Miroslav Novak[1]

METHODS OF LINEAR AND HIERARCHICAL SEQUENCES IN ARCHIVES

Abstract

There are more and more data collected and captured in the archival repositories and in the archival information systems and organized in different ways. Therefore, legitimate archival professional dilemmas and related discussions arise about ensuring their integrity. To solve this problem it is necessary to check the suitability of existing methods or develop and establish new methods of managing data in the archives.

The paper presents two archival methods for managing the archival data consistency and integrity on the base of linear and hierarchical sequences and the method of cardinality of sets of sequences. In archival professional procedures all three methods are already used in various forms. In some areas, their use is systematic, in other sporadic. They are often carried out by inertia and with the limited archival practical experience. The main intention of that paper is to present and theoretically ground these methods with intention of better and systematical implementation in everyday archival professional procedures.

Keywords: *methods, linear sequences, hierarchical sequences, cardinality, data in archives, archival material*

1 INTRODUCTION

In the framework of archival theory and practice, archival professionals have developed several ways and procedures for data management over time. Their solutions base primarily on the practical experience of manually manageable quantities of preserved archival material or related records about them.

A prerequisite for managing data in archives is formalized and established procedures. We can systemize these into the following basic processes of managing archival material:

- evaluation of creators and archival appraisal,
- arranging and describing archival records,
- performing examinations and the inventory of the entire archives, archival warehouses and archival records,
- implementation of procedures for relocation of archives,
- implementation of internal and external loan procedures,
- implementation of restaurant and conservation procedures including the replacement of technical equipment,
- carrying out procedures related to the dismantling and publishing of archival records,
- the implementation of procedures for the use of archival records for scientific, research and other purposes (ZVDAGA, Art. 53).

If the basic principles of the afore processes did not change significantly over time, then in practice we van conclude that the relations between the data generated in these processes and their creators and managers did not change significantly. Only few decades ago people (archivists) collected and processed data by themselves. Today many various devices generate and process data within archival professional processes.

1 Ass. Prof. Dr. Miroslav Novak, Pokrajinski arhiv Maribor, Glavni trg 7, SI-2000 Maribor
Tel.: + (386 2) 22 85 013, miro.novak@pokarh-mb.si

Individuals and machines manage data by implementing different methods and modes of processing. Some of these have been deliberately developed on the basis of the legitimacy of archival records and their contexts, while others appear "ad hoc" and are used to a limited extend regarding time and space. In this context, we should mention the methods of data and content classification (Šauperl, 2007), methods of producing reproductions (Novak, 2018a), scientific research methods in archival science (Semlič Rajh, Šauperl, Šabotič, 2013), etc.

2 PRESENTATION OF THE RESEARCH PROBLEM

The information contained in the procedures of archival professional work must be "accurate" and "credible". It is important to distinguish the "accuracy of the data"[2] from their "precision"[3]. Although both requirements are largely defined by the ratio between the amount of archival records and the available human resources in time and space, the accuracy and precision of data in practice are determined in a methodologically different manner. The accuracy of the data is determined based on the relationship between their incidence and the expected results. Precision is determined with the degree of detail of the definition of the entity under consideration. Different comparative methods in combination with abstraction, reduction, selection, etc. are used to determine the degree of accuracy and precision. (Novak, 2011).

From the point of view of archival practice, however, the use of the comparative method often depends on data previously obtained through various other formalized or non-formalized methods. In archival theory and practice, the method of determining the number of elements of a particular set or the power of the set or the cardinality of the set (Weisstein). Complementary to this, the procedures for the production of different lists and reviews of the entities of archival expert reviews are carried out. On this basis, different metadata are generated, which need to be compared with the long-term consistency of the entities of the sequence in question. That is why it is important to know the kind of data captured and the way of capture, as well as how they are preserved in time and space so that they can be compared with the results of analyzes in another time and space (Semlič Rajh, Šauperl, 2011). However, if we add to this the other activities and procedures implemented in the archival institution, especially in the field of managing the archival institution, human resources, professional and other procedures in the archives, time management and technological infrastructure in the archives, then we can understand the complexity of the whole data management in archives and related methods of capturing, checking, storing, evaluating and deleting data. Several methods and their practical implementations have been developed to carry out these tasks and to achieve goals in the field of data management in archives. Therefore, we will therefore limit ourselves to only three methods:

- the method of cardinality of a set of the sequence elements,
- the method of ensuring the consistency of linear sequence and
- the method of ensuring the consistency of hierarchical sequences in archives.

2 In this context, the accuracy is something that, when measuring something, shows as much as it really is. is completely consistent with the real situation (FRAN).

3 Natančnost pa upošteva, zajema vse, tudi podrobnosti oz. z največjo mogočo popolnostjo kaže, podaja resnično stanje Precision includes everything, also details or with the greatest possible perfection it shows the true state of affairs. (FRAN)

3 THE METHOD OF ENSURING THE CONSISTENCY OF LINEAR SEQUENCE IN ARCHIVES

The method consistency of linear sequences in archives we can define as the one that determines the accuracy of the elements of the sets of the linear sequence in accordance with the known parameters of these sequences.

A linear sequence in archives means that the entities of archival professional processes or states are placed in a row, so we know exactly where the entity is in this sequence. Sequences can be generated in the same or different time and space and in various combinations of both values.

Ensuring the consistency means respecting the same principles, flows or manifestations of entities in a particular order.

Each linear sequence is defined by the following entities:

- **the set or objects of a sequence** (physical[4] and logical[5] entities, content[6], elements[7], data[8], etc.),
- **starting point, begining or grabbing of linear sequence** (numeric, alpha or alpha-numeric value),
- **the direction o**f the sequence (rising, decreasing),
- a sequence **step** (the rate of increase in a sequence that can be defined by value as a value of 1 or greater than 1),
- sequence **consistenc**y (sequencing rules including the principles for the division of entities into a specific sequence) and
- the **power** or **cardinality** of the plurality of sequence elements. This determines the number of elements that the crowd contains.[9]

Linear sequences can be represented in different ways. Let's just mention:

- physically (for example, 10 archive boxes on one shelf),
- logically (for example, 100 records on description units at the level of the fond),
- in a textual, descriptive form,
- in graphic form,
- on the basis of known spatial data or time points, etc.

4 Examples of physical sequence entities include technical equipment: archival boxes, bindings, cords, etc.
5 Logical sequence entities appear at the level of the archival description within one level, for example, sequence of sub-series within one series.
6 In archival theory and practice the contents appear as integral parts of the wider completed whole. For example: chapter in the book, physically deferred archives on the basis of the original arrangement, etc.
7 In this context, let us just mention the elements of the inventory in the census forms, the sequence of columns in the tables or other presentations of archival material, etc.,
8 In this context, we should mention only a set of data that is classified according to a particular criterion or written in accordance with the known rule of formation of wider information units, e.g. when creating the address of the description unit.
9 Example of set X, that begins with 1, in first step we can put down as X = {1, 2, 3, 4, 5}. In this case the power of the set or the cardinality of the set F equal 5, which can be put down as m(F) = 5. (Množice. Retrieved 8. 2. 2019: http://www2.arnes.si/~mpavle1/mp/mnozice.html)

Figure1: A linear sequence model in graphic form with a starting point 1, a clearly defined sequence direction, a corresponding consistency, and known sequence steps, which in this case is a value of 1

The method of ensuring the consistency of a linear sequence we can realize through a process that we can define in six steps:

- 1st step: **Defining the sequence objects.** At this stage it is necessary to establish or define the legality of the objects of the sequence to correspond to the condition that the properties of a set of objects that form a sequence must be similar in at least one characteristic.
- 2nd step: **Defining the starting point or a begining point (SpS).** At this stage, it is necessary to establish the start of the linear sequence and consequently the end of this sequence.
- 3rd step: **Definition of the direction of the sequence.** At this stage, it is necessary to determine or ensure the direction of the linear sequence (vector) and define possible deviations in the form of parallel linear or pseudo-hierarchical sequences.
- 4th step: **Definition of the sequence step (Ss).** At this stage it is necessary to determine or ensure the sequence step value. This we usually define by an integer. Very often, the step is defined as Ss = 1. Because of the particularity of the individual procedures, this rule also deviates , since Ss = n + 1, where n≥0 or n≤0 can also be, so the value of this step should be specified in each sequence.
- 5th step: **Defining the sequence consistency.** At this stage it is necessary to determine the degree of deviation of the results of the first 4 phases and to determine the status of a "consistent" or "inconsistent" sequence. At the same time, check compliance with any specific requirements of the sequence or additional criteria, principles of the division of the sequence objects.
- 6th step: **Determining the power or cardinality (card) of a plurality of sequence (S) elements.** In this phase, we determine or define the actual power of the set card(S) of the linear sequence. The power of the set of sequence elements is defined by the number of elements the set contains.

If the method of ensuring the consistency of linear sequences is implemented on objects of archival professional activities, which in themselves can not be empty values, natural numbers need to be used for their labeling. In this context, let us mention some examples of the implementation of this method:

- in managing the sequences of technical units in archival warehouses,
- when disposing of library material to shelves on the basis of the current number,
- in the management of records on archives that are produced on the same level of enumeration,
- when naming files in accordance with ISO standard (ISO / TR 13028), etc.

Theoretically, the method is used wherever:

- it is necessary to arrange the set of elements into the expected linear sequence,
- it is necessary to determine the integrity of a particular linear sequence,
- it is necessary to establish the legality of the existing linear sequence and the occurrence of deviations of this sequence,
- it is necessary to transpose the elements of one linear sequence into elements of the second linear sequence.

An example of using the management of linear sequence data based on one-to-one transponder transposition is illustrated in the process of digitizing a large-format book, 150 pages. The pages in the book have a sequence (S) established. This is derived from the content and is usually defined by pagination. The order of the page can be defined by the page number, page order sequence and same point sequence. This can be written as follows:

S={1,2,3,4,5,6...150} with SS 1 and SpS in point 1, where card(S)=150.

In the process of digitization, for practical reasons, we first perform digitization, for example, odd and then even sides. After the digitization of the odd pages, the digitized odd pages (Dop) obtain the sequence:

Dop={1,2,3,4,5,6,...75} with Ss 1, SpS in point 1, where card(Dop)=75.

This is followed by digitization of even-numbered pages. Digitalization even pages (Dep) obtained sequence:

Dep={1,2,3,4,5,6,....75} with Ss 1, SpS in point 1, where card(Dep)=75.

Both sequences digitalizatov odd and even pages have exactly the same features and none of them shall not be liable to the original sequence of pages in a digitized book. Therefore, it is necessary to change, transpose, the legality of both sequences of digitizers.

The order of the digitizers of the odd pages should be transposed so that it becomes:

Dop={1,3,5,7,...149} with Ss 2, SpS in point 1, where card(Dop)=75.

The same is true for digitalizate even-numbered pages. They receive the following legality of the sequence:

Dep={2,4,6,8,..150 with Ss 2, SpS in point 2, where je card(Dep)=75.

After the digitalization of the even and odd pages, we combine the digitized files of the odd and even pages in one folder and sort them ascending, then we get the sequence:

Dop+Dep={1,2,3,4,5,6,7,8,...150}withSs1,SpSinpoint1,wherejecard(Dop+Dep)=150.

This corresponds to the condition of the page sequence of the original archival material, which can be defined as a bijective mapping of the set S → the set Dop + ep, since the condition is satisfied:

(S={1,2,3,4,5,6...n} with Ss 1, SpS in point 1) = (D ep+op={1,2,3,4,5,6...n} with Ss 1, SpS in point 1)

If we define the strength of the two sequences, then we can write down:

card(S) = card (Dop+Dep) or |S| = |Dop+Dep|=150

Using the above method, we can conclude that the digitization of the book was carried out correctly, in the object management segment throughout the entire process of converting one form of archival records into another.

Another example of the implementation of this method can be seen through the processes of transforming the signature list signature (Sg), e.g. at the level of the aggregation. It is known that at the record creator, the original linear sequence of cases with the production numbers Sg = {1,2,3,4,5,6,7,8,9, ... n} is created with step (Ss) 1 in the starting point of sequence (SpS) 1. After removing records of non-archival value in the process of transferrinf archival records to archives, the sequence of identification numbers or the case labels (SF) changes to Sg = {2,3,6,8..n} with Ss ≤1 and the SpS of sequence ≤1.

After removing records with no archival value, there is an unordered sequence of archival records. The markings or identification numbers of files of archival value (Sa) form a sequence with the following legality Sa = {1,5, 7,9, ... n} with Ss ≤1 and the SpS in ≤1. This sequence is no longer the same as the original linear sequence, so it needs to be transposed into a new sequence (Sn). The new sequence must have all the characteristics of the linear sequence Sn = {1, 2,3,4,5,6,7,8,9 ... n} with Ss 1, the SpS of sequence 1. It is expressed in the identification number of the description units of archival records as an injective maping[10] the sets Sa → Sn. Forming the identification number is actually a complex archival professional task based on the consideration of basic archival professional principles, since it is necessary to uniquely define the positions of description units within certain sequences, which is not always followed in practice (Novak, 2012). If the afore knowledge is generalized, it can be stated that by a method of ensuring consistency of linear sequences, and related determination of the power of the plurality of the linear sequence, the basic tool for managing data and their structures in the archives is obtained.

The method can be implemented in many basic archival professional processes and subprocesses in solving individual archival professional problems. Depending on the implementation method, it can be systematically supported with simple system tools that allow sorting or not.

4 THE METHOD OF ENSURING THE CONSISTENCY OF HIERARCHICAL SEQUENCES IN ARCHIVES

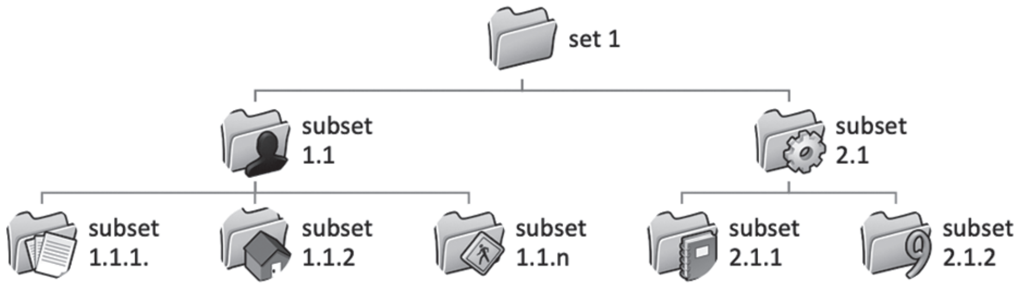
A hierarchical sequence in archives means that the entities of archival professional processes or states are placed in a structure in this way to know the exact position of the entity in the structure. The consistency of a hierarchical sequence, by analogy to linear sequences, means ensuring compliance with the same principles of events or manifestations of entities in a particular order of the structure. Hierarchical sequences are generated in time and space and are subject to the same rules as for linear sequences

Since a hierarchical sequence in some sense with a structure is the upgraded set of linear sequences, we can define this with the following entities:

- **sequence objects** (physical and logical entities, content elements, data, etc.)
- **starting point, begining or grabbing of** (numerical, alpha or alpha-numeric value) of a hierarchically organized set, including a plurality of complementary clasps of subsets od elements od a hierarchical sequence,
- **directions** of the sequence in the horizontal and vertical dimension (ascending - descending, wider - narrower) or as a precursor - successor (larger - smaller, etc),
- **consistency** of sequence (design rules of sequences including the principles of separation of entities in a particular location sequences), and
- a **two-dimensional sequence step** (defines the rate of increase or decrease in the sequence. The step is defined as the value 1 by the hierarchy and as the rate of increase or decrease in the sequence within the subset of the linear sequence by 1, in certain cases also for a value of more than 1 or less than 1, depending on the points and direction of the sequence).

10 An injective function or injection or one-to-one function is a function that preserves distinctness: it never maps distinct elements of its domain to the same element of its codomain. More about in:: Injective function. . Retrieved 8. 2. 2019: https://en.wikipedia.org/wiki/Injective_function.

Figure 2: A model of a hierarchical sequence in graphic form with a starting point in a set 1 defined by the top-down hierarchy direction and the linear sequence within the subsets of the corresponding consistency and the known steps within the sequence of the subset.



Hierarchical sequences can be the same as linearly presented in different ways. Let's just point out:

- in physical form (for example, 10 folders in one archive box),
- in logical form (for example, 1 record at the level of the fond, 10 records at the level of the series, 100 at the level of the aggregation),
- in a textual, descriptive form,
- in graphic form,
- based on previously known spatial data or time points, etc.

The method of ensuring the consistency of a hierarchical sequence is realized through a process that can be defined in six stages:

- 1st phase: **Defining the sequence objects.** In this phase it is necessary to determine whether the objects of the set of hierarchical sequences correspond to the conditions under which the properties of a set of objects that form the sequence structure are similar in at least one characteristic.
- 2nd phase: **Defining the starting point or a beginning point (SpS).** In this phase, it is necessary to determine the starting of the hierarchical sequence of the discussed set of objects of the hierarchical sequence, and the starting point of each of its subsets of the objects.
- 3rd phase: **Defining the direction of the sequence.** At this stage, it is necessary to determine the direction and type of the hierarchical structure of the sequence and the direction of the linear sequences within the subsets of the considered set, including possible deviations.
- 4th phase: **Defining the sequence step (Ss).** This is usually defined as the value of 1 in both the hierarchy and the horizontal within the sequence. Mostly: $Ss = n + 1$, where $n \geq 0$ or $n \leq 0$.
- 5th phase: **Defining the sequence consistency.** At this stage it is necessary to determine the degree of deviation of the results of the first 4 phases. At the same time, check compliance with any specific requirements of the sequence or additional criteria, principles of object division in structure.
- 6th phase: **Defining the strength of the structure or cardinality (card).** At this stage, we determine the actual power of the set and its subsets in a hierarchical structure.

The method of ensuring the consistency of hierarchical sequences in archival theory and practice is, as a rule, implemented on objects that in themselves can not be empty values, therefore they must be denoted in practice by natural numbers.

The method can be used to manage sequences of multi-stage records describing. Also to manage the system of marking technical units – archival maps in relation to archival boxes, to establish all types of relationships between objects in the form of decision trees at the level, for example, of organization of digital and physical storage locations, establishment of relations between descriptors, description and technical units, etc. Theoretically, the method is used wherever:

- arrange the necessary set of elements in the expected hierarchical order,
- to determine the integrity of a particular hierarchical sequence,
- to establish the legality of the existing hierarchical sequence and the occurrence of deviations in the sequence in question,
- transpose the necessary elements of one hierarchical sequence into elements of the second hierarchical sequence.

A typical example of the problem of arranging elements in an expected hierarchical sequence is the management of the sequence of levels of description (Ld) of archival records. In accordance with the ISADg standard, there are 6 levels of inventory defined: fond (F), sub-fond (sF), series (S), sub-series (sS), file (F) and documents (D). They are in the following hierarchical relationship:

$D \subset Zd \subset pS \subset S \subset pF \subset F$

In the archival information systems, the structure of the levels is defined, for example. with numbers of levels 1 to 6 in this way to represent the value of $1 \equiv F$, $2 \equiv pF$, $3 \equiv S$, $4 \equiv pS$, $5 \equiv Zd$ and $6 \equiv D$. In this way, we get a bijection mapping of the levels of descriptions into natural numbers.

Additionally, the following conditions or the following insertion rules are added:

1. Rule A: D1…n ϵ F;
2. Rule B: D1…n v F1…n ϵ sS
3. Rule C: F1…n v sS1…n ϵ S
4. Rule D: F1…n v sS1…n v S1…n ϵ sF
5. Rule E: Fd1…n v sS1…n v S1…n ϵ sF1…n ϵ F

Although relatively weak structure cardLd = 6, due to the established relationships and rules insertion, can be properly managed only by means of the algorithm used within the archival information system

Direct sorting of the linearized form of this hierarchical structure by title does not give the expected result.

Level structure by standard	Activity	The result of the classification of the levels of description
fond **sub-fond** **series** **sub-series** **aggregation** **documents**	sorting values	**documents** **fond** **sub-fond** **sub-series** **series** **aggregation**

A solution to this problem is possible, for example, by adding prefixes with the center of gravity from 1 to 6 in the name of the element, in a way that is disproportionate to the describing level. If we perform this bijective mapping, we get the following results:

fond ≡ [1] fond

sub-fond ≡ [2] sub-fond

series ≡ [3] series

sub-series ≡ [4] sub-series

aggregation ≡ [5] aggregation

document ≡ [6] document

Thus, when sorting the mapped linearized values of the levels of description, the corresponding sequences would be obtained without the intervention of a specific algorithm.

Fairly unsophisticated example shows the actual complexity of managing hierarchical relationships, especially on the problems of management of such sequences, especially if the rules and logic of their operation restricted to specific system solutions. On the basis of this, we can conclude that, in order to ensure the autarktility of such sequences, the data structure and presentation must be made in such a way that it does not depend on the source system. The appropriately designed identification number or tag of individual elements represent in this case a prerequisite for the consistent management of such sequences and their structures.

Another example of using data management with a hierarchical sequence is the labeling of archival boxes of specific archival fond of, for example, 10 technical units and 100 units of the technical equipment of description units (folders or maps) within archival boxes.

Archival boxes (Ab) are defined within the archival fond as a set of linear sequence entities.

Ab={1,2,3,4,5,6,7,8,9,10} with Ss 1, SpS in point 1

The equipment of description units (Edu) can, however, be theoretically sorted in two ways: as a set of elements of a linear sequence or as a set of elements of a hierarchical sequence.

If the set of linear sequences of the technical equipment of the description units is treated within the framework of the archival fond, then this can be recorded as:

Edu={1,2,3,4,5,6,7,8,9…100} with Ss 1 and SpS in point 1

In this case, we abstract the existence of a superior set of technical equipment - that is, archival boxes, and the defined sequence is not in accordance with the requirement that each subset starts with a point sequence 1.

Therefore, it is more correct for the sequence of technical equipment to be considered on the basis of the legality of the hierarchical sequences, according to which in each subset the point sequence is defined by 1. The starting point is that each element of the Edu is also a technical unit element[11], and that the power of each subset Edu1 … Edu10 is always 10. The first condition can write down:

Edu ϵ AŠ

If the set of elements correspond to this condition, then we obtain the following hierarchical order of the technical equipment of description units within archival boxes:

Edu1{1,2,3,…10} ϵAb1+Edu2{1,2,3,…10}ϵAb2+Edu3{1,2,3,…10}ϵAb3+… Edu10{1,2,3,…10} ϵAb10

This raises the question of how to calculate the power of a set of hierarchical sequences in archival theory and practice. If we proceed from the starting point so that each set of elements of hierarchical structures can be defined in a linearized form, then we can calculate the cardinality of the whole set of elements or the power of individual subsets or their combinations.

11 In archival practice this means that books (K), for example, which are deposited independently on the shelf, are not an element of archival boxes, which is written in the form K ∉ AŠ

The strength (cardinality) of a set of archive boxes is:

card(Ab1…Ab10) = 10 or |Ab1…Ab10| = 10

The strength of the set of the technical equipment of the description units is:

card(Edu1…Edu10) = 100 or |Edu1…Edu10| =100

The power of the set of descriptions of the technical equipment is:

card(Ab) + card(Edu)=110 or |Ab| + |Edu|=110

By introducing the method of consistency of hierarchical sequences and determining the power of the set of these elements, with methodologically thought-shaped data structures, we gain in archival practice a powerful tool for data management and their structures. This can be implemented in many archival professional processes, including in the phases of testing the planned solutions of individual archival professional problems. Depending on the implementation, the method can be supported by information system or not.

5 THE METHOD OF CARDINALITY OF A SET OF THE SEQUENCE ELEMENTS

The method of determining the cardinality of a set of elements is based on a comparison of at least two measured power sets of the same elements on the basis of biictivity and injectivity or using cardinal numbers. It can be calculated for the sequence of items that are located in the same time period and in the same place or not. If the difference between the powers of the comparing sequences is zero, then the sets are considered to be consistent, and therefore the entity of the treatment is consistent in this part.

As a typical example of the implementation of the cardinal method, we use an example of 10 archival boxes (Ab) of a fond based on the identification of the cardinal numbers of this set. The cardinality of the multitude of boxes was established at the time of the acquisition (T1) of the archival fond, after processing at a location (L1). When it is established that the arranged archival boxes follow the legality of the linear sequence with step 1 and the point sequence in point 1, this is written:

Ab={1,2,3,4,5,6,7,8,9,10} with Ss 1, SpS in point 1, in T1, L1

In the next step, the power of the sequence is calculated, which is

card(Ab1…Ab10) in T1,L1 = |Ab| in T1,L1 = 10

When the archival records are deposited on archival shelves, the consistency of the sequence at a different time and another environment is revisited based on a biative mapping.

Ab1={1,2,3,4,5,6,7,8,9,10} with Ss 1, SPS in point 1, in T2, L2

The power of the sequence is recalculated:

card(Ab1…10) in T2,L2 = |Ab in T2,L2| = 10

The difference in the power of the linear sequence of archival boxes at the time of the records transmission to archives (T1, L1) and the power of the sequence of archival boxes, after being deposited on the archival shelves, is 0:

[card(Ab1…10) in T1,L1] – [card(Ab1…10) in T2, L2] = 0

This means that the sequence in question is consistent in time and space. If we want to formalize the power of the considered sequence, this is done in the inventory of archival records in accordance with the ISADg2 standard, for example, at the level of the fond description in the element 3.1.5 Scope and medium of the description unit.

In the case of an inventory (at other times = T3) in the archival warehouse it is necessary to re-check the consistency of the sequence of these archival boxes with known linear sequence legality based on the bijective mapping. If we recalculate the power of the sequence, we find:

() in T3, L2 = 10

Then, the power of the sequence from T3 time is deducted from the power of the T2 sequence that is written in the archival description:

[card(Ab1...10) in T3,L2] – [card(Ab1...10) in T2,L2] = 0

The result of this operation is 0, which means that there were no changes in both sequences. By doing this, the consistency of this sequence is verified and confirmed over time. At the same time, we have also set up a model for its implementation to permanently determine the power of this sequence in time and space.

The method can also be used for complex data management activities, such as, for example, arranging a set of records on the description units with non-systemic tools, importing a linear set of descriptions of archival records into a database of the archival information system, etc. Each record has its own identification number (IDn), which from the point of view of the sequence provides the logic of the operation of this linear sequence, which can be written:

IDn = {1,2,3..n} with Ss 1, SpS in point 1 in T1, L1

At the same time, we can perform another operation, such as counting (CIDn) by using the system, or otherwise. In this case, we also obtain a linear sequence:

CIDn = {1,2,3..n} with Ss 1, SpS in point 1 in T1, L1

For each sequence, the power of the sequence is calculated and the countdown operation is performed on the values obtained:

[card(IDn1...n) T1,L1] – [card(CIDn1...n) T1,L1] = 0

If the result is 0, then we have shown that the number of records is consistent with the identification numbers used. However, the power difference between the two sequences is not always the same. The deviation can be in a positive or negative side. If the difference is positive, the increase in the elements of the crowd, e.g. to obtain archival records, to reorganize fonds and collections, or to modify the technical equipment used, or errors in creation of identification numbers or data capture, etc. However, if the gain is negative sequence, it can mean mislaying, in the extreme alienation or even destruction of archival records; however, changes may have occurred due to the implementation of other technical equipment or the reorganization of the archival records, as well as the errors in the design of descriptions or their labels. In any case, this is always an important alarm for archival professionals, because the sequence is inconsistent in time and space.

Another example of determining the cardinality of a set of sequence elements represents the archival records, which is not arranged and in the process of arranging we are giving them the identification number. In the process of arranging archival records, using the method of an injective mapping, we assign to the existing code numbers (eCn) the new teference numbers of the units of description (Rn).

eCn → Rn

The cardinality of the sets is determined on the basis of the cardinal number of the two sets (eCn) and (Rn).

The method is used in different time and space, for example in the case of various archival professional activities on objects that can determine the power of the sequence according to the state of power in a different time or space, or according to the two measured power of the masses of the same elements. Depending on the implementation method, the method can be systematically supported or not.

6 DISCUSSION

Presented methods for checking the consistency of a set of data structures and their elements are often used in archival theory and practice. Normally, this is part of the routine of individual archival professional procedures. They have been developed to provide basic data management needs in archival practice (Semlič Rajh, 2012) and are systematically implemented in some segments in the applicable legislation, standards, etc.

As an example to the legislation of the defined sequence, we should mention the creation of technical metadata, especially the technical unit designations. These primarily serve as information on the position in the linear sequence of the technical units of the entire archival fond. In archival information systems, these are often used as entry points for inquiries. Individual sequencer tags are typically created in the process of records arranging and formalized in the process of describing of archival records. Thus, for example, in Slovenia as a standard is used a numeric code for a technical unit (UVDAG, Articles 24 and 64), which is written on the archival plate in such a way that the data is also well visible in the context of limited visibility, for example in archival warehouses. At the same time, Article 23 of the UVDAG defines that the acceptance list should be included among other basic data also ... *the quantity of archival records, expressed by the number of technical units* - that is, with the power of a set of technical units of the entire collection of archival records respectively. In addition, the same article defines the elements of the takeover record, among which it also includes the designation of the technical unit. From the methodological point of view, the solution is a system of data management of technical units also for relatively complex sequences. At the same time, it allows for a good control over the consistency of the sequences of the technical units of one creator, especially if it has handed over all records within one transferring act.

In the area of management of archival records data, the Guide through Fonds and Collections is the basic utility for use. In it, for each fond, data on the quantity and extent of records are entered (Cvelfar, 2011). If these are expressed in technical units, then the guide through fonds and collections represents the source of the cardinality of this sequence. Takeover lists, archival lists, inventories can be defined in this case as the reference sources of cardinality of individual sequences.

Archival professionals often encounter problems with the management of archival records data, and in doing so implement both cardinal methods and the methods of ensuring the consistency of linear and hierarchical sequences. In this context, some problems have been detected in the archival practice already in the technical unit management, the following should be pointed out:

- the methods discussed are used only in those data management segments where the linear sequences are products of natural conditions and arise from the legality of the creation of physical archives in themselves;

- The methods discussed in existing archival practice can not always be consistently implemented because the sequence of technical units is not properly defined, or the technical units are not properly sorted in sequence;[12]
- Problems in the implementation of all three methods are also poorly performed rearrangement (injective mappings) or addition of values within sequences (bijective mappings), implementation of alpha and alphanumeric labels, in particular in the way that it obtains the legality of a nonlinear sequence[13] or pseudo-hierarchical-linear sequences[14].

Implementation of the methods discussed is often not carried out sufficiently precisely due to limited practical needs in the procedures of archival professional processing, especially the physical manifestations of the elements. However, it is to be expected that these methods will become very topical in order to ensure and control large masses of logically and physically rounded altogether at the level of data (Novak, 2018a) and metadata (Novak, 2018b) in archives. These requirements, however, do not arise only from theoretical considerations, but above all from the fact that many data structures are already accumulated in the archives and by archival professionals only partially can manage to handle them (Novak, 2018c). The process of accumulation of data structures is in a trend of strong increase, which is consistent with the production of archival records in a modern information society. The problem is even greater because, according to the dynamics of general development, we can expect an increase in archiva records, both in the physical and the exponential growth of the data captured in electronic form. However, we should add to this also the direction of development in the segment of describing archival records (Popovica, 2016). Very simplified archival information aids were first upgraded to standardized archival aids, and these are developed into contextual archival information aids. Thus those methods, tools and solutions, which will help archival professionals to control individual entities in the contexts of the management of linear and hierarchical sequences, are becoming necessary requirements. Especially since they will allow the overview of the integrity of preserved archives, regardless of whether these methods are implemented at the level of data from archival records or at the level of content or technical metadata.

7 CONCLUSION

Increasing amounts of data that are collected, organized, processed and are ready for use in a modern society are directly reflected also in archival theory and practice. Before some decades simplified data management methods were sufficient for everyday needs in the archives. Today it is different. Archivists cannot avoid different implementing of complex data management methods in everyday work. We can see, that in theory and practice, the process of chaining of methods or their results occurres. Among important basic archival professional methods we count the methods of ensuring consistency of linear and hierarchical sequences, and the method of cardinality of sequences.

12 This is the case if the last archive box of a linear sequence is not physically set to the last place in that sequence.

13 Typical deviation of technical unit management when one or more books as technical units appear in the linear sequence (n = n + 1) of standard archival boxes.

14 An example of this is the introduction of technical units with a poorly implemented labeling of technical units, usually in alphanumeric form.

The method of consistency of linear sequences in archives is that which determines the accuracy of the elements of the sets of the linear sequence in accordance with the known parameters of these sequences. In the case of dealing with hierarchical data structures, it is necessary to implement a method of consistency of hierarchical sequences. Although both methods have some common characteristics, they have also important differences, since their purposes, and above all the expected results are very different. As some kind of upgrade of the use of these methods represents the method of cardinality of elements of sequences in archives.

The method of ensuring the consistency of linear sequences and the associated determination of cardinality of elements of linear sequence is the starting point for the data management and for the data structures in the archives. We implement it in many basic archival professional processes and subprocesses in solving archival professional problems. In practice, the method itself can be supported in various ways, as with easy-to-use system tools that enable at least sorting elements of individual sequences.

Another important method for today's archivists is the method of ensuring the consistency of hierarchical sequences, including the determination of the power set of elements of the hierarchical sequence. It presents methodically well-formed data structures with a consistent data management tool and their structures. The method can be implemented in many archival professional processes, as in phases of testing the planned solutions of individual archival professional problems. It can be also systemically supported or not.

The method of cardinality of sets of sequences in the archives represents the possible management of those data structures that need to determine consistency over time and space, or in cases where the consistency of one sequence of elements is determined by two or more data structures over time and space.

All three methods can be successfully implemented in archival theory and practice on those sequences of data sets that have standardized and long-term sustainable data structures.

References

Cvelfar, B. (2011). Vodnik po fondih in zbirkah arhiva v knjižni obliki - potreba ali nuja? V: Tehnični in vsebinski problemi klasičnega in elektronskega arhiviranja : zbornik referatov dopolnilnega izobraževanja s področij arhivistike, dokumentalistike in informatike..., No. 10 (p. 387-410). Maribor : Pokrajinski arhiv Maribor. Retrieved 16. 2. 2019: http://www.pokarh-mb.si/uploaded/datoteke/Radenci/radenci2011/35_cvelfar_2011.pdf.

FRAN – Slovarji Inštituta za slovenski jezik Frana Ramovša ZRC SAZU. Retrieved 14. 2. 2019: https://fran.si/.

Hajtnik, T., Škoro Babić, A. (2018). Ali nam lahko pri vrednotenju in odbiranju elektronskega gradiva pomaga tehnologija? V: Moderna arhivistika : časopis arhivske teorije in prakse, 1, No. 1. (p. 169-196). Maribor : Pokrajinski arhiv Maribor. Retrieved 14. 2. 2019: http://www.pokarh-mb.si/uploaded/datoteke/radenci_20181/1_2018_169-196_%C5%A0koro.pdf.

ISO/TR 13028. Information and documentation - Implementation guidelines for digitization of records. Retrieved 14. 2. 2019: https://www.iso.org/standard/52391.html.

Novak, M. (2011). Metode stručnog i istraživačkog rada u savremenoj arhivskoj teoriji i praksi. V: Arhivska praksa, 14 (p. 213-227). Tuzla : Arhiv tuzelskega kantona.

Novak, M. (2012). Signatura kot entiteta arhivskih informacijskih sistemov. V: Tehnični in vsebinski problemi klasičnega in elektronskega arhiviranja : arhivistika in informatika, (p. 605-630). Maribor : Pokrajinski arhiv Maribor. Retrieved 15. 2. 2019: http://www.pokarh-mb.si/uploaded/datoteke/Radenci/Radenci2012/62_Novak_2012.pdf.

Novak, M. (2018a). O relacijah med fizičnim, digitalnim in digitaliziranim arhivskim gradivom = About the relations between physical, digital and digitized archival records. V: Arhivi v službi človeka - človek v službi arhivov : simpozij : zbornik znanstvenih prispevkov z recenzijo (p. 81-90). Maribor : Alma Mater Europea – ECM.

Novak, M. (2018b). Fasetni model oblikovanja naslovov popisnih enot. V: Moderna arhivistika : časopis arhivske teorije in prakse, 1, No. 2. (p. 385-406). Maribor : Pokrajinski arhiv Maribor. Retrieved 10. 2. 2019: http://www.pokarh-mb.si/uploaded/datoteke/2_2018_385-406_novak.pdf.

Novak, M. (2018c). Upravljanje občutljivih podatkov in arhivska informativna pomagala. V: Atlanti, letnik 28, št. 2 (p. 33-41). Maribor : Arhive Centre for Professional and Technical Problems.

Popovici, B. F. (2016). Records in contexts : towards a new level in archival description? V: Tehnični in vsebinski problemi klasičnega in elektronskega arhiviranja. Popisovanje arhivskega gradiva: zbornik mednarodne konference, Radenci, April 13-15 2016 (p. 13-31). Maribor : Pokrajinski arhiv Maribor. Retrieved 10. 2. 2019: http://www.pokarh-mb.si/uploaded/datoteke/Radenci/radenci_2016/013-031_popovici_2016.pdf.

Šauperl, A. (2007). Malo znane lastnosti klasifikacije – Kratek vpogled v teorijo klasifikacije in njeno uporabo v knjižničnih katalogih. V: Tehnični in vsebinski problemi klasičnega in elektronskega arhiviranja: Zbornik referatov z dopolnilnega izobraževanja (p. 83-95). Maribor : Pokrajinski arhiv Maribor. Retrieved 15. 2 2019: http://193.2.142.10/fileadmin/www.pokarh-mb.si/pdf_datoteke/Radenci2007/_auperl.pdf.

Semlič Rajh, Z. (2012). Arhivski zapisi in postopki sledenja v arhivskem informacijskem sistemu. V: Tehnični in vsebinski problemi klasičnega in elektronskega arhiviranja : arhivistika in informatika : zbornik mednarodne conference (p. 631-646). Maribor : Pokrajinski arhiv Maribor. Retrieved 10. 2 2019: http://www.pokarh-mb.si/uploaded/datoteke/Radenci/Radenci2012/63_Semlic_2012.pdf.

Semlič Rajh, Z., Šauperl. A. (2011). Problematika gesljenje in oblikovanja deskriptorjev. V: Tehnični in vsebinski problemi klasičnega in elektronskega arhiviranja : arhivistika in informatika : zbornik mednarodne conference (p. 187-213). Maribor : Pokrajnski arhiv Maribor. Retrieved 10. 2 2019: http://www.pokarh-mb.si/uploaded/datoteke/Radenci/radenci2011/19_semlic_sauperl_2011.pdf.

Semlič Rajh, Z.; Šabotić, I., Šauperl, A. (2013). Znanstvenoraziskovalno delo v arhivistiki : značilnosti uporabe dveh raziskovalnih metod. Tehnični in vsebinski problemi klasičnega in elektronskega arhiviranja : arhivi in ustvarjalci gradiva : stanje in perspektive, (p. 125-144). Retrieved 10. 2. 2019: http://www.pokarh-mb.si/uploaded/datoteke/Radenci/Radenci2013/11_Semlic_Sabotic_Sauperl_2013.pdf .

Uvdag. (2017). Uredba o varstvu dokumentarnega in arhivskega gradiva. Uradni list Republike Slovenije, No. 42/2017. Retrieved 15. 2 2019: https://www.uradni-list.si/glasilo-uradni-list-rs/vsebina/2017-01-2140?sop=2017-01-2140.

Weisstein, Eric W. Cardinal Number. V: MathWorld--A Wolfram Web Resource. Retrieved 15. 2 2019: http://mathworld.wolfram.com/CardinalNumber.html.

ZVDAGA-A. (2014). Zakon o spremembah in dopolnitvah Zakona o varstvu dokumentarnega in arhivskega gradiva ter arhivih. Uradni list RS, št. 30/06, 51/14. Retrieved 15. 2. 2019: https://www.uradni-list.si/glasilo-uradni-list-rs/vsebina?urlurid=20142170.