Patent Search For Engineers And Lawyers
Prof. M. Padmavati
Rajiv Gandhi School of Intellectual Property Law
Indian Institute of Technology, Kharagpur

Lecture - 25
Introduction to patent landscape (Contd.)

Welcome to the aspects of patent data analysis in relation to patent landscapes. In the earlier lecture we have understood the aspects of how to go about with the different steps in relation to patent landscape search, the value of looking at technology focus, the necessity to look at the aspects of whether it is a broader search or a narrow search, what is the importance of looking at citations information in relation to patents.

(Refer Slide Time: 01:01)

So, patent data analysis points can come from various aspects of the patent landscape. So, there could be structured data and they could be unstructured data. Structured data is something in relation to what we have as bibliographic data, which is the metadata in relation to patents that is available. The organization of that data is much simpler so, classifying that is much easier. A lot of unstructured data, that is present in patents typically the drawings, claims, these become difficult in order to classify and visualize in patent landscapes.

Today, we have enhanced tools available. So, that you can actually do a mouse hover and the drawing the relevant drawing comes up. For instance if I am looking at an invention
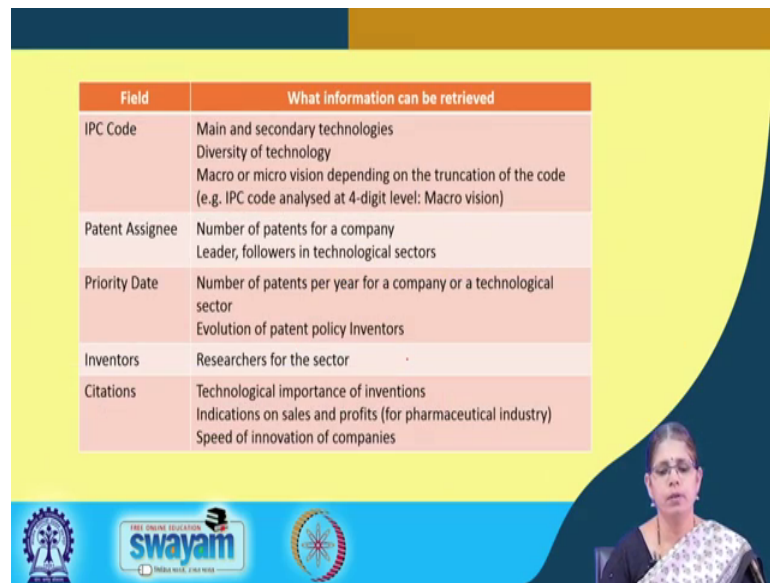
in a given sub domain of a particular technology area, and I am interested let us imagine in the various improvements of a particular device. The concentration here will be what components of that particular device. So, we are looking at specific type of interventions in relation to components. It is not only important to look at the textual part of the data in terms of what are the improvements. It is also sometimes important to look at the information in relation to images that the patent captures for instance drawings.

So, today the subscribe databases provide an opportunity to do a mouse hover and you would actually get the images in relation to patents. So, looking at structured and unstructured data can make a big difference in relation to the way patent landscapes are prepared and also interpreted. A lot of analysis has become easier today, because of statistical tools available in relation to patent datasets. Not only that today one can actually do a network analysis, which means that comparisons can be drawn between one data set to the other data set. And, this again provides a lot of information in relation to looking at correlative information in relation to certain aspects of patents.

So, data mining and data analysis has become much easier today. And, this is an area which has helped actually now classified 1000s of patents as per certain areas. Earlier for a patent searcher one needed to actually look at downloading the data and organizing individually the graphs or the pie charts. Today this is done on the flight. So, this is where a lot of enhanced display which is being provided by software tools is a great opportunity for patent landscape, searchers to look at patent data across different domains or even across different datasets. Text mining is another area which has grown enormously. Today, one can rank the data, similarly one can rank patents.

The patent ranking can be undertaken by different ways based on the content, based on the value, based on certain concepts. It is also today easy to look at similarity indices in relation to patents. Software tools today provide for these also as filter options. So, you can actually ask for it for a given patent you can ask for similar documents. So, the growth of text mining tools has also enabled easier preparation of patent landscapes.

(Refer Slide Time: 06:01)

Let us understand how information is retrieved in relation to a patent landscape. So, if those areas which represent a wide domain IPC becomes the main context for the basis of doing their landscape. So, what is the information that we can get from the IPC based retrieval of information. We can get the information relation to the main and the secondary technologies, one can understand what is the diversity of technology or rather the growth of a particular technology.

Since, IPC codes represent the classification to different levels. So, one can actually go to the different digit level analysis. So, that is where one can see the depth of information in relation to IPC. Today patent landscapes can also provide us great opportunity to do what we call assignee analysis. So, in the entire patent landscape one can actually position different assignees based on technologies. So, many a time landscapes also provide us some surprising information for instance there are instances where you would realize that, here is a cement company, but working on energy. And, there is a construction company working on certain areas of electronics.

So, patent landscapes do reveal such interesting information and it also helps to provide, where is the assignee moving with respect to technology. So, the portfolio which we talked about is understood from also the assignee analysis. So, in a given period of time has the assignee diversified from the point of view of main technologies to other technologies. So, it actually helps plot the information in relation to various companies in

that particular segment. This is particularly relevant for the area of what we call platform technologies, these are a set of technologies which are operable across different areas.

So, they are lead technologies, they are not necessarily with respect to certain apparatuses or certain specific methods some of these technologies are operable across the different segments. So, wherever platform technologies are involved, it is always a important question that a patent researcher keeps in mind from the point of view of looking at what are all the sectors relevant to that particular set of technologies. Another, important information that a patent landscape can provide is the information in relation to priority of patents.

What is the information that we get from priority data? Priority data provides us in relation to a particular technology sector, how much of the patents have been filed. So, one can get priority vis a vis assignee priority vis a vis country. Today, search is also done to understand inventor information, researchers are important from the point of view of development of certain technologies. So, tracking inventor information is also one part of the patent landscape search. So, one could go for top 10 inventors. And, sometimes today inventions are brought out in a collaborative mode. So, it is also helps to understand invent of movement or invent of work in a particular technology area.
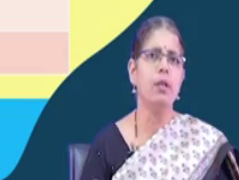
Citations are an important aspect of patent landscapes. They provide you information on the importance of a given technology in a particular area. In fact, the simplest of the patent valuation methods is what is the citation analysis? And, these are relevant for all industries, irrespective of the technology area. The citation value also provides information on the rapidity of the implementation of a particular technology by looking at the forward citations.

So, the when we look at citation information based on a given time point or based on a given technology timeline one can actually go for backward citations and as well as forward citations. Today there are a lot of databases, which specifically provide information in relation to citation analysis. So, what one can do is? From the landscape a set of patents can be taken out and put in this particular citation tool which provides for maps in relation to understanding citation data.

So, there are different data sets, which are useful in relation to patent landscapes. Another important aspect of the patent landscapes is what we call cross referencing. Now, cross referencing again helps in understanding relationships between the data. So, one can have a cross referencing in relation to priority year and assignee publication as well as let us say country.

So, among the different fields that are considered for analysis, one can actually also utilize the cross referencing information in relation to a patent landscape. So, one can actually go for inventor priority year, priority country, patent family IPC code assignee. Each of these will provide for a cross reference with respect to. So, one can actually draw a grid like this and identify what type of cross referencing information is useful in relation to a particular landscape.

So, for instance if you are looking at the IPC code and priority country, what is it that we can look at we are looking at the activity fields per country. Let us say we are looking at sometimes information relation to patent family and country. How does that help us, when you look at cross referencing between a patent family and a priority country you get information on the filing strategies at national levels.

So, it helps you to actually draw a lot of conclusions out of a patent landscape on different aspects. Another aspect of looking at let say patent family and priority here would be to look at the evolution of interest in the national markets. So, unless there is a

need for the product in a market, that particular geography would not be selected for filing. So, these insights are provided from the cross reference information that none can do in relation to patent landscape analysis.

So, you can imagine a whole lot of data can be obtained by cross referencing information. So, in a patent landscape you can have individual information, for instance you are looking only at the application and grant data in relation to a time period. So, that is specific information. And, then you have cross referencing information where you can have a whole gamut of this particular data correlation between different fields. So, all of that can be assembled as what we call the conclusions that come out of that particular patent landscape.

(Refer Slide Time: 15:45)
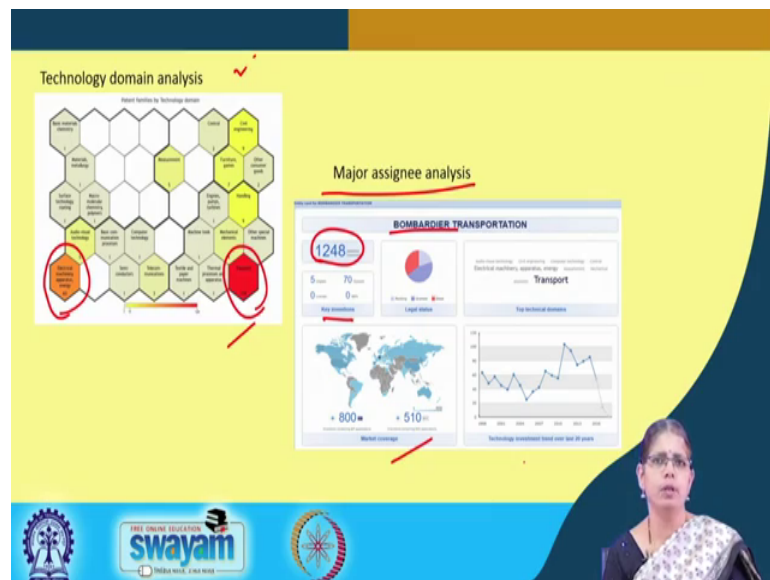
This is an illustration to indicate how visualization tools are helping us today to look at data in relation to patent landscapes. What we see here in this panel is the technology domain analysis. So, based on the keyword or the IPC classification code that you have used to search for the patents, you get a particular patent collection.

Now, if you ask for the visualization tool to organize your patent data set as per technology domain focus you may get a visualization information, which is shown in this particular illustration. The color coding will represent the number of patents present in that particular data set. Not only that it also provides you where is the major focus of inventions in a particular area. For instance here we are looking at this particular block.

These blocks represent the major interventions in that particular technology area in relation to inventions.

So, one can actually take out that specific data set and they do the analysis further. So, today in the early phases of the patent landscape, you have tools available which will help you to sift through the data and get the enhanced information from a patent landscape. In the earlier times when patent landscapes were done, one needed to actually wait till the end of the patent landscape from the point of view of data analysis to derive conclusions. Today, midway through the patent landscape one can actually take out the information and continue the analysis on a different mode.

Let us look at some other information which visualization tools also provide these days, for instance major assignee analysis. Now, here is a technology which is about a specific aspect of transportation. And, in this case bombardier is the assigning, which is identified as a major assignee, representing a patent dataset collection of 12 48. What are the key inventions are represented? What is the legal status of the patents? Where are the different geographies or the market coverage in relation to that particular area is represented. And, what sort of technology investment is there over a period of 20 years.

So, what is the unique aspect that we are actually looking at here, that here we have simultaneous representation of different sets of information on the same block. So, this provides for easier interpretation in relation to a data.

(Refer Slide Time: 19:17)
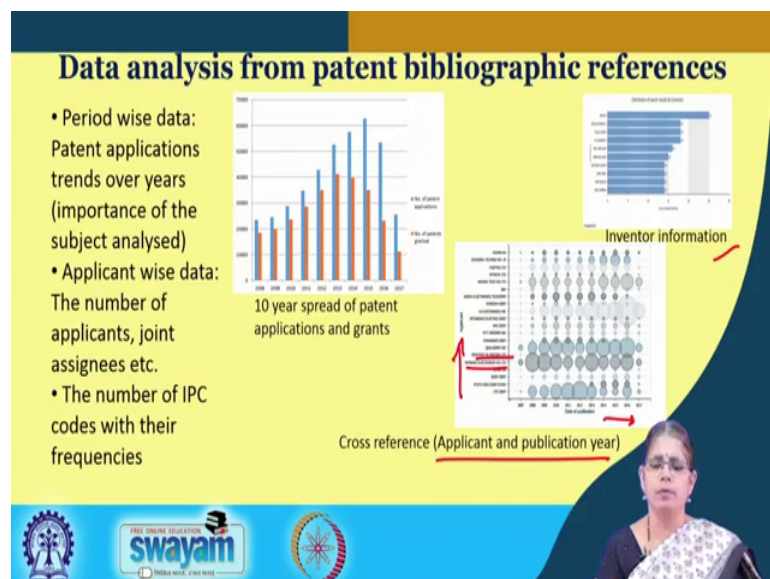
A lot of data can be analyzed based on patent bibliographic information. We can look at period wise data, applicant wise data, and understand patent trends, for instance here we are looking at a 10 year spread of patent application and grants, how is this information helpful? So, among the total set of patents that are available you have year wise information in relation to that particular area, it is also telling us the number of granted patents in a given area.

So, how much of application filing is happening, this is typically the information that you get from published patent application information and how many of them are moving into grants? So, it pretty much gives us an idea about the growth of a particular technology in relation to patent filing and grants. And, one when one is looking at emerging areas. So, if one way to look at the objective of a patent landscape is to plot just the emerging areas in a given technology area. Normally, till last 5 year or the last 10 year information is relevant.

In which case you are doing a patent landscape taking into the consideration of a time period, which is 5 years or 10 years. And, this is a simple information which you can get in terms of the period wise data. One can actually get a lot of information in relation to who are the people, who are actually practicing this particular technology by the assignee wise information and also one can also look at the inventor information.

So, what you see in this panel is the inventor information in relation to a particular area. So, you know which are the inventors, which are the major inventors in a particular area, again here either the top 10 inventors or the top 5 inventors are a consideration, because those inventors have significant number of patents in relation to given area. We talked about cross referencing one aspect of cross referencing is looking at the applicant and the publication your information.

So, here you have on the x axis the date of publication and on the y axis you have the applicant information. So, when you overlay both of these you can see the spread of the information and one can actually look at who are the people, who have consistently maintained for instance here you are looking at research in motion, Samsung, Qualcomm and ZTF some of these are consistently moving with patents. So, the spread of patents information in the form of these visualization tools helps us understand in a more easier way the relevance of certain data.

(Refer Slide Time: 23:05)
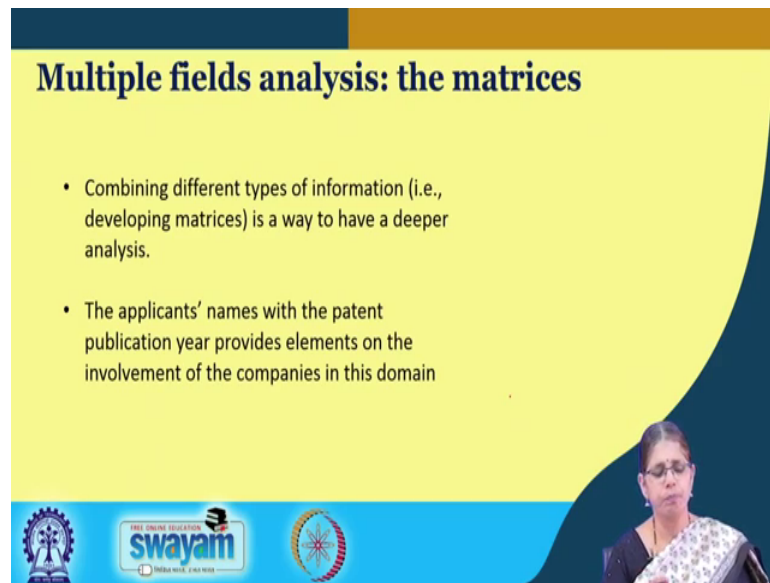
So, top 10 inventor analysis information is useful for instance today the one of the important areas is the areas of gene editing, there have been institutional disputes in relation to gene editing between two large institutes; one of the important things there would be to look at the inventor portfolio, who are the inventors in relation to those set of technologies. And inventor information is therefore, useful in many other industries. So, the inventor information is useful from the point of view of understanding the background research being undertaken and the growth of the inventors from the point of view of the growth of filing in that particular area.

Country based filings information can be obtained by geometrical distribution of patents in a given from a given data set. So, patent status information is also useful from the point of view of understanding, how many patents are filed, how many patents are abandoned, how many patents are valid? And, through a pie chart representation either in the overall technology or based on the sub technologies one can understand the patent status information in relation to a particular landscape.

(Refer Slide Time: 24:47)
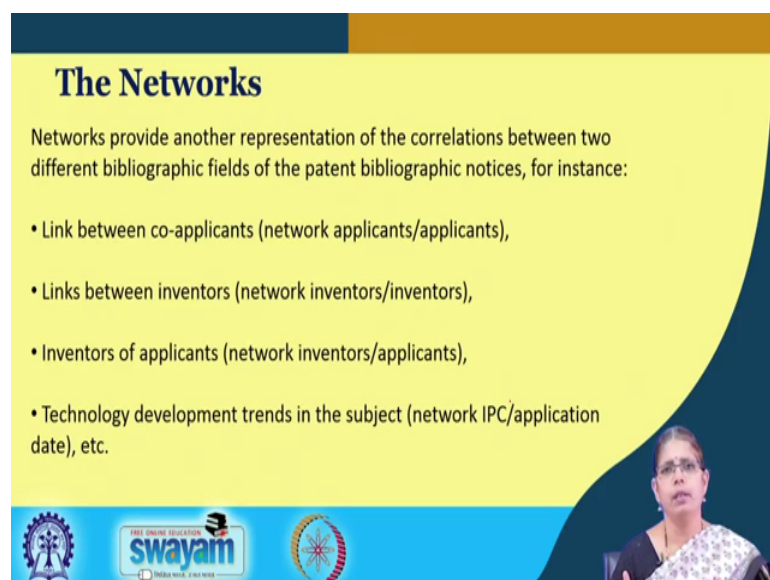
The entire enormity of a patent landscape is from the point of view of extraction of the value of the patent landscape. So, one can actually combine different types of information to get a deeper analysis of the patent landscape. A lot of time is invested not only in collecting of the patent dataset, but actually in the analysis in relation to patent landscapes.

(Refer Slide Time: 25:33)

Today, patent landscapes can provide network information, because of the visualization tools that are available. A lot of information can be obtained in relation to who are the co
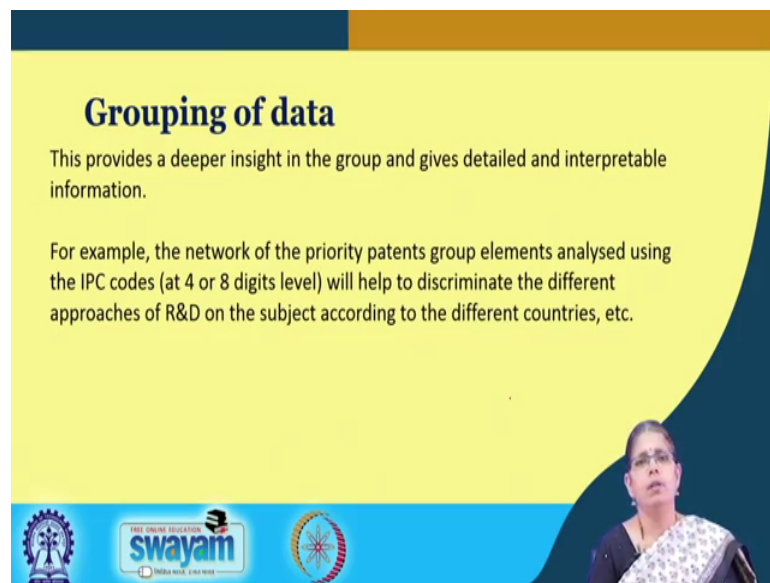
applicants, link between inventors, inventors and applicants, subject area based information spread. So, today from the analysis tools that are available, one can actually see a lot of network information in relation to the patent data set that is collected. And, typically this involves the use of at least two field of for the preparation of the network.

(Refer Slide Time: 26:39)
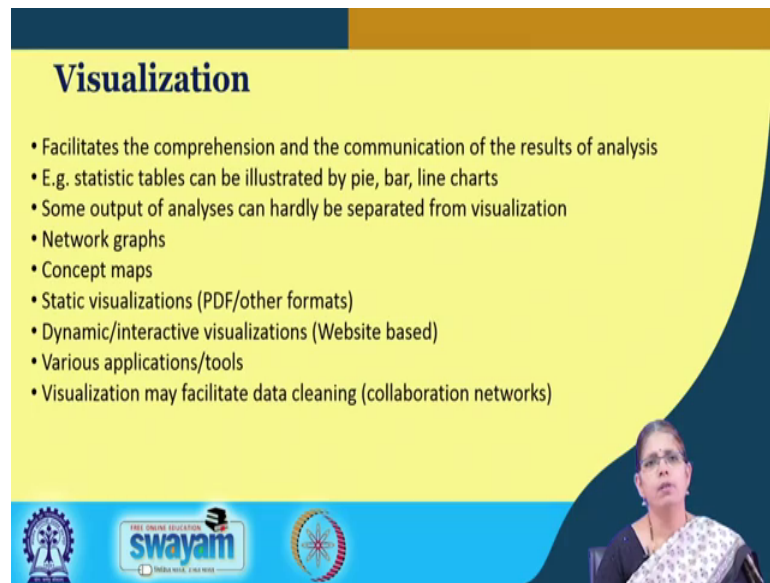
Grouping of data is important in relation to patent landscapes. So, how would one group data depends on actually the objective of the patent landscapes. So, you can have grouping of data for the ancillary purpose of understanding the main focus of the patent landscape or you can have different sets of groups, which are understood from the point of view of a objective. Again this helps in providing the interpretation of the patent landscape in a given scenario.

(Refer Slide Time: 27:23)

So, visualization as we have discussed provides us this great opportunity and this is also interestingly an area where the lot of patents have been filed. So, if you look at a keyword called patent landscape and let us say you put it in Google patents or you put it in any one of those patent databases. You would actually get patents in relation to tools and methods available for preparing patent landscapes. So, this is where the value of the patent landscapes is being realized from the point of view of providing information not only in a rapid manner, but also value based information.

(Refer Slide Time: 28:13)

The conclusions and recommendations are an important part of the report. Based on the objective of the patent landscape, there are several conclusions one can draw. Conclusions may be made on the R and D opportunity available who are the potential players with whom the merger can be taken up on acquisition of I P portfolio is of value. Potential licensing information in terms of who are the licensees trends in relation to patent information.

So, a whole lot of conclusions can be drawn and based on that recommendations will need to be made in relation to the patent landscape. So, that the value of patent landscape to a particular entity is realized.

(Refer Slide Time: 29:17)

Is it also a scenario where lead inventors will need to be hired such decisions can be taken by companies, what kind of a growth strategy is provided let say quarter wise. And, where are the gaps available for further work in relation to that particular technology?

From here on we look at certain examples of how a patent landscaper can be taken up from the point of view of the 4 steps that we discussed. So, this is an example what you see on the screen is a colony stimulating factor which represents a whole lot of developments. So, this is a molecule now for CSF one is the aspect of how this molecule is made? What are the conjugates? What are the different type of formulations that are there, purification techniques are there any fusion proteins.

So, in this illustration what you are trying to understand is the mapping of the technology? So, in relation to this given molecule, what are all the developments that are possible? A simple way of understanding the technology focus area is to read a good review paper in this particular area, or as a person from this specific area you would already actually be having the knowledge of what are the different aspects of this particular molecule that are being worked currently.

Once you have this mapping done then one embarks on either looking at a general landscape we are looking at all improvements in relation to CSF or you are looking at a narrow landscape, where you are looking at certain aspects of CSF and the growth of technology in that.

(Refer Slide Time: 31:47)
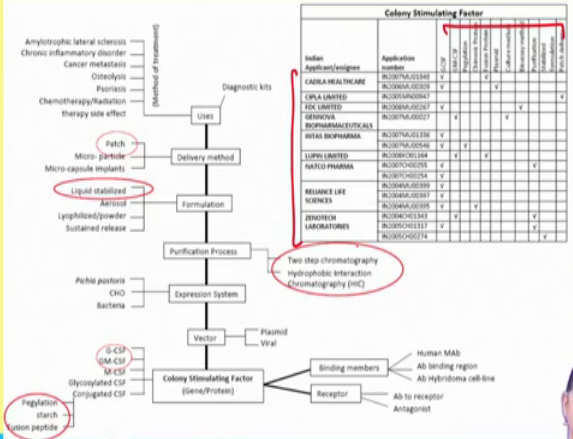
So, the search strategy will depend on where the focus of the landscape is. Now based on the data that is collected one can actually go in for the analysis of information, where the idea would be to either just look at the filing trends or to look at developments in relation to this particular molecule. So, the landscape study can look at different aspects, it may look at the focus of patenting in a specific geography for instance only in a specific country, in which case only those set of patents are analyzed on the global landscape of the technology.

(Refer Slide Time: 32:31)

So, this is a simple landscape analysis of colony stimulating factor, taking into consideration both the aspects of gene as well as protein. What you see in the inset here is the excel sheet, which is the basic sheet in relation to cataloging all the information in relation to different patent applications. And, then looking at what are the areas that they represent from the point of view of the remember the original technology map and then who are the applicants, who are actually having the patents here.

So, you can actually have a gridded sheet like this. And based on that understand where the innovations are coming up. Now, what are circled in red color here are the second level of innovations that are coming up on the CSF molecule. So, if you look at the general spread that is on a general map patent line in this particular patent landscape, you are looking at which are the emerging areas in particular in relation to that particular molecule.

(Refer Slide Time: 33:59)

Another example is the analysis of patenting trends for earth mobility technologies. In this case we are looking at the post Kyoto protocol that is when the standards were put in place in relation to commitments for different countries. So, if one way to do the patent landscape analysis the time point for consideration is the period post the Kyoto protocol, incidentally in this particular area there is no specific classification that are that is available, what is earth friendly mobility technologies is a general term.

So, how do you approach to prepare a patent landscape in an area like this? One needs to look into consideration that where is the sector that you are looking at. So, is it a automobile segment.

So, one needs to look at what are the, what is the sector? Now, even within the sector one needs to look at specifically sub sectors. So, there are different approaches let us say approach one is to look at redesigning the conditions under which combustion takes place, that is one is looking at pre combustion. So, here an understanding of the technology is first important, before we actually embark on the entire patent landscape. So, what are the technologies involved in relation to pre combustion. So, here you see a list of the technology that would fall under this particular category.

(Refer Slide Time: 36:01)

Another approach could be in relation to mitigating the entire aspect of emission from the point of view of post combustion technologies. How well to prevent emissions, how well to capture back in form of saving energy? So, those fall into these set of technologies.

So, here at this stage it is important to now look at what is the specific focus. If, you are looking at approach one or you are looking at approach 2 and in this particular case one can also look at a specific option of looking at only certain technologies, for instance you are looking at only a 4 stroke. In which case, so this is the way in which you can actually look at specific focus in relation to a given technology area.

(Refer Slide Time: 37:07)
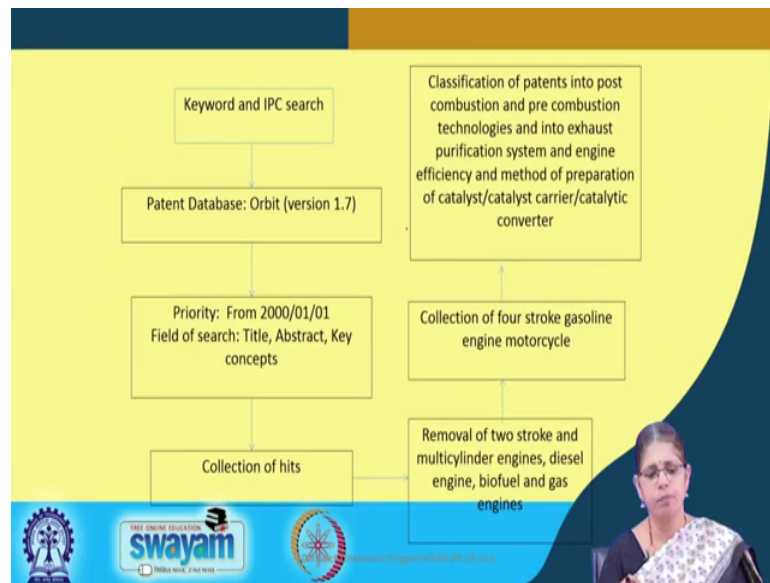
Once you are ready with this you can actually go for a keyword and IPC based search choose this relevant database, identify from when you want to look at the patent data set the time period, collect the hits screen classified order the data, then look at the analysis of the data. So, from the basic technology in terms of the general area to the specific area to a patent data set which represents the type of innovations that you are looking at in a given invention area. And, those patents are then classified in terms of the technology domain and the various aspects of looking at bibliographic information as well.

(Refer Slide Time: 38:11)
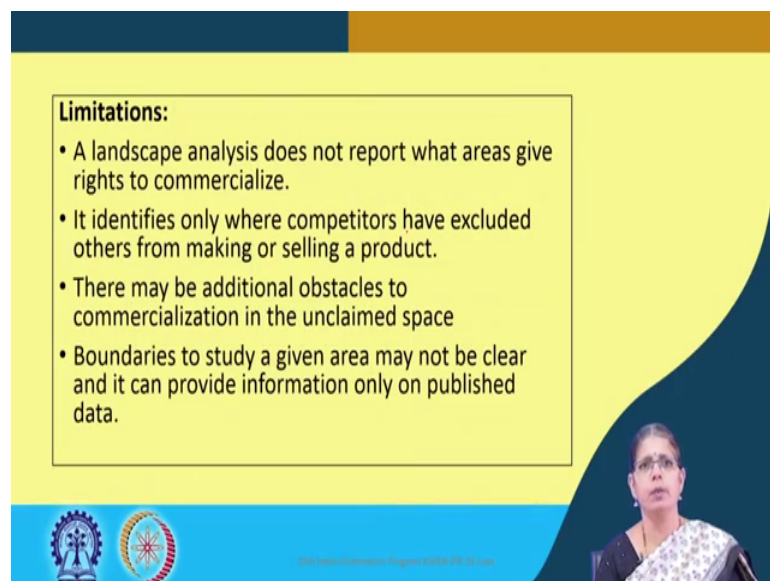
While, patent landscapes provides great opportunity to do all these things there are certain limitations of the patent landscapes. Because, patent landscapes provide you an overall idea, they do not give you an idea on what are the patent rights that can be taken such decisions come post the patent landscape. So, an overall idea is what a patent landscape can give you.

Further they could be other additional obstacles in commercialization of the, but of the product, when you identify gaps in relation to a particular area; it is not always that you can operate, because there are other laws that are applicable. So, though there is an unclaimed space, which the patent landscape provides information on there could be other obstacles, other regulatory hurdles, and other laws, that could be applicable, which may lead to a scenario where a patent landscape becomes only an indicatory information of value to business.

Many a time the boundary of a study may not be clear for instance you are looking at platform technologies, you are looking at let us say, the technologies in relation to business methods, technologies in relation to the electronic segment, for instance you are looking at LTE technologies. It is very very difficult to really set the boundaries, because patent landscapes evolve with respect to the changing objectives that can happen in relation to a particular concern or an issue that a company or an institution is looking at. So, while there are a lot of advantages in relation to a patent landscape, there are these limitations which are there which one needs to keep in mind.

(Refer Slide Time: 40:31)

Patent landscapes are different from other types of patent searches, particularly if one is looking at a patentability search the view of prior art is very narrow whereas, patent landscapes are state of the art searches. So, whole lot of information has to be looked at. And, patent landscapes can be narrow or broad in nature depending on the technology area whereas, if you look at the FTO search or the patentability search, it is pretty much a very focused search from the point of view of a narrow area in which you are looking at information.

(Refer Slide Time: 41:21)

In summary patent landscape represents state of the art search, the data can be classified according to different categories, the value in relation to specific technology clusters can be realized and new developments can be identified based on patent landscapes. The opportunity value of a patent landscape can be for the R and D can be in an area where you are looking at I P transactions and also in an area where you are looking at future R and D.

Certain important aspects of patent landscapes must be kept in mind one from the point of view of understanding the focus of the patent landscape search, second from the point of view of data categorization, adopting the right type of visualization tools, looking at cross referencing of information, screening and deriving conclusions is one important facet of the patent landscape search.

(Refer Slide Time: 42:27)

This figure sums up the entire aspects of patent landscapes and their relevance to the reasons for why patent landscape searches are done from understanding just the distribution of data looking at who are the players in the market, what are the kind of technology collaborations that are available, what type of licensing information an opportunity is available? The value to an organization whether it is a research institution or a company can be best realized by doing what we call the state of the art searches, which the extensive aspect of looking at patents.

(Refer Slide Time: 43:15)

There are a few references for this lecture the WIPO website has a guideline for preparation of patent landscapes, which is a document available, which can be accessed, WIPO also publishes regularly landscape reports which are available for public access in relation to the landscapes are in relation to different technology areas. So, one can actually download from this link that is available. The patent search workbook which has been prepared under this course also provides examples of how the patent landscape has to be conducted and what comes out as a value out of patent landscape.

Thank you.