**VISUALIZATION OF INFORMATION ANALYSIS AS A TOOL FOR CLUSTER IDENTIFICATION OF GLOBAL PRODUCTION NETWORKS**

*Luis ORTEGA-SEGURA[[1]](#footnote-1)*

*Mariel GAYTÁN-RODRÍGUEZ*

**Abstract**

The so-called ICIO (Inter Country Input Output) models, properly defined, allow for developing more accurate indicators on the interdependency in the production structure of different countries. Until now, this framework has been used to analyze groups of selected countries and periods of few years. One of the reasons is the large volume of information needed and the difficulty to handle it.

This paper aims to visualize the global production process using the WIOD (World Input-Output Data Base).This database provides disaggregated input-output matrices for 40 countries and 35 economic sectors, equivalent to more than 2 million data for each year.

The authors evaluated a visual information analysis technique associated with the network theory which is able to map and cluster large data without excluding any information. This is a new tool that is alternative and/or complementary to Multidimensional Scaling Models or Graph Layouts and leads to identify the main Global Production Networks and, most importantly, the Global Value Chains, thereby opening a new path for relevant analytical procedure.

Thus, our goal is to make visible by means of computational techniques of information display, the global production chains in terms of their performance producing domestic and external value to an economy, based on the network of relationships that are captured by a model ICIO and the enormous volume of information provided by WIOD.

Key words: Information Analysis Visualization; ICIO Models; Global Production Networks; Global Value Chains.

1. **ICIO models: fundamental to study international production fragmentation**

Since more than two decades, and ever increasingly, most international goods traded are primarily intermediate products, i.e., imported and exported inputs related to productive processes that take place in several countries before satisfying their final demand as finished products.

Hence, the measurement of trade in gross terms is gradually being complemented by the measurement of trade in value-added terms, where the domestic value added represents the country’s exports that contribute to total Gross Domestic Product (GDP), while the foreign value added represents the country’s gross exports made up by intermediate inputs imported from other countries.

The aforementioned are more precise and detailed records that require supply and use tables from the National Accounts, which are not available in all countries and, for that reason, they have been estimated by the OECD and other international organizations like the UNCTAD through the UNCTAD-Eora Data Set and the WIOD (World Input-Output Data Base).

In parallel, and in order to improve and update the aforementioned existing information, joint efforts have been made to carry out a series of studies using inter-country models based on the interconnection of their respective input-output matrices. In the following pages we will describe a model developed by Koopman et al. (2010) which will allow us to go from breaking down value added trade to identifying global productive chains.

An Inter Country Input Output model (ICIO model)

Assumptions:

1. There are two countries that produce tradable goods in N sectors of their economies.
2. Such goods may be part of the final demand (directly consumed products) in each country or are intermediate inputs from a further production in the other country.
3. Each country exports both final consumer goods and intermediate goods to the other.

Based on the above it may be inferred that each country’s Production Gross Value (Xr) can be used as an intermediate or final consumer good in both the domestic economy and the global economy.

On the basis of the general input-output model X = Ax + Y, the above mentioned can be expressed as:

where r,s = 1,2 ……………… (1)

For a bilateral production and trade system, we would have:

………………….………….(2)

………………….………….(3)

If the system is expressed in matrices, we would have:

…………………….………..(4)

where:

X1,2 represents the production gross value vector (2x1) of each country.

Ar,s represents the (total) direct and indirect requirement matrix (2x2) needed to satisfy an increase in demand in any of the two countries given their respective productive structures. Therefore:

A11 is the technical coefficient matrix (2x2) for (final) goods produced and consumed in country 1.

A12 is the technical coefficient matrix (2x2) for (intermediate) goods produced in country 1 and used in country 2.

A22 is the technical coefficient matrix (2x2) for (final) goods produced in country and consumed in country 2.

A21 is the technical coefficient matrix (2x2) for (intermediate) goods produced in country 2 and used as inputs in country 1.

Yr,s is the final demand vector (2x1) of each country, both for final and intermediate goods:

Y11 the final demand vector (2x1) for (final) goods produced and consumed in country 1.

Y12 is the final demand vector (2x1) for (intermediate) goods produced in country 1 and used as inputs in country 2.

Y22 is the final demand vector (2x1) for (final) goods produced and consumed in country 2.

Y21 is the final demand vector (2x1) for (intermediate) goods produced in country 2 and used as inputs in country 1.

As is well known, a key question in the input-output analysis is: in the event of a demand shock originated by external factors, how much of each sector production gross value should be modified to satisfy the changes in final demand? In this bilateral model, the final demand vector Yr,s of goods consumed as both final goods in each country and intermediate goods used in the other country (on the basis of assumption (2) of the ICIO model) needs to be determined. We must remember that, in the general model (I-A) X = Y, i.e., for each final demand vector there is a production value vector with n linear equations of unknown variables; thus, the solution of the unknown variables would be given by X= (I-A)-1 Y. Using equation (4), we would have:

…………………………..…………(5)

Since Y11 + Y12 is the total demand for goods produced in country 1 and Y21 + Y22 is the total demand for a good produced in country 2, equation (5) may be rewritten as:

…………………………………………………..…….(6)

Where matrix B is the Leontief inverse matrix including a country’s production value amounts needed for a one-unit increase in the final demand of the other country. It is a block matrix since it contains final and intermediate good requirements needed to satisfy demand in any of the two countries.

We then measure the content of domestic and foreign production inputs through value added. Using the same process used for determining the technical coefficients of matrix A (which shows, through coefficients, each country’s inter-industrial structure), and normalizing the production gross value (X), we obtain vector Vr,s which is the direct (domestic) value added coefficient vector Vr,s/Xr,s.

This vector shows the domestic value added proportion in domestic production and equals 1 minus its reciprocal number. The domestic value added reciprocal number comprises all the other intermediate inputs (which include both imports and domestically produced intermediate inputs). It can be expressed as follows:

………………………………………………………….. (7)

To show this more clearly in our bilateral model, by substituting *r* and *s*, we would have:

………………………………………………………….. (7a)

………………………………………………………….. (7b)

Where *u* is a unit vector (1xN) and the domestic value added for each country equals 1 minus the sum of all total intermediate goods. It should be noted that A12 is the technical coefficient matrix (2x2) for (intermediate) goods produced in country 1 and is used as inputs in country 2, and that A21 is the technical coefficient matrix (2x2) for (intermediate) goods produced in country 2 and used as inputs in country 1. If this vector is diagonalized, we would have:

………………………………………………………………….(8)

What do we have at this point? On the one hand, we have matrix B, which reveals which are the production requirements for both final and intermediate goods that are needed to satisfy an increase in demand in any of both countries. On the second hand, we have matrix V, which shows the proportion of domestic value added in each country’s gross production value as measured through its reciprocal number.

By combining these two matrices, which are derived from formulas (6) and (8), a measure of the value added proportions according to its source (VAS) is obtained. Hence:

………………………(9)

Each of the elements in this matrix represents the following:

V1B11 is the value added proportion of the goods domestically produced in country 1 by each sector.

V2B21 is the value added proportion of country 2 of the goods produced by the same sectors in country 1.

The sum of these two elements is equal to the total value added needed to produce an additional unit of product in country 1 and, since they are coefficients, they are equal to one. Consequently,

V1B12 is the value added proportion of country 1 of the goods produced in country 2 by each sector.

V2B22 is the value added proportion of the goods domestically produced in country 2 by sector.

“The VAS matrix contains all the information needed to separate the national and international content of production and trade in each country, at sector level” [Koopman, et al.].

The importance of constructing the VAS matrix in the context of the ICIO model is worth underlining. On the one hand, it allows for building vertical specialization indicators of countries, i.e., it allows measuring the imported content in exports made by each country. On the second hand, by using a VAS matrix it is possible to determine all the productive chains existing among countries. Last but not least, the matrix can be used to identify value added global chains (VAGC). These two last concepts are not synonyms: the first allows mapping who produces what, while the second allows identifying who really provides value added to world production. Let us begin with the vertical specialization.

As stated above, one of the first interrelated indicators of world production is the so-called vertical specialization (VS, for its acronym in English), which is nothing else than measuring the use of imported inputs to produce goods that will be exported later. Based on the seminal paper by Hummels et al. (2001), efforts have been made to measure this indicator more precisely in the domestic (see Chen H. et al., 2005) and international context (see Amador, J. and Cabral, S. 2008).

On the basis of the VAS matrix, which indicates the proportions of production value added by source, total exports (measured as per traditional statistics) can be used as weights to calculate the proportion of intermediate and final good exports. The assumption (common to all input-output literature) that the proportion of both domestic and foreign value added of intermediate goods is the same as that of final goods in each sector of different countries is implied.

The next step is therefore to combine the VAS matrix with the gross export vector E of each country. Like in the case of value added, matrix Ê, which is a diagonal matrix, was defined for the exports of both countries:

……………………………………………………. (10)

The combination of the VAS matrix with the export matrix will certainly generate a matrix that allows separating the value added content of each country’s gross exports (by sector), both domestic and foreign. Hence:

…………………………………………..…….. (11)

Unlike other vertical specialization measurements, this indicator includes all direct and indirect value added contributions included in the exports of a specific sector. The elements in the main diagonal define the domestic value added of exports in each country, while the elements outside the diagonal represent the foreign (imported) value added integrated in each country’s exports.

Only as an example we present the identification of global value chains according to the information provided by the WIOD for 1995 and 2011 through VOSviewer. As mentioned in the introduction, these are graphic interfaces used to locate the underlying structures within a large volume of information. The list of countries and activities is found in Appendix1.

1. **The emergence of big data in input-output analysis: the WIOD information.**

As stated by Auray et al. (2000), we live in a time that can be defined as the information era. This is particularly true in the field of input-output. For a long time, the input-output matrices were rarely or not calculated at all for various reasons, among others, the paramount cost for countries of building the statistical device needed to derive such matrices.

However, in this field as in many others, we are currently facing overflowing (and at times, excess) information. “We benefit today […] from a considerable improvement of information transmission, storage, and exploitation. Hence, thanks to the recent development of informatics and its organization, the amount of available information in the field of human sciences is so significant, so diversified, and so complex that, if we are not careful, this information will serve a more distressing than modeling purpose. Excess of information may become a lack of information and this excess may be even more dangerous when the abundance of computer resources may generate the creation of more and more information!” (Auray et al., 2000).

The conclusion that knowing how to use information is a top priority as well as devising a key tool that helps to guide us in the ocean of complex and bulky data is inevitable.

In the next pages we will briefly describe the content of WIOD. The reader interested in delving into this information source should refer to the text by Timmer et al. (2012).

The first thing to highlight is that this database is based on the construction of the World Input Output Table (WIOT) which, in turn, uses the Supply and Use Tables (SUTs) of countries. In other words, the construction of WIOT is based on domestic SUTs, rather than on input-output matrices (which many countries do not generate). This fact allowed us to use the final production and consumption series from the national accounts of each country.

The advantage of using SUTs is that they provide information on products and industries, as compared to an input-output matrix which exclusively provides information product by product or industry by industry. SUTs do not require the number of industries to equal the number of products nor an industry to produce only one product.

Besides, the construction of SUTs is based on two accounting principles: 1) Total use must equal total product supply, and 2) for each industry, the total value of inputs must equal the total value of production. In formal notation:

3)

4) j

The first equation shows that the supply of each product is given by the sum of domestic supply plus imports, while the second shows that the total consumption (use) is given by the sum of final domestic consumption (F), exports, (E) and intermediate consumption (I). The third equation shows that, for each industry, the total value of production equals the total value of inputs, a value given by the sum of value added (VA) plus intermediate use (I).

By combining a SUT with foreign trade information we obtain an international SUT. This is mainly composed of domestically manufactured products and imported products. This breakdown must be done ensuring that total domestic supply equals the use of domestic production of each product and that total imports equals the use of imported products.

Imports are allocated by classifying use as follows: the percentage of use categories (intermediate, final consumption, and investment) was used for dividing total imports in the SUT. This same classification is made for each product to finally generate an import matrix. Once the import matrix is obtained, each cell is divided between country of origin (since countries may differ in terms of product classification, but the division is not made between categories).

An Input-Output Table (IOT) is built over a SUT at base prices, taking into account additional changes in technology. The assumption of a “fixed product sales structure” is used, according to which each product has its own sales structure, regardless of the industry producing it.

Summing up, each country generates a SUT and, if those of several countries are added, we obtain a world SUT which becomes the WIOT using the assumption of “fixed product sales structure”. Hence, bilateral export and import flows mirror themselves: imports by country A from country B are exports by country B to A.

The WIOD includes 27 European countries and 13 countries from other continents. Countries are selected on the basis of the quality and public availability of their data and its economic relevance as well. All SUTs are classified into 35 industries of 59 products and all IOTs refer to 35 industries. In total, there are 40 countries and the rest of the world.

To conclude with this section, two elements deserve mention:

1. The information in WIOD, although based on real data, includes assumptions and estimates based on these assumptions; hence, the results must be seen only as indicators. Undoubtedly, it is a major source of information in an on-going improvement process (with the information provided by countries); nevertheless, it should be used with due caution.
2. All the information must be understood as a network and is therefore susceptible to be analyzed according to the instruments proposed by the network theory (including visualization).
3. **Visualization of information analysis: a new tool for solving new problems**

Information visualization may be defined as "the science of analytical reasoning facilitated by interactive visual interfaces" (Chung Wong, Park, 2007). At least three elements may be perceived in this definition: a) analytical reasoning; b) visual interfaces; and, c) interaction with the researcher. However, analytical reasoning may only be applied once the inventory and characteristics of study population have been determined. From here, separating the essential from the secondary, identifying underlying structures to explain the phenomena observed through visual interfaces, and conducting research at different levels, is a continuous process.

A network like the one we are studying may be mathematically approximated in two ways, either as a double entry table (as the matrices mentioned in previous sections) or as a graph, a system of logically structured signs. However, as stated by Bertin, J. (1981):

“A graph is not merely a drawing, but, at times, a heavy responsibility in decision making. A graph is not “drawn” once and for all, but “built” and rebuilt until it reveals all the relationships between data interaction. The best graphic operations are those carried out by the researcher himself. […] Graphs can be used at least in two different manners that should not be confused: first, as a means for communicating information (in which case the communicator already understands this information), and second, for graph processing (in which case, a person manipulates and perceives graphic objects for problem solving)”.

Indeed, evaluating the configuration and dynamics of population or of the study phenomena and approaching a determination of causes and effects, are part of the daily scientific practice. Thus, information visualization must be considered as an inevitable method of social sciences.

In a world where computer science is constantly evolving and being updated, where the quantity of available information might be overwhelming, the ability to interpret data from various sources and translate them into a multimedia, interactive, and real-time interface is a valuable tool for decision making.

“Generally, information visualization (VI) has been recognized as a way of presenting the results from a process in a graphical, concise, and interesting manner. However, the study and application of VI goes far beyond this concept, i.e., in the last few years, development has focused on creating a series of state-of-the-art techniques regarding information retrieval and organization, using visualization as an exit means. This field studies how to build bi- or tri-dimensional images capable of producing a set of visual stimuli for the observer [researcher] to give him/her enough elements to form an opinion on a certain complex field or problem, alleviating him from the paramount efforts of reviewing enormous data volume of all types and favoring the detection of underlying patterns or structures in the information reviewed” (see Nivón, A., 2011).

There are several VI techniques, among which the Multidimensional Scaling (MDS) and Graphs Layouts can be mentioned. These are applied to the analysis of social networks. The MDS is an analysis method of one distance or similarity (dissimilarity) matrix whose objective is to model distances or proximities between nodes to represent a low dimension (generally bi- or tri-dimensional) space as precisely as possible. The main focus is that a clear and effective visualization of data may speed up understanding and timely response to various scenarios.

However, when data do not have a low dimensionality and do not follow the symmetry and inequality inherent to metric spaces, graph layout algorithms are more adequate.

“One alternative and complement of MDS are Graph Layouts (GL). These are based on the considering that metric spaces have the advantage of being a very close representation of the surrounding physical space (distances, heights, etc.) and their application to drawing and visual creation is effective in transmitting information on certain phenomena” (Solís, V., 2012).

This author describes different GL algorithms, from Eades’ “Spring Embedder” to the Fruchterman and Reigold algorithm, and down to the Kamada-Kawai algorithm. All these models comply with the five basic principles of Graph Layouts, namely: 1) All nodes are visible at the same time; 2) all available space is used; 3) line lengths are uniform; 4) intersection between lines is minimal, and 5) vertices are separated at a reasonable distance.

Now, when faced with networks with millions of nodes that are related through irreducible asymmetries, we need techniques that allow for visualizing more favorably the large volume of information and interact and carry out a multi-level analysis. Through interactive visual interfaces the analyst may thus concentrate on the most relevant aspect of the massive amount of data available.

In section 4 we will describe the VOSviewer computer program which we consider a most useful tool for visualizing information due to its ability to handle large data volumes and provide a more satisfactory representation of the data than the two previously described approaches.

1. **Visualization of Similarities (VOS): global approach without missing information**

As mentioned above, graphs usually consist of a network of points (nodes or vertices) and a set of lines. Nodes and lines become integrated or join based on a relationship. The characteristics of nodes and lines, such as size, color, thickness, etc., are used to represent additional information that is used to improve the visualization of a network; points may have many different shapes, sizes, and colors and may have text labeling or even different letter fonts and hence they may represent several attributes of nodes. Relationships may also be represented with lines of different thicknesses, sizes, and colors.

VOSViewer has all of these features. As stated in the Manual Introduction (see Van Eck, N.J. and L. Waltman, 2013), VOSviewer is a computer program that can be used for the following purposes:

* To create maps based on network data. Maps are created using a special mapping and clustering technique.
* To visualize and explore maps or graphs. This exploration may be done in different ways to emphasize various aspects of the graph.
* The program offers functions such as zoom, displacement, and search, hence facilitating a detailed examination of the graph. This program allows analyzing a considerable amount of data and provides answers in real time.
* The program can also be used to create maps based on any other type of networks because it uses cartography techniques to build maps. It can show graphs that are done using any other adequate mapping techniques, such as MDS.
* Finally, the program can run in several hardware platforms and operative systems and may be directly started from Internet.

The purpose of VOS is to provide a visualization in few dimensions in which the nodes are located in such a way that the distance between any pair of them reflects their similarity as precisely as possible. Nodes considerably similar will be closer, while those with low similarity will be far from each other.

VOS minimizes the weighted sum of squared Euclidean distances between all node pairs: the greater the similarity between two nodes, the greater the weight of their square distance will be in the sum. To avoid solutions where all nodes are within the same coordinates, a restriction is established so that the sum of all distances equals a positive constant. Let us briefly go over the mathematics underlying this technique.

We have *n* nodes, from 1..n, a similarity matrix of *n* x *n* S=(Sij) Sij>0, Sij=0, Sij=Sij for all i,j ∈ ⎨1,….,n⎬. We assume that similarities in S may be considered as measurements in a proportion scale. Matrix X of *n* × *m*, where *m* is the number of space dimensions used, includes the coordinates of objects 1, ..., *n*. Vector xi = (xi1, ..., xim) ∈ Rm is the i line of X and includes the coordinates of object i.

In mathematical notation, the objective function to be minimized in VOS is given by E (X, S) = (1)

…………………………………………………(1)

Where is the Euclidean norm

Minimization in (1) is subject to

………………………………………………………… (2)

are not squared in the restriction; this restriction is more satisfactory for visual analysis.

Under certain conditions, VOS is equivalent to Sammon cartography, a special variation of Multidimensional Scaling (MDS). The mathematical notation will be the same we are using. D = ( dij ) is used to express a dissimilarity matrix *n* × *n* satisfying dij > 0 y dij = dji for all i , j ∈ { 1 , ..., n }. The element dij in matrix D indicates the dissimilarity between objects i and j.

Like Multidimensional Scaling, Sammon mapping is aimed to offer a space with few dimensions in which nodes 1 , . . . , n are placed in such a way that the distance between any pair of nodes i and j reflects their dissimilarity dij as precisely as possible. Nodes with a high dissimilarity are located far from each other, while those with a low dissimilarity are close to each other.

Sammon mapping and VOS share a very similar purpose, with the difference that Sammon mapping uses dissimilarities, while VOS uses similarities.

VOS and EMD may provide quite different solutions. In applications where indirect similarities along third nodes may include relevant information, VOS probably offers better solutions than EMD. An example of an application where VOS may be more appropriate then MDS is the visualization of associations between concepts based on co-occurrence data. Typically, several pairs of concepts may not co-occur at all, and these following pairs of concepts have a similarity of 0. EMD aims to provide a visualization in which, for each pair of concepts with 0 similarity, the distance between concepts is the same.

Conversely, VOS focuses more on indirect similarities along third nodes and hence may locate concepts with closer indirect similarity than concepts with low indirect similarity. Based on this property we believe that VOS may provide more interesting visualizations of concept associations than EMD.

In section 5, the information included in the VAS\_Ê matrix (see equation 11 in section 1) will be shown for years 1995 and 2011.These matrices include both the direct and indirect value added contributions in exports of each of the 35 sectors.

In the main diagonal of these matrices there is the domestic value added of each of the 35 sectors in 40 countries, while the foreign value added included in exports of each sector in these countries is located outside the diagonal. This is how the chains of productive sectors with higher value added, i.e. the Value Added Global Chains (VAGC), is identified.

At this point, we expect the reader is able to clearly differentiate vertical specialization (the imported content of countries’ exports), global production chains, and value added global chains (VAGC). However, the purpose of this paper is not to discuss each of these three indicators of world production fragmentation, but to visualize effectively a considerable volume of information.

We have broadly described the methodology used to identify VAGC because, as pointed out by Brandes (2005) “… to produce effective visualization, the relevant information must be clearly identified, i.e., we need to filtrate, transform, and process the collection of players, links, and features of productive sectors to identify the core elements [in the present case, productive sector chains that generate more value added within the world economy], define an appropriate map for graphic representation, and generate the corresponding image with no devices included.”

1. **WIOD graphic results (1995 and 2011)**

The use of WIOD entails handling considerable volumes of information (big data) because, as described in section 2, it includes the input-output matrices broken down into 35 sectors from 40 countries. In other words, there are 1435 x 1435 matrices, which imply two million records for each year. Fortunately, today there are programs that help handle considerable volumes of information. Nevertheless, a computer strategy had to be designed implying the use of programs such as Mathematica[[2]](#footnote-2), Pajek, and finally VOSviewer. Although the last two where devised to solve problems not related to economic matters, a relatively novel approach will be to use them for input-output analysis, particularly to address the subject of information visualization.

In the following section, we will show the main functionalities of VOSViewer applied to matrices that facilitate the identification of VAGC. By interacting with the first graph results, four value-added global chains for each of the years will be set out in a second level of analysis.

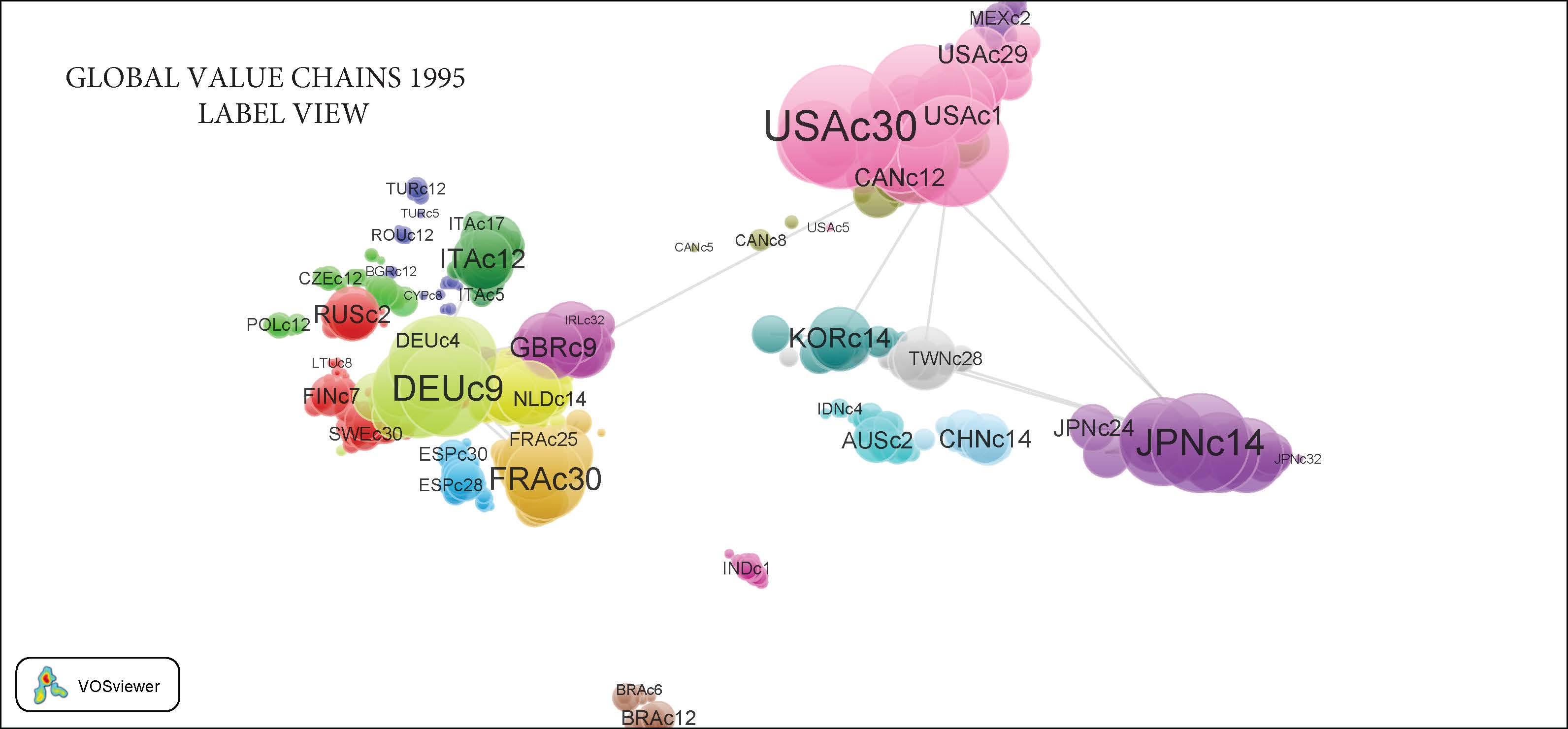
Once the graph of the VAS\_Ê matrix (which includes 1 million and 960 thousand nodes) is created, we will obtain four different views of this graph in the *Main Panel*:

In the *Label View*, each node is shown by name and by a circle (default). The size of both the name and the circle depend on the weight of each node; the weight, in turn, depends on the strength or intensity of the links established between this node and the remaining. In this case, the visualization of the graph will allow us to see the main economic sectors that provide value added to world production, as well as the countries where the sectors are located.

As can be clearly seen, the main sectors generating value added within the 40 countries studied for the year 1995 (according to WIOD) were the following: Renting of machinery and equipment (M&Eq) and other business activities (c30); Electrical and optical equipment (c14); Chemicals and chemical products (c9), and Basic metals and fabricated metal (c12).

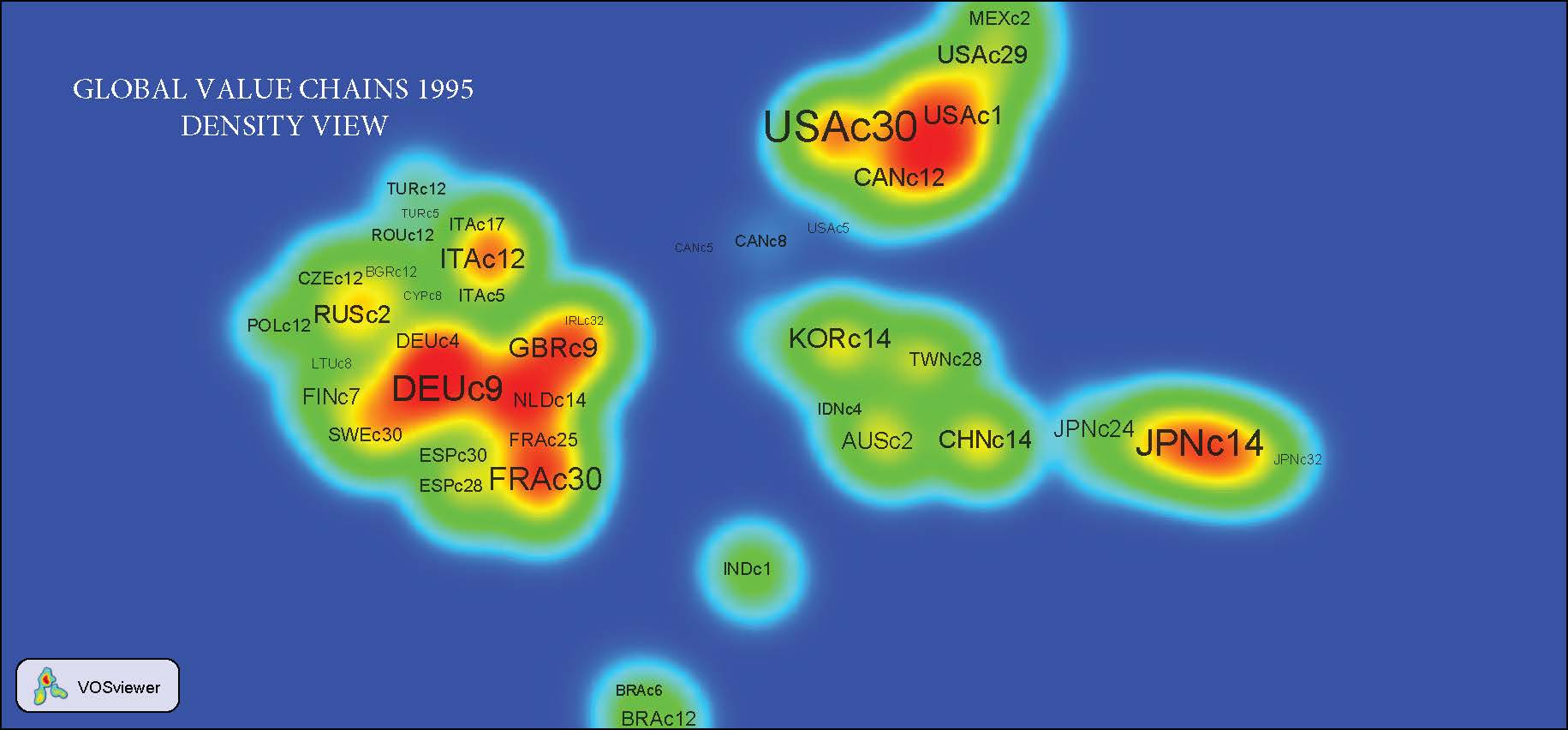
The main countries heading these sectors were the US (USA), Japan (JPN), Germany (DEU), and Italy (ITA), respectively

**Graph 5.1 Label View 1995**



An alternative visualization may be achieved through Density View, where each node color depends on the intensity in each point. By default, the color pallet ranges between red and blue. The greater the number of nodes converging around one node, the closer to bright red the color is.

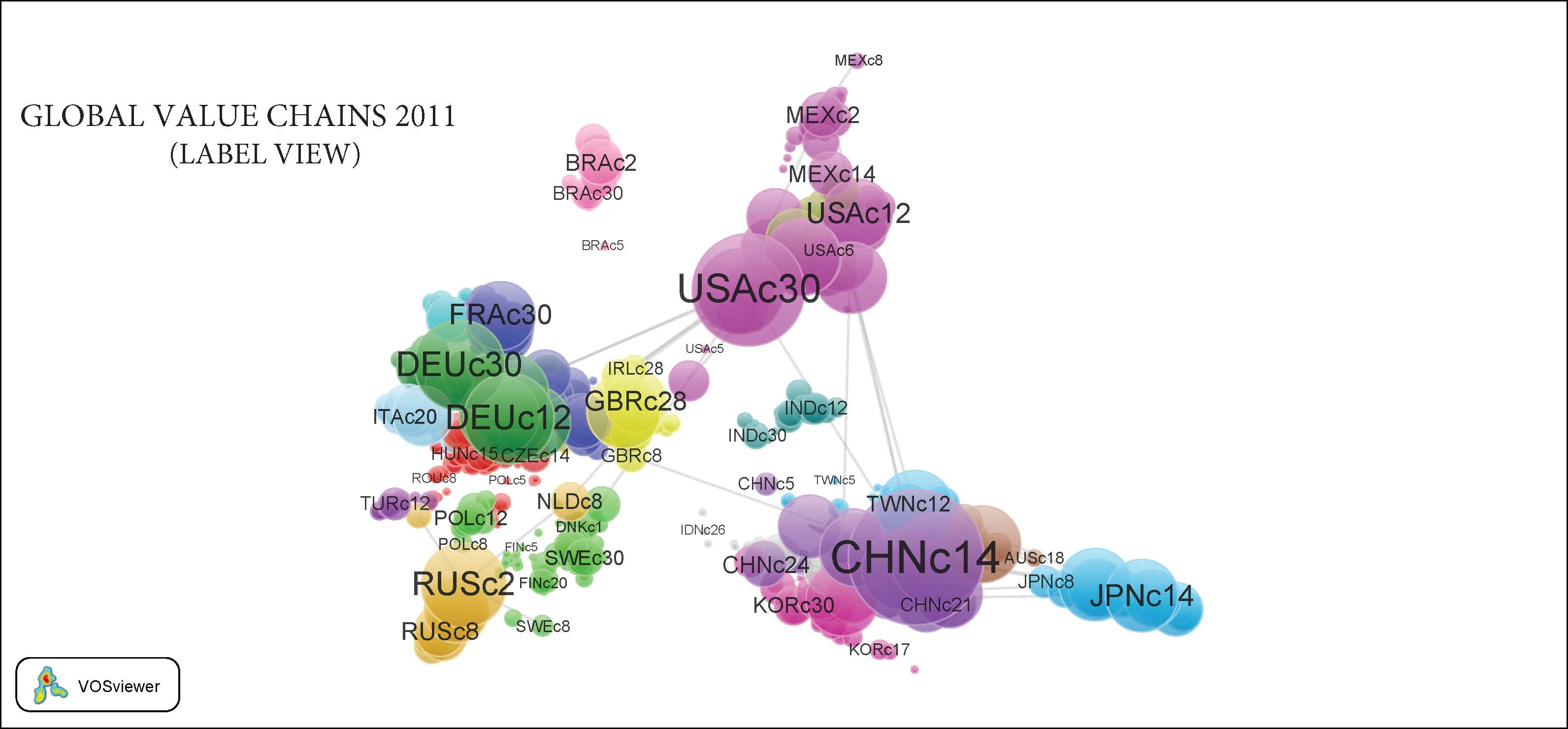
**Graph 5.2 Density View 1995**



The “triad” formed by the Asian economic blocks (headed by Japan at that time), Europe (headed by France and Germany), and North America is clearly outlined. The countries and most important economic sectors in terms of value added are included in this triad.

One of the researcher’s main reactions will be to ask himself/herself whether the sectors generating value added in the 1995 world economy were the same for 2011. We generated the same graphs for this last year and obtained the following images:

**Graph 5.3 Label View 2011**



A simple visual comparison allows detecting a change in relative importance in both the sectors generating value added and the countries implied. It is most evident that the sector Electrical and optical equipment (c14) and China (CHN) now have a much more important weight in the world economy.

On another front, almost two decades later, although the economic blocks still continue, they now include more countries. In other words, the network generating world value added has become more linked and incorporates more sectors and countries. However, the relative importance of both sectors and countries is quite different.

Indeed, at that time Italy practically disappeared from sector c12 (Basic metals and fabricated metal). Although sector c30 (Renting of machinery and equipment and other business activities) upgraded its ranking to second place, it included other countries, while in 1995 the main cluster was made up by Denmark, Estonia, Finland, Lithuania, Latvia, Russia, and Sweden. In 2011, the main cluster comprised Austria, Bulgaria, Cyprus, the Czech Republic, Greece, Hungary, Malta, Romania, Slovakia, and Slovenia. As shown in Table 1, practically all countries in cluster 3 become part of cluster 1 and form the most important group in the sector. France and the Netherlands (Belgium, Holland, and Luxemburg) also rank in 2011 as an important group in terms of value added in the mentioned sector.

It is worth noticing that the main countries (not necessarily linked to any cluster) generating value added in sector c30 show a higher relative importance during this period, namely France and the US.

**Table 1**

**Two Main Global Value Chains and their Mapping/Location**



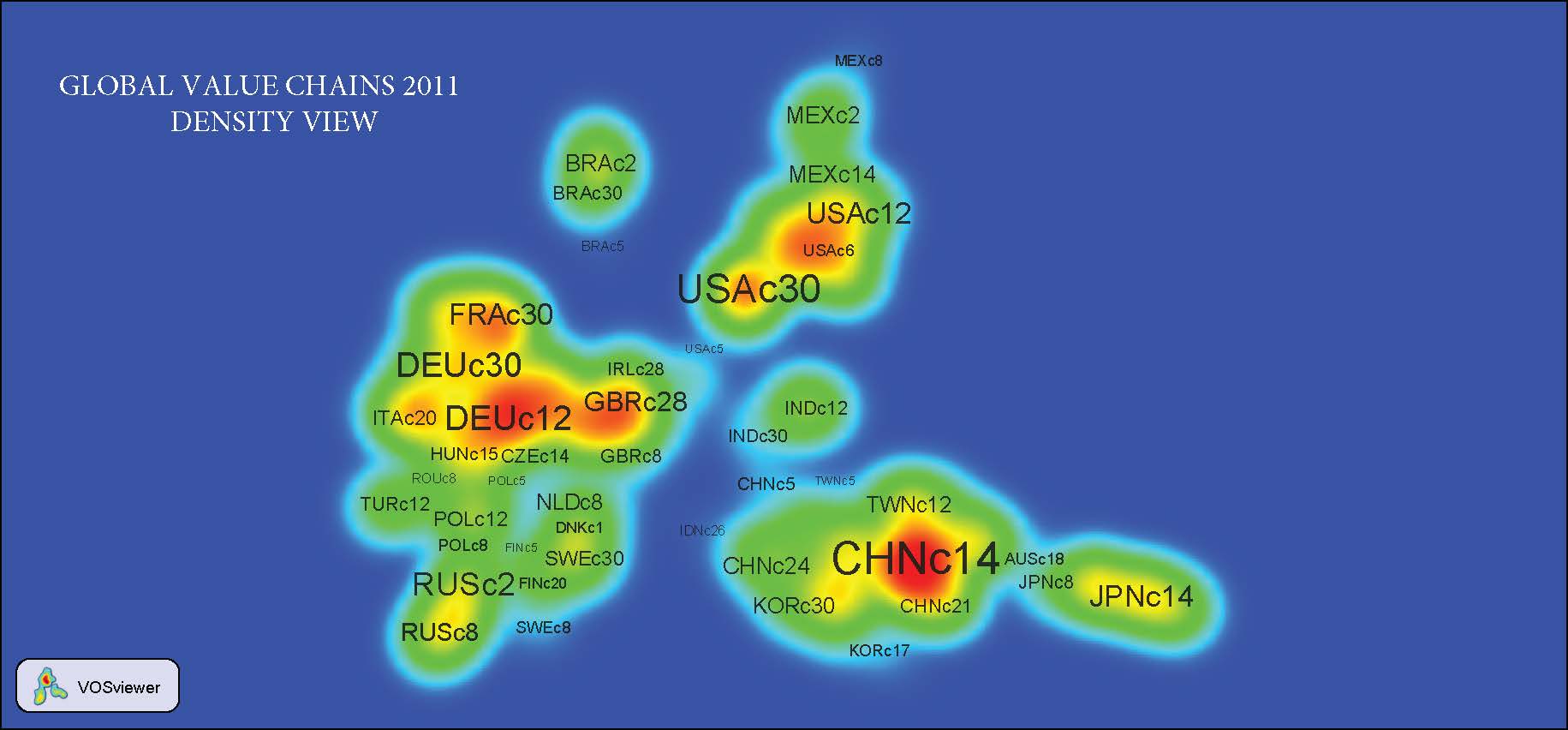
\*Includes the renting of automobiles and other transport equipment (land, water and air), and of machinery and equipment such as agricultural, construction and civil engineering machinery; office and equipment machinery including computers, personal and household goods. Computers and computer-related activities (hardware and software consultancy; data processing, maintenance and repair). Research and development and other business activities (including legal, accounting, book-keeping and auditing activities).

\*\* Includes the manufacturing of office machinery, computers, electrical machinery and devices; and of radio, television and communication equipment. Manufacturing of medical, precision and optical instruments, watches, and clocks.

Source: Authors’ construction using data from WIOD.

In the case of Electrical and optical equipment industry (the second most important chain generating value added worldwide), results are quite similar: the group comprising Austria, Czech Republic, Hungary, Poland, Slovakia, and Slovenia, moves from second place in 1995 to first place in 2011. The same happens with The Netherlands, which upgrade their ranking from fourth to third place, together with France. As for the main countries, hierarchy is different in this sector, which is headed by the US, followed by South Korea, Japan and China. All these countries increased their importance during the period, except for South Korea.

**Graph 5.4 Density View 2011**



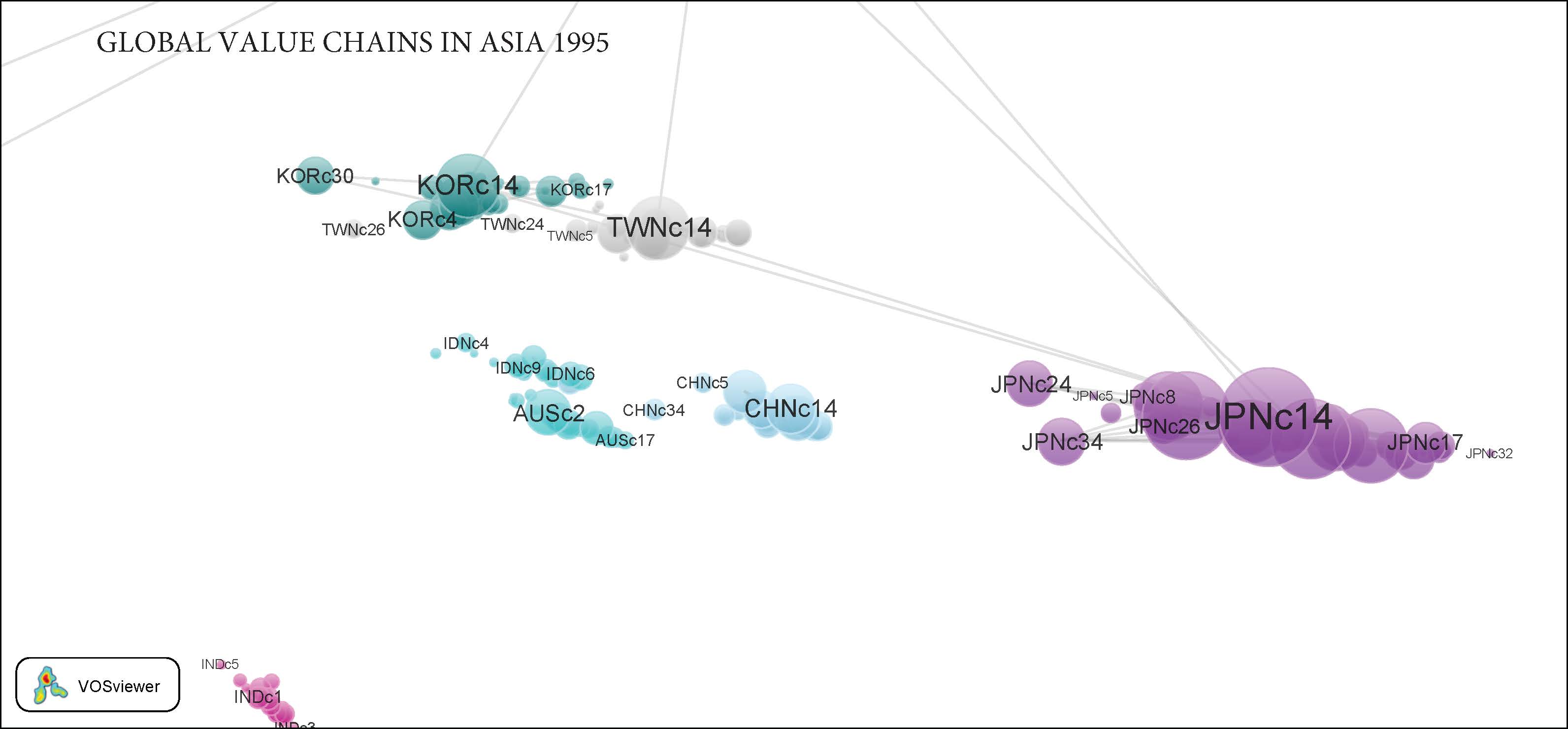
Much more could be said about the evolution on VAGC during the period considered, but it suffices to highlight that in 2011 already Russia, Brazil, Mexico, and India, are more clearly located as value added generating countries, with the first three particularly in the energy value chain included in sector c2 (Mining and quarrying).

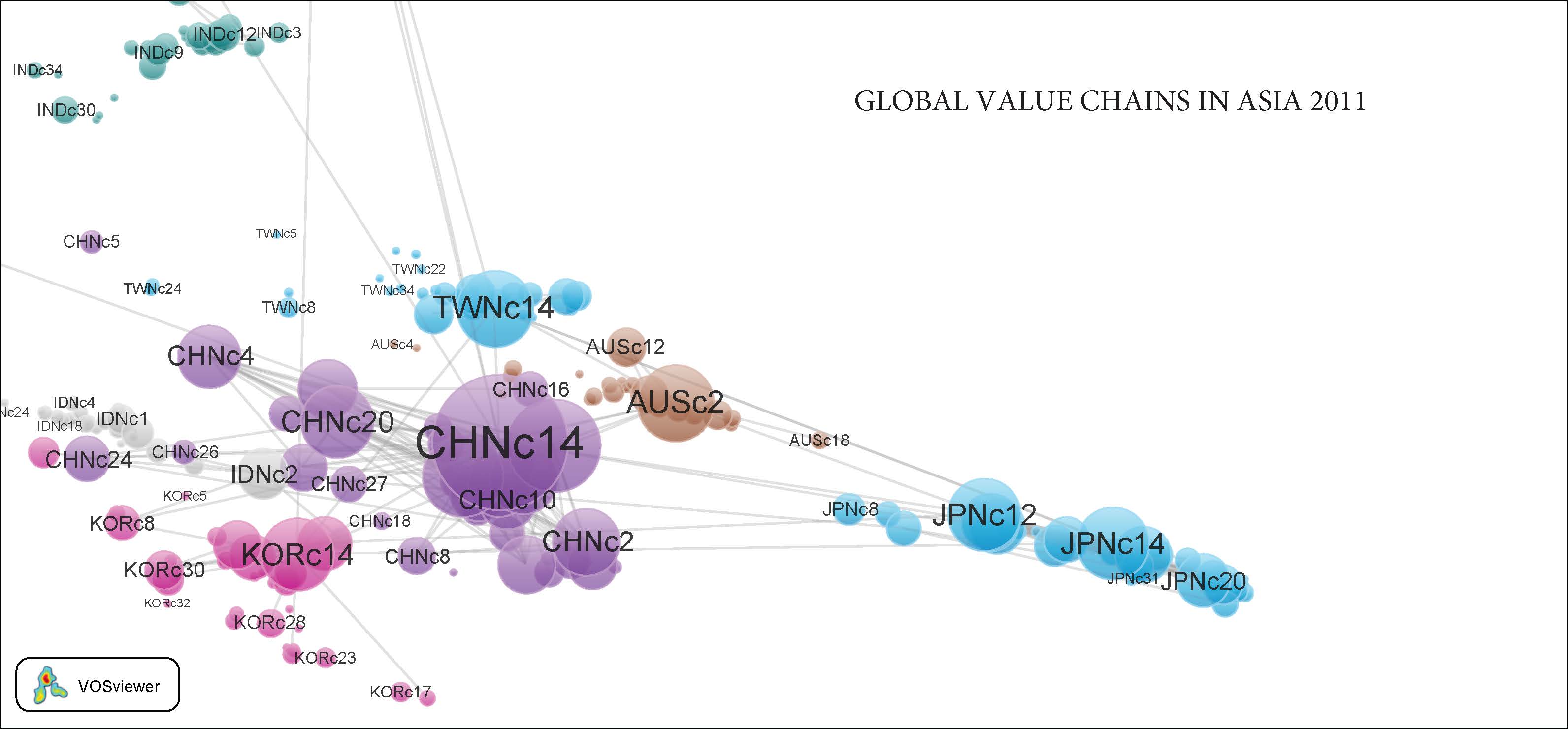
At this level of analysis, visualization of information techniques allow identifying the phenomena general characteristics, the main regional structures, and the underlying structures, as well as the main productive sectors similar in different countries. The analysis can be done for all and each of the 35 countries included in the database.

Another feature of VOSViewer is graph rotation and zooming to focus over specific areas in the graph. These are extremely important functionalities since, as pointed out by Munzner, T. (2000), “One of the main challenges in a visualization system is the manner of representing, as close as possible, all the relevant information given for a finite visualization area. When the structure of interest is too big for perceiving all the details at once, the easiest solution is allowing the user to zoom-in or zoom-out the visible area”.

In the following two graphs we will show the VOSViewer zooming feature using the Asia-Pacific block as an example.

**Graph 5.5 Global Value Chains in Asia**





It is worth emphasizing that the global context prevails: the graph that was created with the 1,960,000 nodes still subsists. We simply zoomed on the area of interest for the researcher. Besides, the zooming may reach the smallest node in an interactive, ongoing process.

In the previous graphs, we can clearly see the significant change that took place in the countries belonging to the Asia-Pacific Basin, where China’s development in terms of value added stands out. As stated a few lines above, the small relative importance of China in 1995 in the Electrical and optical equipment industry (c14), as compared to Japan, Taiwan, and South Korea, is noticed. However, by 2011, China dominated the sector. The zooming also allows visualization of all the sectors linked to this development, mainly the sectors of Basic metals and fabricated metal (c12), Chemicals and chemical products (c9), Renting of machinery and equipment (M&Eq) and Other business activities (c30), and Transport equipment (c15).

**Table 2**

**China’s development (1995 and 2011)**



Source: Authors’ construction using data from WIOD.

Once again, this analysis may be applied to all countries or groups of countries to help re-visualize or focus the analysis without losing sight of the global context.

A multi-level analysis may also be carried out. Once the sectors/countries of interest are identified within the global graph, the nodes or relationships between them can be exclusively visualized.

Unfortunately, VOSViewer automatically maps all nodes. Therefore, if the researcher prefers to map a set of nodes, he first needs to manually inactivate one by one to isolate the nodes of interest. This procedure turns out to be very impractical, particularly when we are speaking of 1,400 nodes, like in this case, or more.

An alternative solution is dividing the VAS\_Ê matrix which, as already explained, includes all direct and indirect contributions to value added included in a specific sector’s exports so that we can exclusively consider the sector of interest.

For clarification purposes, we will consider the sectors c30 Renting of M&Eq and other business activities, c14 Electrical and optical equipment, and c12 Basic metals and fabricated metal. In each case, the process is repeated and, since we have been already working with 40 X 40 pre-cut VAS\_Ê matrices, the relative place of countries in the graph changes *versus* the total graph. However, with VOSViewer it is possible to analyze clusters in each sector and, if information from several years is available, see how they change in time.

For example: in 1995, countries in sector c30 were grouped around 5 clusters, of which the most important was the group of European countries, headed by Germany and France (26 countries). The second cluster comprised the Asia-Pacific countries (6). In 2011 there were already 4 clusters and their configuration had changed: the first cluster included two other countries (28), with Holland as the leader. However, the most surprising feature is that in the second cluster, the NAFTA countries (US, Canada, and Mexico) are integrated with some Asia-Pacific Basin countries: China, Japan, Taiwan, and Australia.

A similar analysis may be applied to all and each one of the value added chains, identifying country’s groupings and their development overtime, zooming in and out the respective graphs, and rotating them as the researcher considers in order to obtain the needed knowledge on the studied phenomenon. Data, information, knowledge…Indeed, most networks only include data but, as Brender (2005) said: “… visualization of networks may be designed in such a way as to include information, hence becoming a basis for knowledge.”

The graphs of the three main value chains mentioned above (comparing 1995 and 2011) are found in Appendix 2.

1. **Final remarks**

Summing up, the extensive use of information in WIOD (Big Data), together with a methodology that provides ICIO models, the development of adequate programs, and the use of new programs as VOSViewer for handling considerable data volumes, particularly for information visualization, opens new and promising possibilities for analysis in social sciences.

However, much remains to be done. As the attentive reader must have noticed, the available information is extremely grouped. In Appendix 2 we have described in more detail the economic activities included in the main sectors used as examples, and their diversity is such that indicators may be used only for orientation for more accurate research and for broader economic policy decision making.

Besides, part of the information included in WIOD represents mere estimations. In the most extreme cases, an entire period was estimated using actual information from a single year. A consensus would be needed among countries to generate statistics susceptible of enhancing the effort developed by WIOD.

To conclude, the same attentive reader must have perceived that the visualization technique introduced in this work does not allow for identifying the origin or the destination of a global value chain, or whether it is oriented to production or consumption. It does however allow determining the main productive chains generating value added and their contributing countries, and this is definitely a considerable achievement!

1. **References**

Baldwin, R. (2012) Global Supply Chains: Why They Emerged, Why They Matter, And Where They Are Going. Centre for Trade and Economic Integration. *CETEI Working Papers*. Graduate Institute of International and Development Studies. Geneva and Oxford University.

Bertin, Jacques (1981) Graphics and Graphic Information Processing. Ed. Walter de Gruyter. Berlin. New York. 273 pp.

Brandes, Ulrik et al. (2005) La explicación a través de la visualización de redes. *REDES-Revista Hispana para el análisis de las redes sociales*. Vol. 9, No. 6. <http://revista-redes.rediris.es>

Dietzenbacher, E. and Isidoro Romero (2007) Production Chains in an International Framework: Identification by Means of Average Propagations Lengths. *International Regional Science Review*. 30, 4: 362-383. <http://irx.sagepub.com/content/30/4/362>

Koopman, R. et al. (2010) Give Credit Where Credit Is Due: Tracing Value Added in Global Production Chains. *NBER Working Paper* 16426. <http://www.nber.org/papers/w16426>

Munzner, Tamara (2000) Interactive Visualization of Large Graphs and Networks. *Ph. D. Dissertation*. Department of Computer Science of Stanford University.

Nivón Zaghi, A. (2011) Visualización de la Información Estructural de Matrices Insumo Producto, por medio de Técnicas de Trazado de Grafos. Instituto de Investigaciones Matemáticas Aplicadas a Sistemas. Universidad Nacional Autónoma de México.

Solís Arias, V. (2012) Algunas relaciones entre los sectores real y financiero en la Euro-Área (mimeo).

Timmer, Marcel (ed.) (2012) The World Input-Output Database (WIOD): Contents, Sources and Methods. WIOD Project funded by the European Commision. *Working Paper Series*. Working Paper Number 10.

Chung Wong, Park (2007) Information Visualization. *Visual analytics science and technology* 6, 1–2. doi:10.1057/palgrave.ivs.9500149,

Van Eck, N.J. and L. Waltman (2010) Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, 84(2), 523-538.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ et al. (2010) A Comparaison of Two Techniques for Bibliometric Mapping: Multidimensional Scaling and VOS. *Journal of the American Society for Information Science and Technology*, 61(12), 2405-2416.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ (2013) Manual for VOSviewer. Version 1.5.5 Universiteit Leiden. CWTS Meaningfull Metrics.

**Appendix 1**

List of countries and codes in WIOD-database

|  |  |  |  |
| --- | --- | --- | --- |
| **European Union** | **CODE** | **North America** | **CODE** |
| Austria | AUS | Canada | CAN |
| Belgium | BEL | United States | USA |
| Bulgaria | BGR |  |  |
| Cyprus | CYP | **Latin America** |  |
| Czech Republic | CZE | Brazil | BRA |
| Denmark | DNK | Mexico | MEX |
| Estonia | EST |  |  |
| Finland | FIN | **Asia and Pacific** |  |
| France | FRA | China | CHN |
| Germany | DEU | India | IND |
| Greece | GRC | Japan | JPN |
| Hungary | HUN | South Korea | KOR |
| Ireland | IRL | Australia | AUT |
| Italy | ITA | Taiwan | TWN |
| Latvia | LVA | Turkey | TUR |
| Lithuania | LTU | Indonesia | IDN |
| Luxembourg | LUX | Russia | RUS |
| Malta | MLT |  |  |
| Netherlands | NDL |  |  |
| Poland | POL |  |  |
| Portugal | PRT |  |  |
| Romania | ROU |  |  |
| Slovak Republic | SVK |  |  |
| Slovenia | SVN |  |  |
| Spain | ESP |  |  |
| Sweden | SWE |  |  |
| United Kingdom | GBR |  |  |

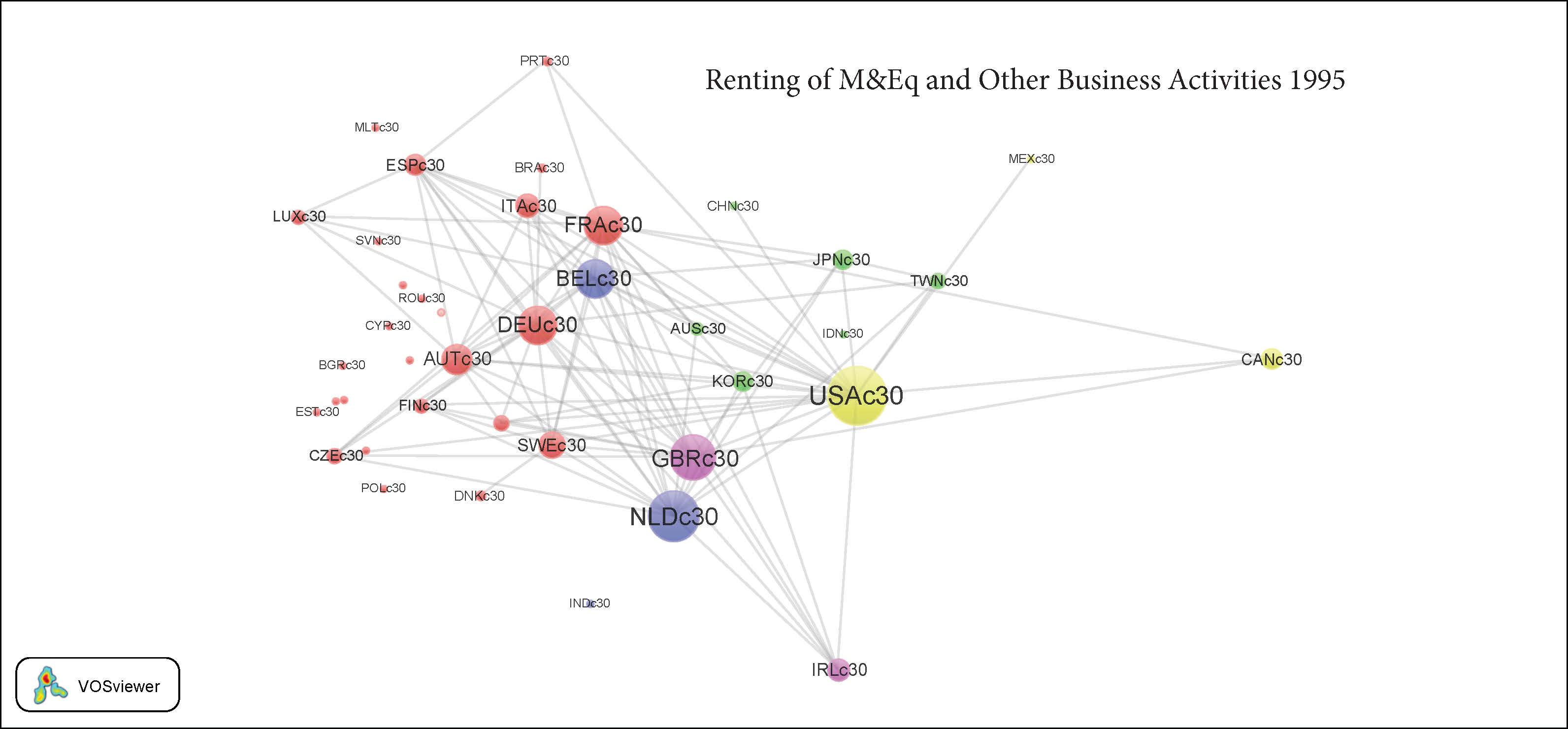
List of sectors in WIOD-database

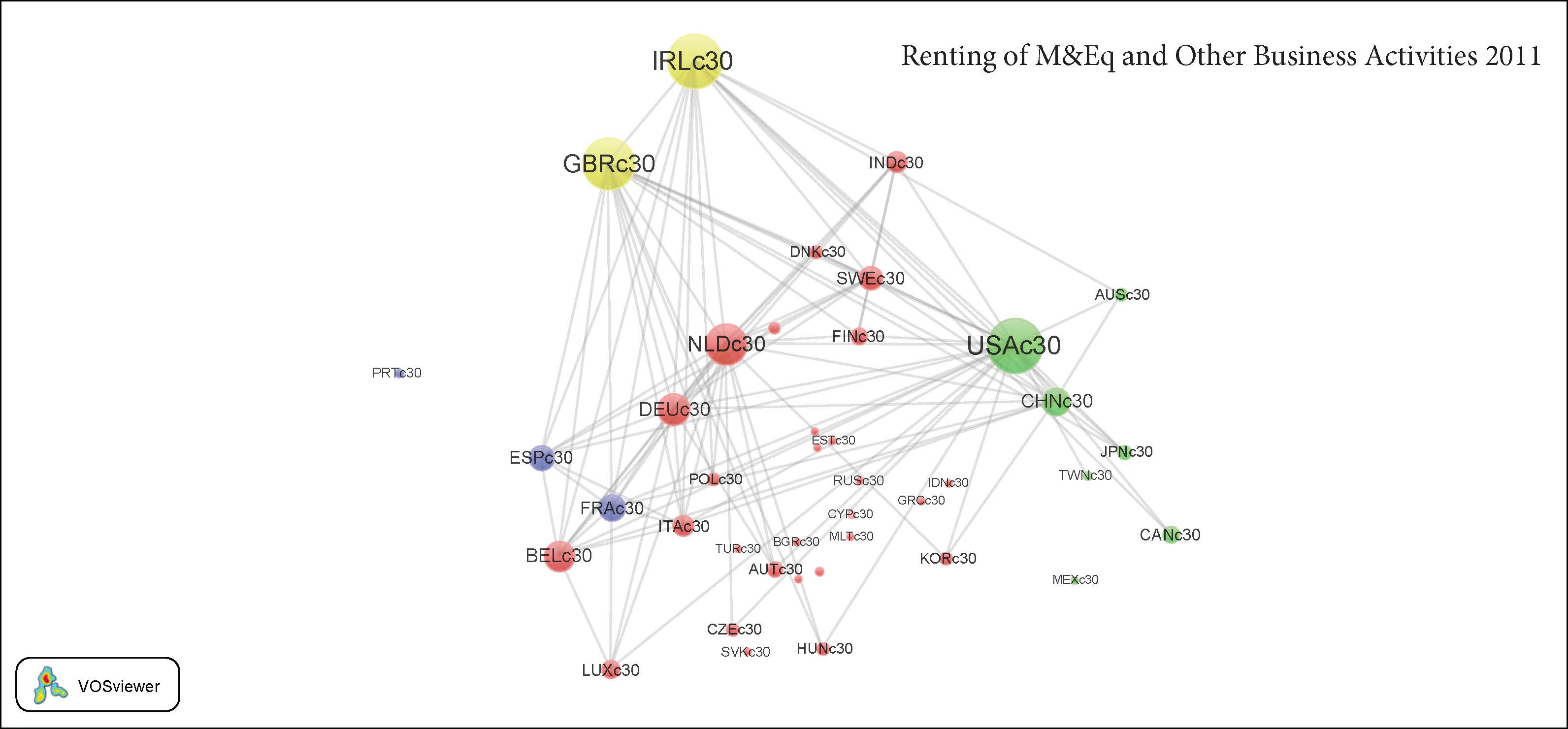
|  |  |
| --- | --- |
| C1 | Agriculture, hunting, forestry and fishing |
| C2 | Mining and quarrying |
| C3 | Food, beverages and tobacco |
| C4 | Textiles and textile products |
| C5 | Leather, leather and footwear |
| C6 | Wood and products of wood and cork |
| C7 | Pulp, paper, printing and publishing |
| C8 | Coke, refined petroleum and nuclear fuel |
| C9 | Chemicals and chemical products |
| C10 | Rubber and plastics |
| C11 | Other non-metallic minerals |
| C12 | Basic metals and fabricated metals |
| C13 | Machinery, Nec |
| C14 | Electrical and optical equipment |
| C15 | Transport equipment |
| C16 | Manufacturing, Nec; recycling |
| C17 | Electricity, gas and water supply |
| C18 | Construction |
| C19 | Sale, maintenance and repair of motor vehicles and motorcycles; retail sale of fuel |
| C20 | Wholesale trade and commission trade, except of motor vehicles and motorcycles |
| C21 | Retail trade, except of motor vehicles and motorcycles; repair of household goods |
| C22 | Hotels and restaurants |
| C23 | Inland transport |
| C24 | Water transport |
| C25 | Air transport |
| C26 | Other supporting and auxiliary transport activities; activities of travel agencies |
| C27 | Post and telecommunications |
| C28 | Financial intermediation |
| C29 | Real estate activities |
| C30 | Renting of machinery and equipment (M&Eq) and other business activities |
| C31 | Public administration and defense; compulsory social security |
| C32 | Education |
| C33 | Health and social work |
| C34 | Other community, social and personal services |
| C35 | Private households with employed persons |

**Appendix 2 Visualization of Main Global Value Chains**

**C30 Renting of Machinery and Equipment (M&Eq) and Other Business Activities**

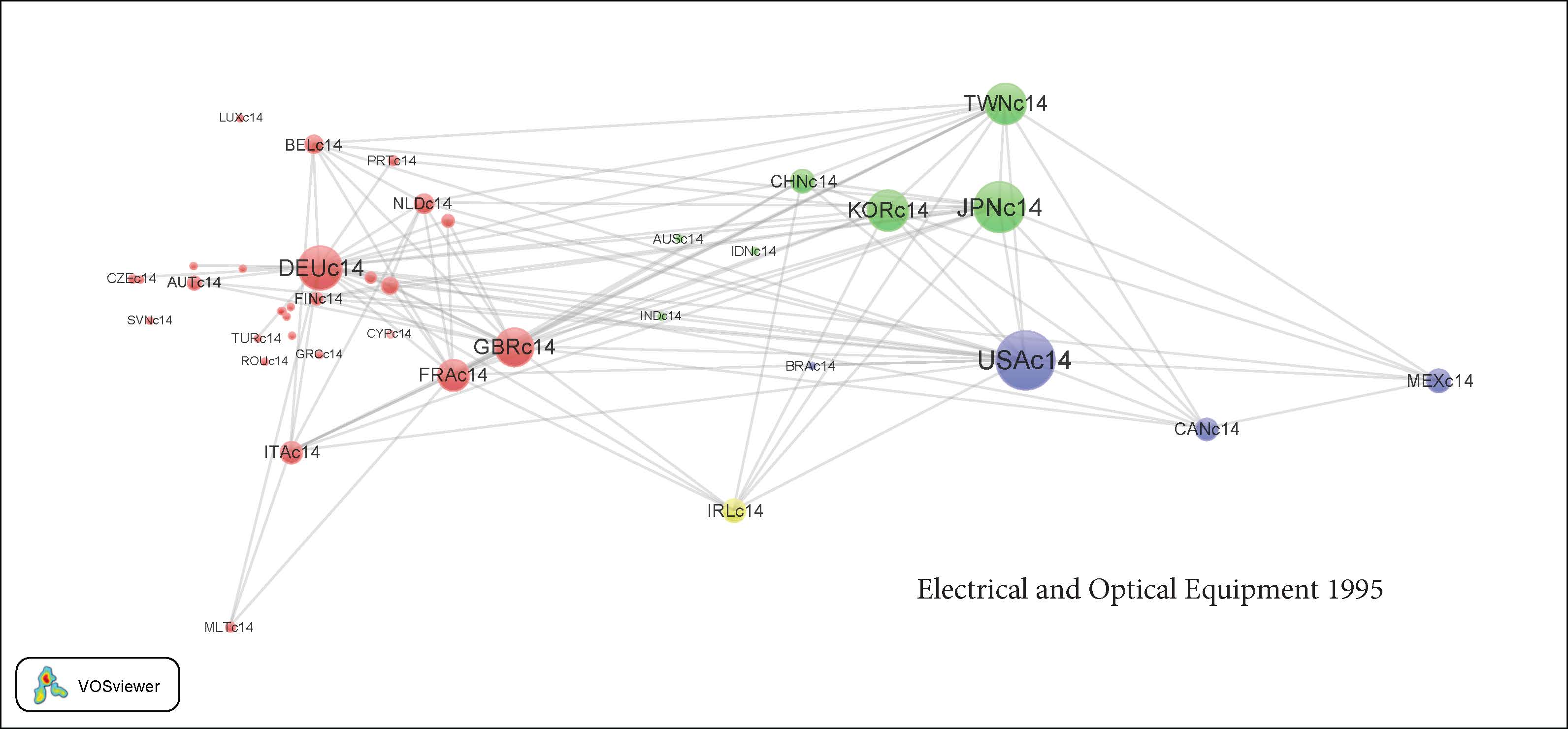
Includes the renting of automobiles and other transportation equipment (land, water, and air); the renting of machinery and equipment such as agricultural, construction and civil engineering machinery; office and equipment machinery including computers, and personal and household goods. Computers and computer-related activities (hardware and software consultancy, data processing, maintenance, and repair). Research and development and other business activities (including legal, accounting, book-keeping and auditing activities).

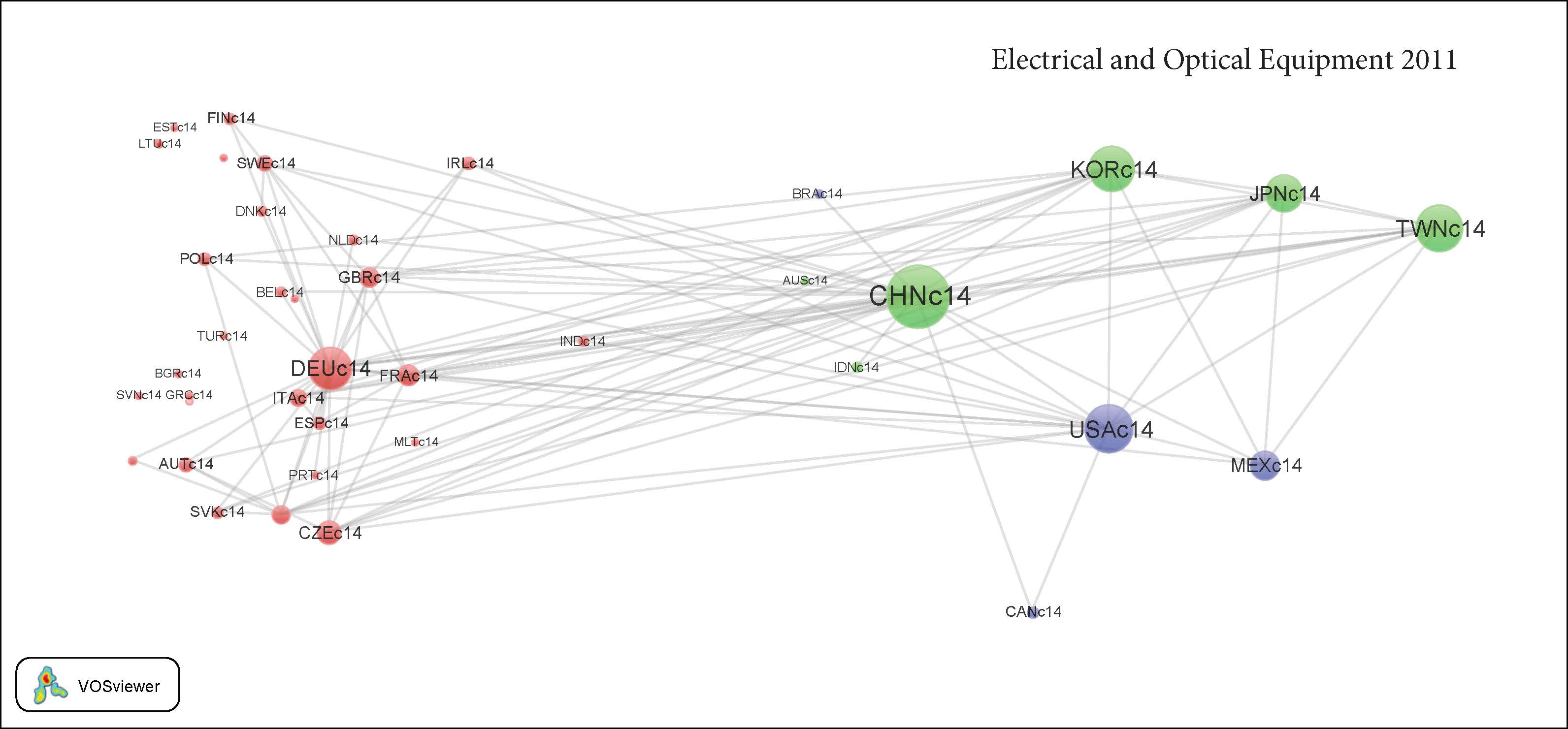
****

****

**C14** **Electrical and Optical Equipment**

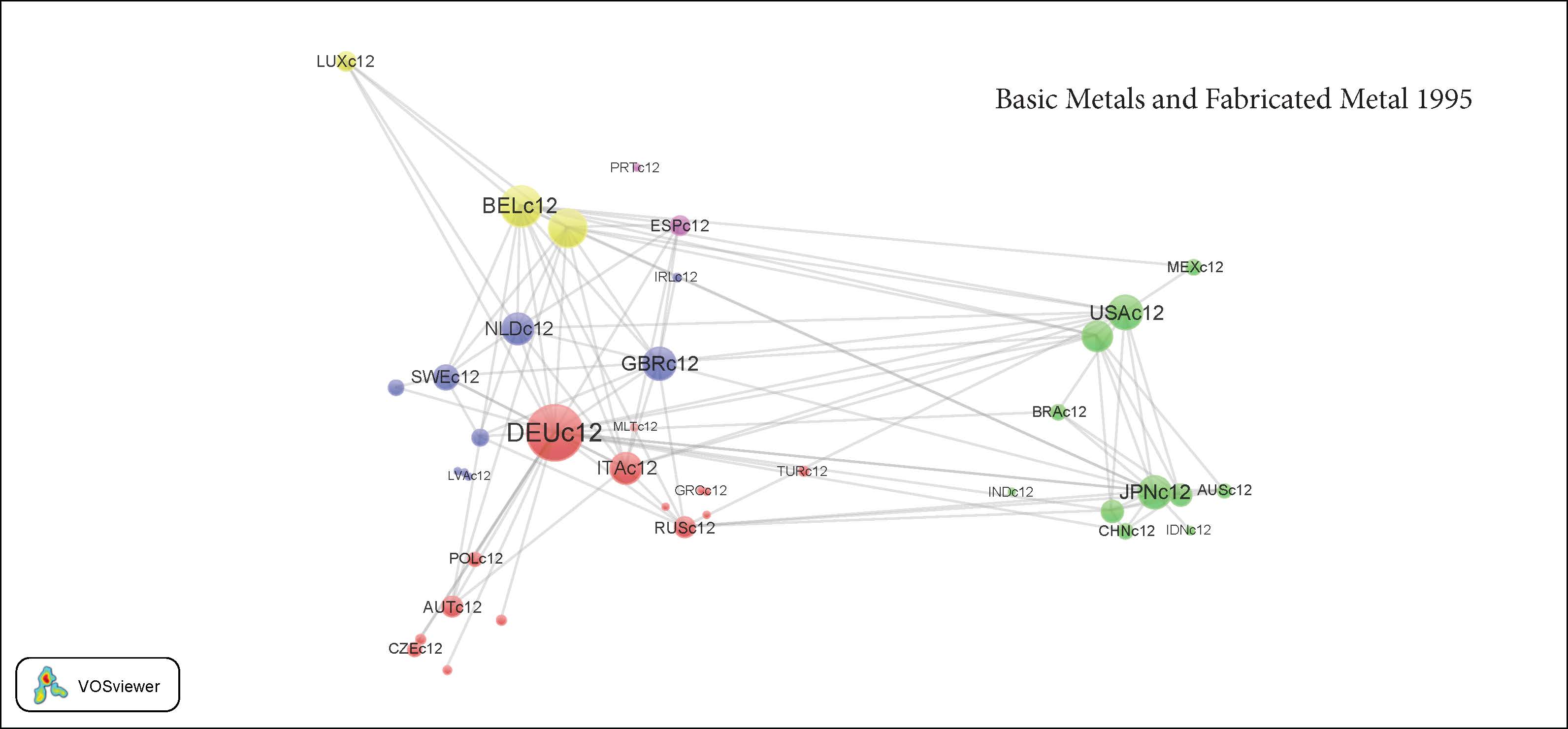
Includes the manufacturing of office machinery, computers, electrical machinery, and devices. Manufacturing of radio, television and communication equipment and of medical, precision and optical instruments, watches, and clocks.

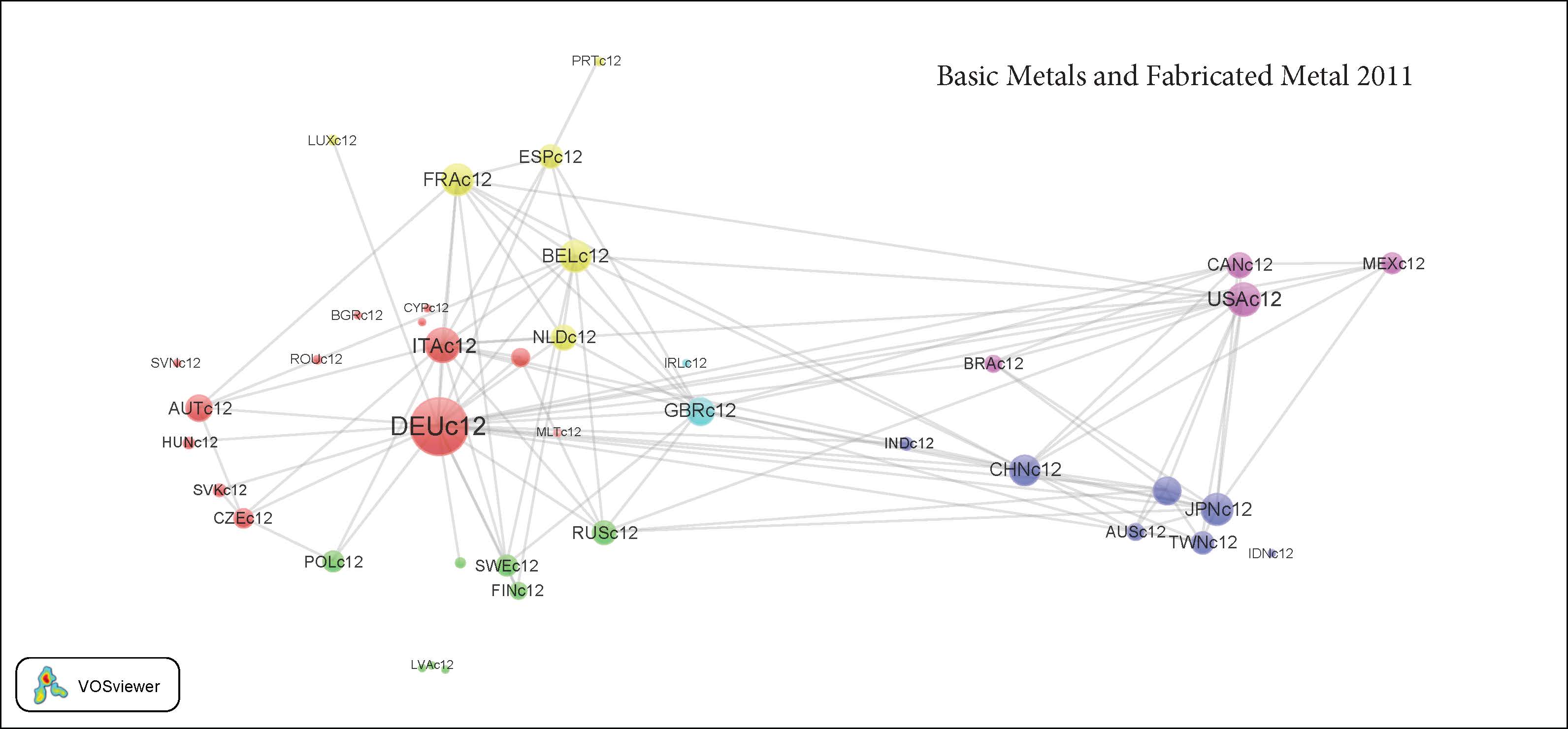




**C12** **Basic Metals and Fabricated Metals**

Includes the manufacturing of basic iron, steel and iron alloys; the manufacturing of pipes, and basic precious and non-ferrous metals (aluminum, zinc, cooper). The manufacturing of fabricated metal products, (except machinery and equipment), like metal structures, tank reservoirs and metal containers; central heating radiators and boilers; steam generators; forging, pressing, stamping and roll forming of metal; and of cutlery, tools and general hardware.





1. The authors would like to thank Valentín Solís for research guidance. Martín Puchet provided valuable comments for the preliminary version. The points of view herein stated are exclusively of the authors. [↑](#footnote-ref-1)
2. The authors acknowledge the work of Roberto Carlos Orozco in implementing this program. [↑](#footnote-ref-2)