsəqqəndə confightipliken/tincovenent (1,8 - Seaqqəndə APA Avia Nocling 🔍			
public Operations openly available to public	^	CONFECTIONS senser conversion operations (conections creation and betrach, queries, etc.)	~
CET /companies searches companies inventory	~	Corpora Corpora-related operations (i.e., indexodelete corpora))	~
GLT /companies/(id) exhect company information	~		
GET /place searches place tenders (moldes) investory	~	Models Models-related operations (i.e., indepublicate models))	Ý
GET /place/(4d) estad tenderinformation	~	Oueries Specto Ser austes.	~
GL1 /place/iefe entred RUCE summary data	~		
CRT /place_memores southorplace londers (wines) investory	~	GET /queries/getCorpusNetadataFields/	
GRT /place_memores/(id) whistimoderinformation (minor()	~	QET /queries/getDocsRelatedToWord/	
GG3 /place_merces/info withd PLACE (minor) summay data	~	0ET /queries/eptDocsSimilarToFreeTextEmb/	
GET / ppv searches CPV ids and descriptions	~	Land Maria and Be constraints and a constraint	
GRT /cpv/(id) enterCPV internation	~	CET /queries/getDocs5imilarToFreeTextTN/	
GBY /cpv/Sefe what CPV	~	GET /queries/getDocsWithHighSimWithDocByid/	
Schemas	^	GET /queries/getDocsWithString/	
CompanyList >		GET /queries/getWetadataDocById/	
		CET /aueries/cetMadelInfo/	

The main API provides the complete set of tenders (about 4 million). And it returns the information within reasonable periods (taking into account the overall volume and the volume that is brought in each consultation).

This API can also be called specifying the fields you want to bring (out of the approximately 100 that exist), which is optimal to adjust it to the needs of different types of queries.

In addition to the use in NextProcurement in the Search Engine and Viewer, during the project they have been integrated into the tools of a partner, such as the Zaragoza City Council, as shown in the following images:

• Bidding company queries

Generalitat de Catalunya gencat.cat	
Licitación electrónica	
nextProcurement	
Cerca consultora	
Consultar empreses Consultar CPV Consultar Pla	
	Empreses
H a	123456789 H
id	Nom
21141202q	myb consultora lebaniega
a14069298	consultora tecnológica
a28215937	consultora actuarial pens mapfre vida
a50142041	consultora aragonesa de ingenieria
a78134004	consultora de riegos
a79389698	gesinca consultora de pensiones y seguros
a83263772	tinsa consultora
a84684406	consultora de energias renovables conersa
b01377118	civilarch consultora
b01582709	atari consultora sostenible
н	1 2 3 4 5 6 7 8 9 🕨 M

• CPV code queries

Generalitat de Catalunya gencat.cat			0 0 0 0
Licitación ele	ectrónica		
nextProcurement			
Cerca	telecomunicaciones		
Consultar empreses	Consultar CPV Consulta	r Place	
		0	PV
		H = 1	2 ► H
	id		Descripcio
32000000			Equipos de radio, televisión, comunicaciones y telecomunicaciones y equipos conexos
32412100			Red de telecomunicaciones
32500000			Equipo y material para telecomunicaciones
32510000			Sistema inalámbrico de telecomunicaciones
32520000			Equipo y cable de telecomunicaciones
32521000			Cable de telecomunicaciones
32522000			Equipo de telecomunicaciones
32523000			Instalaciones de telecomunicaciones
32524000			Sistema de telecomunicaciones
32533000			Estaciones terrestres de telecomunicaciones por satélite
		H 4 1	2 H

• Free-text search

Resultad	os		
Ver so v regi	stros por pagina	Buscar:	
Code	Titulo	Enticled	Fecha
3597520 mp00551631	Servicio de alconrende y administración da rendoma vinuella, eriopación y anglisición de recurso mantenes y ordenes para los Soltennas de Interinguna Tortales (ET) e entenetuccia da Dotos Turístico a suburiente con Windowa y SGL y las inductans vinuellas necesarias para las istemas elevantidade para las de do Besticos intergeneras, la meramenta de autovaluación y la prepa pagna web oe Destros turísticos intelligentes.	Presidencia y Dirección de Administración y Finanzas de la de la Socieciad Mercantil Estatal para la Gestión de la Innovación y las Tecnologías Turísticas, S.A.M.P. (SEGITTUR)	15 julio 2019
4358997 mtp00105493	Actualización del entorno de vinualización y almacenamiento de las máquinas virtuales	Gerente	
2709963 mp00522067	Ampliación del sistema de virtualización VI/WARE	Rectorado de la Universitat Politècnica de València	22 marzo 2019
8396392 ntp00239383	Subministrament d'un sistema de gestió de servidors virtuals, escriptoris virtuals (VDI), emmagattematge i punts de treball del Consell Constroil del Valles Occidental i del Consorti per a la Gestió de Resalas del Valles Concental	Consell Comarcal del Vallès Occidental	

4.6.3 Recommendations

- **Develop and implement APIs** across all public procurement platforms to improve access to data.
- Enable flexible queries while avoiding reliance on bulk data downloads.

5 Other recommendations to make the most of NLP and AI technologies

This section proposes a series of recommendations in addition to those made in NextProcurement related to NLP and AI that could be considered in other projects in the future.

NLP and AI technologies applied to NextProcurement have demonstrated their potential to improve public procurement, but their adoption still faces technical, regulatory, and organisational challenges. To maximise their impact, it would not be enough to apply AI in specific tasks as we have developed at NextProcurement, but it would be necessary to develop an integrated ecosystem that leverages these technologies for additional tasks.

5.1 Repository of processed tender documents for AI training

- Provide structured datasets to improve NLP models in public procurement.
- Guarantee the publication of models under open licenses.

5.2 Repository of pre-trained AI models for public bidding

- Centralise optimised language models for public procurement.
 - Promote the use of multilingual models to ensure fair access across the EU.

5.3 Generation of tender models based on previous tenders

- Use AI to create optimised draft sheets based on previous success criteria.
- Facilitate the standardisation and improvement of contract drafting.

5.4 Risk analysis, detection of irregularities and discriminatory clauses

- Identify suspicious patterns in awards and restrictive clauses in specifications.
 - Create early warning systems to mitigate risks in public procurement.
 - Assess potential breaches and contractual issues prior to awarding.
 - Detect clauses that may generate conflicts or cost overruns.

5.5 Automatic generation of bid summaries

- Apply NLP to extract the most relevant information and reduce the reading load.
- Facilitate quick access to key aspects of each tender.

5.6 Development of virtual assistants for the tender

- Create AI-based chatbots and assistants to assist bidders in the interpretation of specifications.
- Automate responses to frequently asked queries about hiring processes.
- 5.7 Predictive analysis of concurrency and trends in hiring
 - Apply AI to predict the participation of companies in future tenders.
 - Identify opportunities to improve competition in certain sectors.

5.8 Development of recommendation systems for SMEs

- Help small businesses identify relevant tenders based on their profile.
- Reduce barriers to entry through personalised recommendations.
- 5.9 Development of recommendation systems for contracting bodies
 - Suggest hiring strategies based on historical data and best practices.
 - Optimise the structuring of specifications according to market needs.

5.10 Automatic Frequently Asked Questions (FAQs) Generation for Tenders

- Apply NLP to identify recurring questions and generate automated responses.
- Facilitate the understanding of the requirements and procedures for bidders.

6 Conclusions

The implementation of Natural Language Processing (NLP) and Artificial Intelligence (AI) technologies in public procurement represents an opportunity to improve the efficiency, transparency and accessibility of data and therefore, in market competitiveness.

However, the difficulty of extracting data from official tender documents, low data quality, missing data, data errors or incorrect data formatting, is one of the main challenges faced by all previous initiatives when analysing public procurement data.

Since the quality of the knowledge derived will depend on the quality of the data entered, it is essential to improve the quality of the data. The higher the degree of data quality and automation, the newer technologies, such as AI and NLP, can take advantage of:

And this is especially important for smaller entities involved in public procurement (SMEs and local authorities), which have the least access to advanced tools such as AI and NLP. These emerging technologies can expand the competitiveness of SMEs by facilitating access to the public procurement market. It can also help smaller public administrations to improve their tenders, seeking the greatest optimisation of the process. In other words, while large companies or large public bodies have easier access to emerging technologies such as AI and NLP, it is the SMEs and smaller corporations that must be supported the most so that the market is balanced. For this reason, the implementation of these advanced technologies should be done on global platforms, to be on an equal footing.

Throughout this paper, a number of areas have been identified where these technologies can provide significant value, from optimising the search for tenders to improving metadata.

6.1 Main findings

1. Deficiencies in data quality and accessibility

- Multiple problems have been identified in the way public procurement data is published and structured, making reuse and automated analysis difficult.
- The lack of homogeneous standards in the identification of bidders and contracting bodies creates difficulties in analysing competition and the transparency of processes.

2. Lack of interoperability between recruitment platforms

• There are discrepancies between data published on different platforms, creating barriers to a unified analysis of public procurement.

 The implementation of open APIs and data standards would allow for better integration between systems and facilitate the reuse of information.

3. Need for greater transparency in information on bidders

- The publication of all participating bidders, together with unique identifiers, is essential to improve competition and avoid irregularities in the award processes.
- Currently, bidders' data is incomplete, inconsistent, and does not allow for proper analysis of market dynamics.

4. Importance of metadata validation

- The lack of rigor in the structuring of metadata negatively affects the quality of the information published.
- Implementing AI-based validation processes can help improve data accuracy and consistency.

5. Need for advanced information search and extraction tools

- CPV-based search is limited and does not respond to the real needs of users. Incorporating search by object (by text and semantics) can significantly improve information retrieval.
- Automatically extracting data from PDF documents remains a critical challenge, but the use of AI can help structure information more effectively.

6.2 Summary of key recommendations for the future

- 1. Adoption of NLP and AI technologies in all phases of the public procurement process
 - Automate the extraction and validation of metadata from tender documents.
 - Implement semantic search engines that allow more intuitive and accurate queries.

2. Strengthening data interoperability and standardisation

- Develop open APIs that allow the integration of public procurement systems with external analysis tools.
- Standardise the identification of tenderers and contracting bodies at European level by means of standardised unique identifiers.

3. Improvements in data accessibility and reuse

- Ensure that all recruitment data is accessible in open and structured formats.
- To promote the development of visualization and analysis tools so that both administrations and companies can better interpret the available information.

4. Increased transparency in the publication of information on bidders

- Establish the obligation to publish the complete list of bidders participating in each process.
- Use technologies to unequivocally link each company to its official records.