

## XML-CREPČ (verzia 21/003)

Zmeny v štruktúre  
pracovná verzia

## Účel a cieľ

Dokument vznikol na základe vývojového plánu v projekte CREPČ2 ako rozdielový dokument obsahujúci špecifikáciu úprav štruktúry s kódovým označením XML-CREPC. Dokument pre CVTI SR vypracoval dodávateľ, spoločnosť SVOP spol. s r.o. a je určený pre všetky subjekty, ktoré s formátom XML-CREPČ akýmkoľvek spôsobom pracujú.

## Dokumentácia a rozdeľovanie

Dokument podlieha starostlivosti o projektovú dokumentáciu systému manažérstva kvality.

Archív všetkých verzií dokumentu je uložený elektronicky v archíve spoločnosti.

Zadávateľovi bude dokument zaslaný elektronicky.

## Zmeny

**Zmeny v tomto dokumente môže robiť len autor dokumentu.**

Návrhy na zmeny dokumentu môžu podať zástupcovia dodávateľa aj zadávateľa písomnou formou, následne budú odsúhlasené zmeny zapracované do tohto dokumentu.

	Funkcia	Meno a priezvisko	Podpis
Vypracoval	Vývojár	Ing. Peter Matkulčík	
Preveril a schválil	Projektový manažér	Mgr. Ján Grman, PhD.	

Typ dokumentu:	Riadený dokument	
Dátum vydania	Dátum revízie	Dátum účinnosti
04.06.2021	-	04.06.2021
Číslo vydania	Číslo revízie	
1	1	04.06.2021
Poznámka:		

## OBSAH

Zmeny v štruktúre XML-CREPČ (verzia 21/003).....	1
Účel a cieľ.....	2
Dokumentácia a rozdeľovanie.....	2
Zmeny .....	2
1. Zoznam skratiek.....	4
2. Zoznam zmien.....	4
2.1 Rekapitulácia základných faktov.....	4
2.2 Evidencia smernice pre vykazovanie kategorizácie.....	5
2.3 Verzia schémy pre entity.....	7
2.4 Typ biblio formulára.....	8
2.5 Pozvaný príspevok.....	9
2.6 Študijný odbor.....	10
2.7 Zápis ID CREPČ nahrádzaných entít do nahradzajúcej.....	12
2.8 Zmena referenčného elementu Remark.....	16
2.9 Zmena kódovnikových hodnôt pre typ dokumentu.....	17
3. Zoznam príloh.....	18
3.1 Kódovnikové hodnoty pre typ dokumentu.....	18

## 1. Zoznam skratiek

Skratka	Hodnota
CREPČ	Centrálny register evidencie publikačnej činnosti
CVTI SR	Centrum vedecko technických informácií (zadávateľ)
SVOP	SVOP spol. s r. o. (dodávateľ)

## 2. Zoznam zmien

### 2.1 Rekapitulácia základných faktov

Umiestnenie schémy:

<https://www.crepc.sk/schema/crepc2.xsd> (po nasadení schémy do ostrej prevádzky)

<https://www.crepc.sk/schema/crepc2-21003.xsd> (schéma podľa tohoto dokumentu)

Použitia schémy:

- Interný formát záznamov systému CREPČ
- Výmenný formát v systémoch CREPČ OAI PMH a OpenSearch

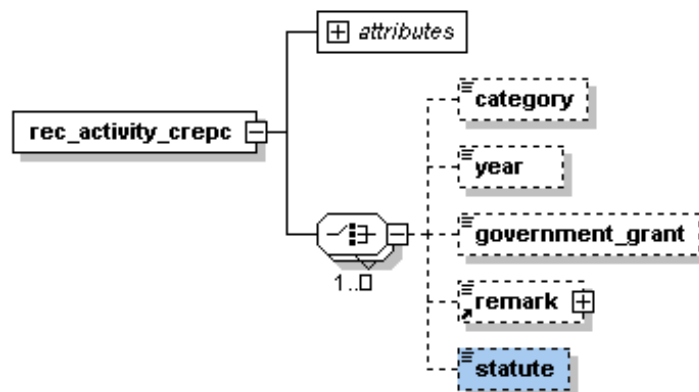
## 2.2 Evidencia smernice pre vykazovanie kategorizácie

Zmeny:

1. Doplnený nový typ vyhlášky pre vykazovanie kategorizácie – kód: **397/2020**
2. Zmena elementu **statute** na atribút **@legislation**. Presun z opakovateľnej položky `cross_biblio_activity` do atribútov `rec_biblio`. Atribút slúži na označenie legislatívy/vyhlášky, podľa ktorej je biblio záznam uložený a validovaný.
3. V systéme CREPČ budú koexistovať záznamy rôznych vyhlášok. Je dôležité vedieť, že záznamy nebudú unifikované na najnovšiu vyhlášku. Údaj je preto pri spracovaní dôležitý, najmä s ohľadom na sémantiku niektorých údajov (napr. kategorizácia).

Pôvodná položka: `rec_biblio/cross_biblio_activity/rec_activity_crepc/statute`

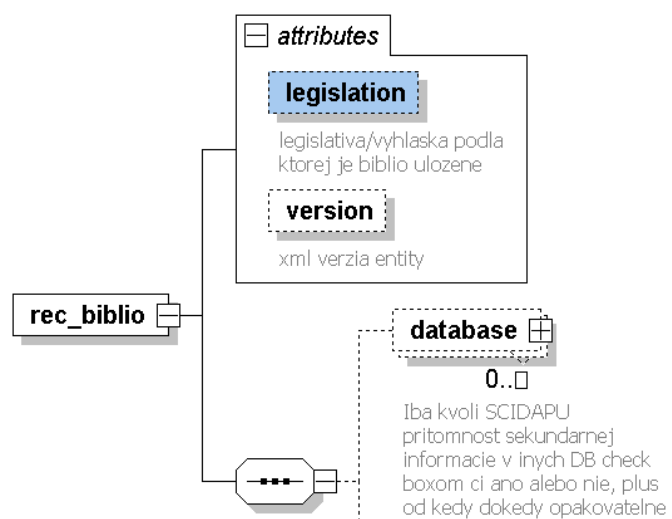
```
<xs:element name="statute" minOccurs="0">
  <xs:simpleType>
    <xs:restriction base="xs:string">
      <xs:enumeration value="13/2008-R"/>
      <xs:enumeration value="456/2012"/>
      <xs:enumeration value="397/2020"/>
    </xs:restriction>
  </xs:simpleType>
</xs:element>
```



Nové umiestnenie informácie: `rec_biblio/ @legislation`

```
<xs:element name="rec_biblio">
  ...
  <xs:attribute name="legislation">
    <xs:simpleType>
      <xs:restriction base="xs:string">
        <xs:enumeration value="none"/>
        <xs:enumeration value="13/2008-R"/>
        <xs:enumeration value="456/2012"/>
        <xs:enumeration value="397/2020"/>
      </xs:restriction>
    </xs:simpleType>
  </xs:attribute>
  ...
```

</xs:element>



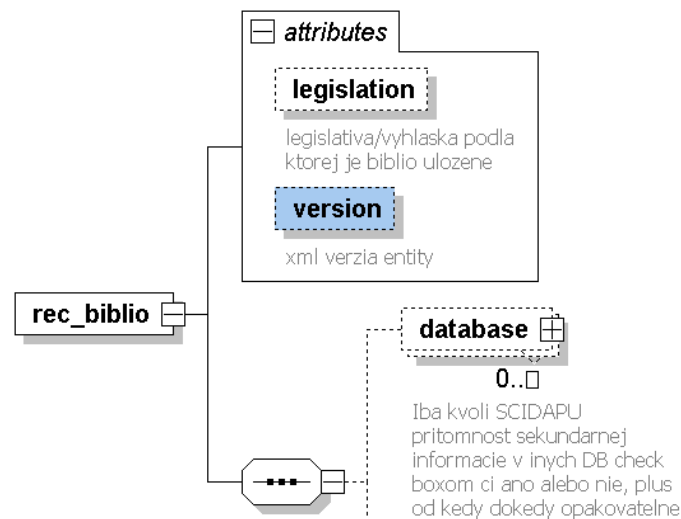
### 2.3 Verzia schémy pre entity

Doplnené označenie každej entity verziou, pomocou atribútu **@version**. Určuje verziu pod akou sú uložené jednotlivé entity. Ak sa verzia entity nezhoduje s aktuálnou verziou schémy, entitu je potrebné pregenerovať. Zmeny verzií budú oznamované a popísané v dokumentoch ako je tento. Verzia schémy v písomnom styku sa označuje rokom (21 ako 2021) a číslom verzie v roku na 3 miesta, oddelené znakom "/". V XML schéme a prípadne v databázových položkách sa verzia uvádza ako celé číslo. Toto číslo vznikne vynechaním lomítka. Aktuálna verzia „21/003“ má teda v číselnej reprezentácii hodnotu 21003 (dvadsaťjedentisíc tri).

Položka je pre spätnú kompatibilitu označená ako nepovinná, nové implementácie schémy ju však budú povinne vyplňať pri elemente najvyššej úrovne (top-level). Oficiálne verzie (release) budú označované

Atribút (typ: xs:int, nepovinné):

- **rec\_biblio/@version**
- **rec\_person/@version**
- **rec\_institution/@version**
- **rec\_project/@version**
- **rec\_meeting/@version**
- **rec\_database/@version**
- **rec\_edition/@version**

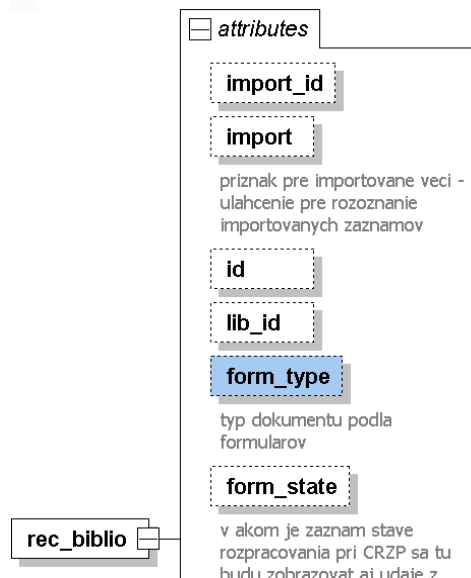


```
<xs:element name="rec_biblio">
...
  <xs:attribute name="version" type="xs:int">
  </xs:attribute>
...
</xs:element>
```

## 2.4 Typ biblio formulára

Z popisu atribútu **form\_type** boli odstránené názvy dokumentov a presunuté do enumerácií. Bol doplnený nový typ biblio formuláru Knižná publikácia – kód: **formBookPublication\_conf.xml**

Položka: **rec\_biblio/@form\_type**



```
<xs:attribute name="form_type">
  <xs:simpleType>
    <xs:restriction base="xs:string">
      <xs:enumeration value="formCasopis_conf.xml"/>
      <xs:enumeration value="formClanok_conf.xml"/>
      <xs:enumeration value="formBookPublication_conf.xml"/>
      <xs:enumeration value="formMonografia_conf.xml"/>
      <xs:enumeration value="formPoster_conf.xml"/>
      <xs:enumeration value="formPrezentacia_conf.xml"/>
      <xs:enumeration value="formPrispevokZbornik_conf.xml"/>
      <xs:enumeration value="formSprava_conf.xml"/>
      <xs:enumeration value="formZbornik_conf.xml"/>
      <xs:enumeration value="formZbornikPer_conf.xml"/>
    </xs:restriction>
  </xs:simpleType>
</xs:attribute>
```

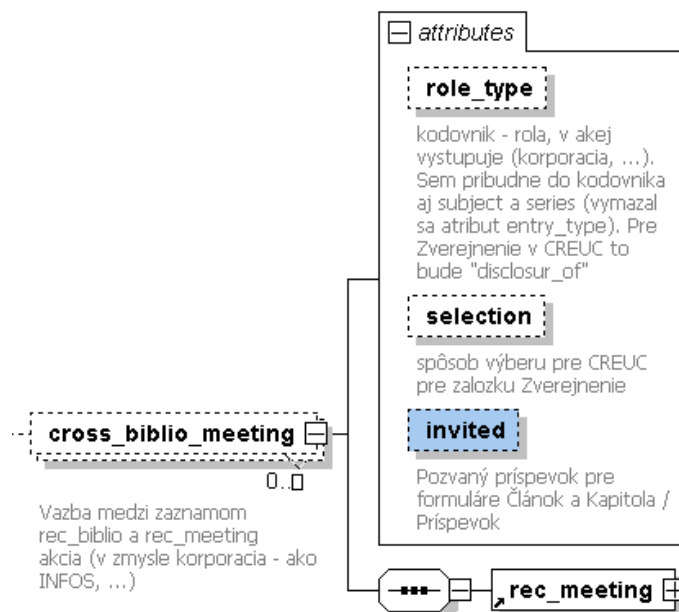


## 2.5 Pozvaný príspevok

Doplnené označenie konkrétneho podujatia atribútom **@invited**. Určuje či je daný záznam považovaný za pozvaný príspevok na danej akcii.

Atribút: **rec\_biblio/cross\_biblio\_meeting/@invited**

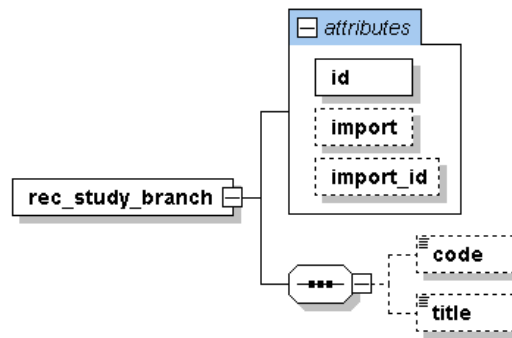
Typ: xs:boolean, nepovinné



```
<xs:element name="cross_biblio_meeting">
  ...
  <xs:attribute name="invited" type="xs:boolean"/></xs:attribute>
</xs:element>
```

## 2.6 Študijný odbor

Evidencia pre študijný odbor sa aplikuje vo viacerých základných entitách: biblio, osoba. V entitách bol použitý existujúci globálny element **rec\_study\_branch**. Poznámka: element a jeho definícia existovali v historických návrhoch schémy, nakoniec sa využili len v projekte CRZP. Do ostatných projektov sa dostávajú na základe zmien legislatívy až touto verziou a v nižšie uvedených kontextoch použitia.

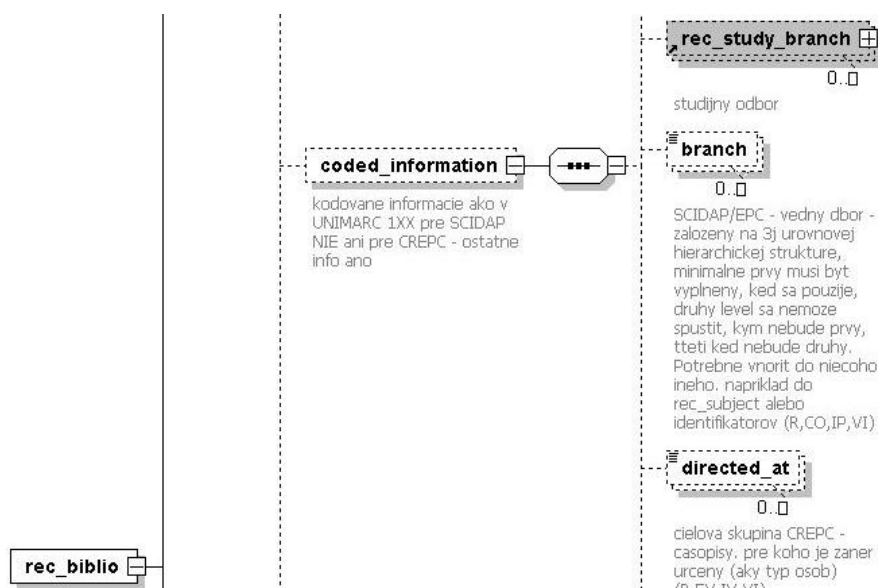


```

<xs:element name="rec_study_branch">
  <xs:complexType>
    <xs:sequence>
      <xs:element name="code" type="xs:string" minOccurs="0"/>
      <xs:element name="title" type="xs:string" minOccurs="0"/>
    </xs:sequence>
    <xs:attribute name="id" type="xs:int" use="required"/>
    <xs:attribute name="import" type="xs:string"/>
    <xs:attribute name="import_id" type="xs:string"/>
  </xs:complexType>
</xs:element>
  
```

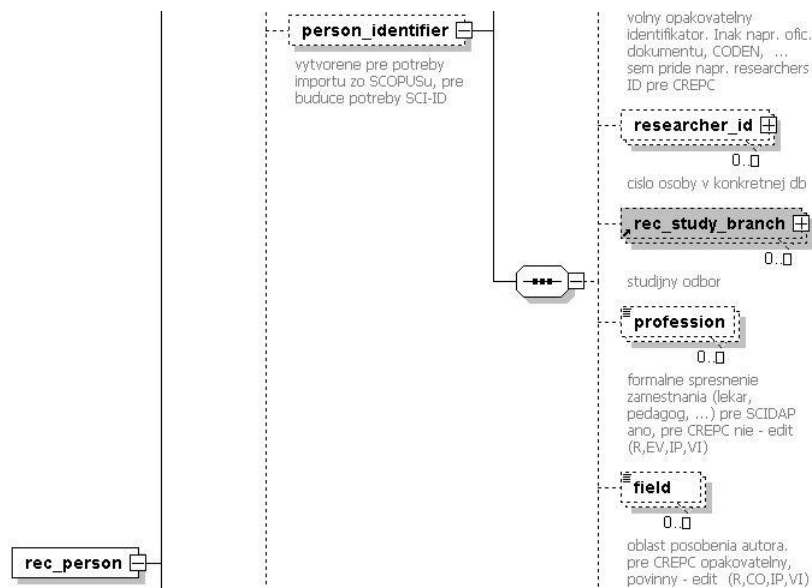
Kontexty použitia sú tieto:

- **rec\_biblio/coded\_information/rec\_study\_branch**



```
<xs:element name="coded_information" minOccurs="0">
  <xs:complexType>
    ...
    <xs:element ref="rec_study_branch" minOccurs="0"/></xs:element>
    ...
  </xs:complexType>
</xs:element>
```

- **rec\_person/person\_identifier/rec\_study\_branch**



```
<xs:element name="person_identifier" minOccurs="0">
  <xs:complexType>
    ...
    <xs:element ref="rec_study_branch" minOccurs="0"/></xs:element>
    ...
  </xs:complexType>
</xs:element>
```

V prípade bibliografického záznamu je študijný odbor povinnou položkou v niektorých kontextoch kategorizácie a typoch dokumentov. V prípade osoby položka dopĺňa metadátovej popis osoby o jeho aktivity (skills) v oblasti jej odborného pôsobenia v študijných odboroch.

## 2.7 Zápis ID CREPČ nahrádzaných entít do nahradzajúcej

Proces deduplikácie je kľúčovým prvkom pre čistenie a reguláciu kvality dát. Pri deduplikácií nenávratne zanikajú prebytočné informácie, niektoré informácie sa prenášajú (napr. externé identifikátory). Jedna z informácií, ktorá sa požadovala na zachovanie, je zoznam ID záznamov, ktoré boli deduplikáciu s daným záznamom zmazané. Informácia je dôležitá najmä pre vyhľadávanie informácií v prípade, že externý systém alebo dátový analytik disponuje referenciou na záznam pomocou ID záznamu a ten už vďaka deduplikácii neexistuje.

V systéme tak pribúda funkcia, ktorá v nahradzajúcej entite zobrazuje aj ID CREPČ pôvodných entít. Tieto ID CREPČ sú evidované ako väzba medzi záznamami (vzťah medzi záznamom ktorý zostal a zmazaným záznamom) s vyplneným atribútom **bond\_type= deduplicated**.

Kontexty použitia sú tieto:

- **BIBLIO - rec\_biblio/cross\_biblio\_biblio/@bond\_type**

```
<xs:element name="cross_biblio_biblio" minOccurs="0" maxOccurs="unbounded">
  <xs:complexType>
    <xs:attribute name="bond_type">
      <xs:restriction base="xs:string">
        ...
        <xs:enumeration value="deduplicated"/>
      </xs:restriction>
    </xs:attribute>
  </xs:complexType>
</xs:element>
```

- **OSOBA - rec\_person/cross\_person\_person/@bond\_type**

```
<xs:element name="cross_person_person" minOccurs="0" maxOccurs="unbounded">
  <xs:complexType>
    <xs:attribute name="bond_type">
      <xs:restriction base="xs:string">
        <xs:enumeration value="see_also_under"/>
        <xs:enumeration value="see_under"/>
        <xs:enumeration value="other"/>
        <xs:enumeration value="deduplicated"/>
      </xs:restriction>
    </xs:attribute>
  </xs:complexType>
</xs:element>
```

- **INŠTITÚCIA - rec\_institution/cross\_institution\_institution/@bond\_type**

```
<xs:element name="cross_institution_institution" minOccurs="0"
maxOccurs="unbounded">
  <xs:complexType>
    <xs:attribute name="bond_type">
      <xs:restriction base="xs:string">
        <xs:enumeration value="parent_child_level"/>
        <xs:enumeration value="parent_child_general"/>
        <xs:enumeration value="prev/next"/>
        <xs:enumeration value="parent_institution"/>
      </xs:restriction>
    </xs:attribute>
  </xs:complexType>
</xs:element>
```

```

                <xs:enumeration value="deduplicated"/>
            </xs:restriction>
        </xs:attribute>
    </xs:complexType>
</xs:element>

```

- **PODUJATIE - rec\_meeting/cross\_meeting\_meeting/@bond\_type**

```

<xs:element name="cross_meeting_meeting" minOccurs="0"
maxOccurs="unbounded">
    <xs:complexType>
        <xs:attribute name="bond_type">
            <xs:restriction base="xs:string">
                <xs:enumeration value="deduplicated"/>
            </xs:restriction>
        </xs:attribute>
    </xs:complexType>
</xs:element>

```

- **PROJEKT - rec\_project/cross\_project\_project/@bond\_type**

```

<xs:element name="cross_project_project" minOccurs="0"
maxOccurs="unbounded">
    <xs:complexType>
        <xs:attribute name="bond_type">
            <xs:restriction base="xs:string">
                <xs:enumeration value="deduplicated"/>
            </xs:restriction>
        </xs:attribute>
    </xs:complexType>
</xs:element>

```

- **DATABÁZA - rec\_database/cross\_database\_database/@bond\_type**

```

<xs:element name="cross_database_database" minOccurs="0"
maxOccurs="unbounded">
    <xs:complexType>
        <xs:attribute name="bond_type">
            <xs:restriction base="xs:string">
                <xs:enumeration value="parent_child_level"/>
                <xs:enumeration value="parent_child_general"/>
                <xs:enumeration value="prev/next"/>
                <xs:enumeration value="parent_institution"/>
                <xs:enumeration value="deduplicated"/>
            </xs:restriction>
        </xs:attribute>
    </xs:complexType>
</xs:element>

```

- **EDÍCIA - rec\_edition/cross\_edition\_edition/@bond\_type**

```

<xs:element name="cross_edition_edition" minOccurs="0"
maxOccurs="unbounded">
    <xs:complexType>
        <xs:attribute name="bond_type">

```

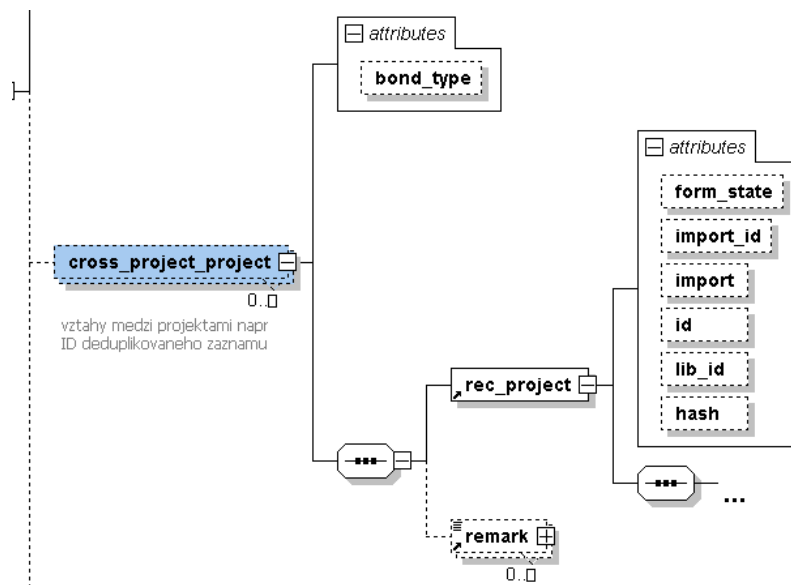
```

        <xs:restriction base="xs:string">
            <xs:enumeration value="deduplicated"/>
        </xs:restriction>
    </xs:attribute>
</xs:complexType>
</xs:element>

```

Z dôvodu neexistencie vnútorných väzieb (entita sa viaže na záznam rovnakého typu) pre entity projekt a edícia boli rozšírené štruktúry o príslušný element typu "cross".

Teda rec\_project o element **cross\_project\_project**

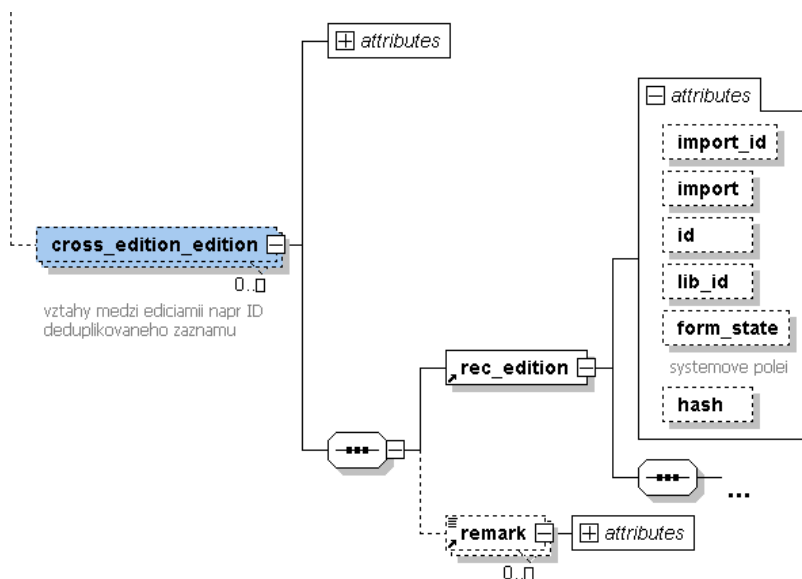


```

<xs:element name="cross_project_project" minOccurs="0"
maxOccurs="unbounded">
    <xs:complexType>
        <xs:sequence>
            <xs:element ref="rec_project"/>
            <xs:element ref="remark" minOccurs="0"
maxOccurs="unbounded"/>
        </xs:sequence>
        <xs:attribute name="bond_type">
            <xs:simpleType>
                <xs:restriction base="xs:string">
                    <xs:enumeration value="deduplicated"/>
                </xs:restriction>
            </xs:simpleType>
        </xs:attribute>
    </xs:complexType>
</xs:element>

```

a rec\_edition o element **cross\_edition\_edition**



```
<xs:element name="cross_edition_edition" minOccurs="0"
maxOccurs="unbounded">
  <xs:complexType>
    <xs:sequence>
      <xs:element ref="rec_edition"/>
      <xs:element ref="remark" minOccurs="0"
maxOccurs="unbounded"/>
    </xs:sequence>
    <xs:attribute name="bond_type">
      <xs:simpleType>
        <xs:restriction base="xs:string">
          <xs:enumeration value="deduplicated"/>
        </xs:restriction>
      </xs:simpleType>
    </xs:attribute>
  </xs:complexType>
</xs:element>
```

## 2.8 Zmena referenčného elementu Remark (poznámky)

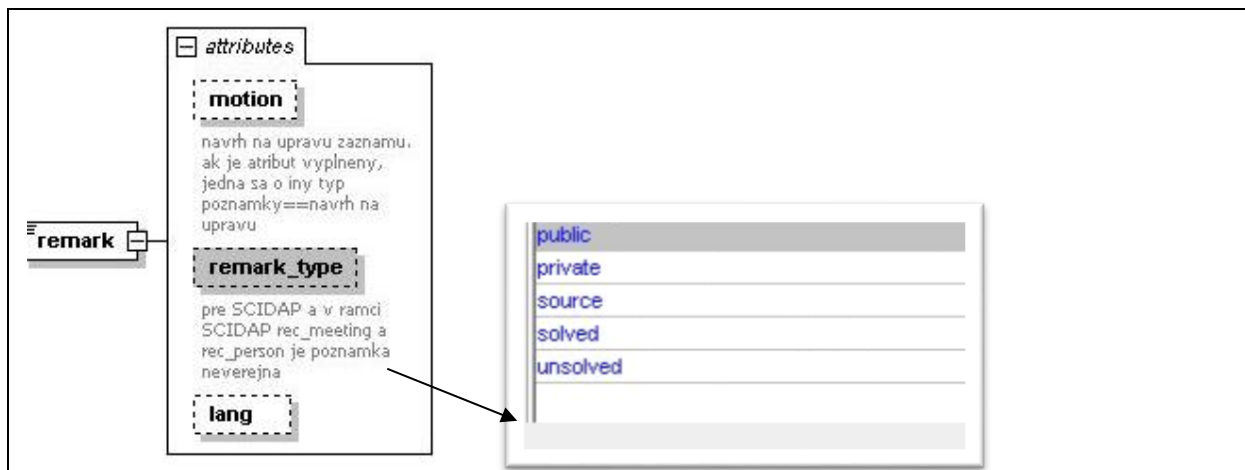
Referenčný element **remark** prešiel kompletnou úpravou. Dôvodom je rozširovanie sémantických typov poznámok s tým súvisiaci potreba vytvárať nové popisné kontexty. V atribúte **@remark\_type** sa miešalo viacero nesúvisiacich kontextov, teda miešalo sa označenie typu poznámky s jej stavom. Keďže počet kontextov sa zvýšil, bolo potrebné túto dualitu odstrániť a urobiť transparentnejšie. Položka **@motion** bola zovšeobecnená na typ údajov a **@remark\_type** sa premenovala na stav údajov.

Element sa používa v každej entite, je potrebné teda venovať pozornosť zmenám v spôsobe čítania a pri systémoch, ktoré dáta XML-CREPC aj tvoria, zmeniť spôsob generovania poznámok do záznamu.

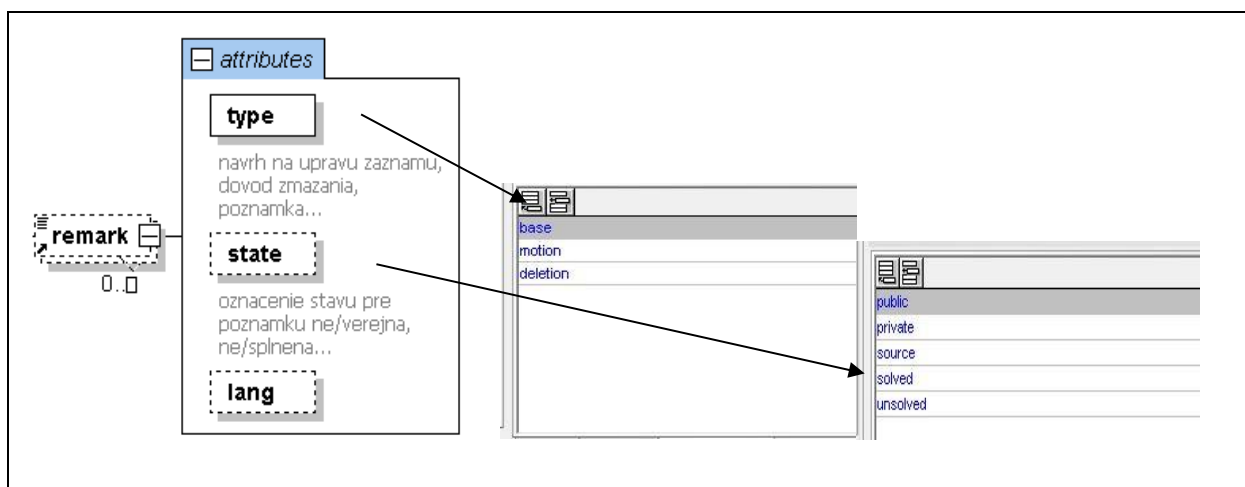
Zmeny v schéme pre element **remark** sú nasledovné:

- vznikli nové atribúty **@type** a **@state**
- pôvodný atribút **@motion** zanikol a je používaný ako enumerácia v atribúte **@type**
- pôvodný atribút **@remark\_type** sa premenoval na **@state**
- v **@type** pribudla nová enumerácia "deletion", ktorá slúži na popis dôvodu zmazania danej entity

Pôvodný referenčný element **REMARK**:



Upravený referenčný element **REMARK** (aj s enumeráciami):

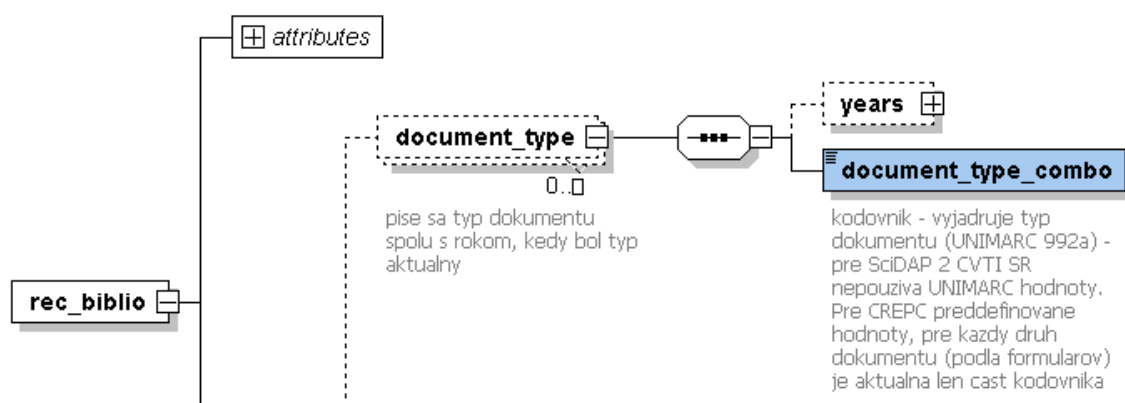




## 2.9 Zmena kódovnikových hodnôt pre typ dokumentu

V databázovom kódovníku typ dokumentu v biblio zázname došlo k viacerým zmenám (Príloha 3.1) položka: **rec\_biblio/document\_type / document\_type\_combo** :

- všetky používané kódy boli preložené do angličtiny (napr. kódy pre článok a časopis boli v slovenčine)
- rozlišovanie kódov pomocou postfixov (napr. kód esej existoval aj pre článok aj pre kapitolu)



```
<xs:element name="rec_biblio">
  <xs:element name="document_type" minOccurs="0" maxOccurs="unbounded">
    <xs:element name="years" type="periodType" minOccurs="0"/>
    <xs:element name="document_type_combo" type="xs:string">
  </xs:element>
  ...
</xs:element>
```

## 3. Zoznam príloh

### 3.1 Kódovnikové hodnoty pre typ dokumentu

Pôvodná hodnota	Nová hodnota
iný	other
časopis odborný	professional_magazine
časopis populárno-náučný	popular_science_magazine
časopis vedecko-odborný	scientific_professional_journal
časopis vedecký	scientific_journal
noviny	newspaper
webová stránka/portál	website_portal
abstrakt	abstract_article
abstrakt z podujatia	event_abstract_article
abstrakt zo stretnutia	meeting_abstract_article
bibliografia	bibliography_article
článok	article
diskusia	discussion_article
esej	essay_article
hlavný článok	main_article
hudobný záznam	music_recording
informácia o jednotlivcovi	individual_info_article
informácie, správy, oznámenia	information_article
konferenčný abstrakt	conference_abstract_article
konferenčný článok	conference_article
list	letter_article
materiál od redaktora	editorial_article
novinka	news_article
oprava	article_correction
poster z podujatia	event_poster_article
poznámka	comment_article
príspevok z kongresu	congress_article
príspevok z podujatia	event_article
príspevok zo zborníka	contribution_article
recenzia	review_article

rozhovor	interview_article
štúdia	study_article
úvodník, editoriál	introduction_article
výťah	summary_article
autobiografie, biografie	biography
bachelor_thesis	bachelor_thesis
diploma_thesis	thesis
dissertation	dissertation
doctoral_thesis	doctoral_thesis
encyklopédie	encyclopedic_monography
habilitation_thesis	habilitation_thesis
jubilejné publikácie	jubilee_publications
other	other_monography
populárno-náučné publikácie	popular_science_publication
ročanky	yearbook
slovníky (technické, výkladové, odborné)	dictionary_vocational
vedecké odborné publikácie	scientific_professional_publication
abstract	contribution_abstract
conclusion	contribution_conclusion
contribution	contribution
digest	contribution_digest
essay	contribution_essay
event_abstrakt	contribution_event_abstrakt
event_contribution	contribution_event
event_poster	contribution_event_poster
chapter	contribution_chapter
information	contribution_information
interview	contribution_interview
introduction	contribution_introduction
review	contribution_review
study	contribution_study
grantová správa	grant_report
priebežná správa	interim_report
projektové správy	project_report
správa o stave	status_report

správa z prieskumu	survey_report
štatistická správa	statistical_report
štúdie	study_report
technická správa	technical_report
výročné správy	annual_report
výskumné správy	research_report
záverečné správy	final_report
inštitucionálne (periodické)	institutional_periodic
jubilejní (pôvodných príspevkov)	jubilee_original_posts
konferenčné (periodické)	conference_periodic
konferenčné (pôvodných príspevkov)	conference_original_posts
odborné (neperiodické)	professional_non_periodic
tematické (periodické)	thematic_periodic
tematické (pôvodných príspevkov)	thematic_original_posts
vedecké (neperiodické)	scientific_non_periodic
ČSN	standard_csn
DIN	standard_din
EN	standard_en
EN ISO	standard_en_iso
ISO	standard_iso
STN	standard_stn
STN EN	standard_stn_en
STN EN ISO	standard_stn_en_iso
STN ISO	standard_stn_iso
patentová prihláška	patent_application
prihláška dizajnu	design_application
prihláška na udelenie šľachtiteľských osvedčení	breeders_certificate_application
prihláška ochrannej známky	trademark_application
prihláška označení pôvodu výrobkov	origin_of_product_application
prihláška úžitkového vzoru	utility_model_application
prihláška zemepisných označení výrobkov	geographical_indication_application
žiadosť o udelenie dodatkových ochranných osvedčení	supplementary_protection_application
annotated_interpretation	annotated_interpretation
annotated_translation	annotated_translation

anthology	anthology
architectural_study	architectural_study
art_catalog	art_catalog
book_publication	book_publication
cartographic_work	cartographic_work
dictionary	dictionary
didactic_manual	didactic_manual
dramatic_work	dramatic_work
edited_book	edited_book
encyclopedia	encyclopedia
fiction	fiction
high_school_textbook	high_school_textbook
learning_text	learning_text
monography	monography
other_publication	other_publication
poetry	poetry
primary_textbook	primary_textbook
proceedings	proceedings
review_work	review_work
sheet_music	sheet_music
short_stories	short_stories
skriptum	skriptum
source_edition	source_edition
university_textbook	university_textbook
workbook	workbook