



IFLA
2005
OSLO

World Library and Information Congress: 71th IFLA General Conference and Council

"Libraries -A voyage of discovery"

August 14th - 18th 2005, Oslo, Norway

Conference Programme: <http://www.ifla.org/IV/ifla71/Programme.htm>

août 9, 2005

Numéro de code: 094-F

Réunion: 107 Généalogie et histoire locale avec les bibliothèques de géographie et les cartotheques (partie I)

Présentation des cartes et de toute autre information spatiale sur internet

Peter Korsgaard

Kort et Matrikelstyrelsen (Service national du cadastre)
Copenhague, Danemark

*Traduction : Samira Cuny
(Titulaire Maitrise des Sciences de l'information et de
la documentation, Université Paris I, France)*

Le Danemark– de nombreux registres numériques décentralisés

Etant un petit pays, le Danemark a une longue tradition pour les registres centralisés ou au moins pour les registres conservés par des établissements publics décentralisés. La plupart de ceux-ci sont maintenant numérisés ou en voie de l'être au nom de l'e-gouvernement. Un des premiers à l'avoir été est le registre central des personnes (CPR) combinant le nom d'une personne avec sa date de naissance et son adresse. Maintenant un grand nombre ont été ajoutés. La plupart des registres ne sont pas accessibles sans permission spéciale, mais dans mon domaine - propriété foncière et utilisation du sol- presque toute l'information est accessible par tous. En partie en raison du souhait de transparence, en partie en raison des grands avantages administratifs, une grande partie des informations est maintenant disponible sur internet. Pour une parcelle de terrain vous pouvez trouver le nom du propriétaire, la

valeur estimée du terrain et des bâtiments, la carte cadastrale, le nombre et l'usage des bâtiments sur le terrain, savoir si une pollution a été enregistrée là et à court terme aussi les hypothèques et autres responsabilités. Si vous avez une autorisation spéciale, vous pouvez voir plus d'information.

Ces registres ne sont pas situés en un lieu unique, ils sont dispersés et mis à jour par plusieurs établissements. La grande tâche est alors de voir les registres communiquer les uns avec les autres de telle manière que l'utilisateur n'ait pas à se déplacer d'un établissement à l'autre. Le Service national du Cadastre joue le rôle principal dans le travail consistant à faire fonctionner ensemble les données spatiales. En fait, c'est notre raison principale d'exister, car nous ne faisons plus de cartes imprimées (juste quelques cartes militaires), juste les données utilisées en cartographie.

Les archives de cartes du Service national du Cadastre – un lot mixte

Cela n'a pas toujours été ainsi. Nous faisons des cartes navales depuis environ 1650, des cartes topographiques depuis 1757 et des cartes cadastrales depuis 1806. De plus, au moment de la réalisation de cartes cadastrales, nous avons rassemblé environ 1800 cartes destinées à des usages économiques, bon nombre d'entre elles montrant les villages à l'époque des fortifications. Dans nos archives nous avons jusqu'en 2004– quand nous avons commencé à envoyer la partie la plus ancienne aux archives nationales– environ 2.000 cartes navales, 70.000 cartes cadastrales (et autres cartes du même genre) et approximativement 30.000 cartes topographiques. À ceci on peut ajouter une collection de photos aériennes prises depuis 1945. Il y a plusieurs centaines de milliers de ces dernières. La plupart du matériel concerne le Danemark, mais les cartes navales incluent également la Baltique et les cartes topographiques incluent les Iles Féroé et le Groenland ainsi que les anciens duchés danois du Slesvig-Holstein.

De loin la majeure partie– disons 90%– de celles-ci sont uniques. Ce sont des dessins, pas des copies. Le seul endroit où ils peuvent être consultés c'est dans notre salle de lecture à Copenhague. Ce n'est pas pratique pour beaucoup de gens, et la détérioration et les déchirures sur ce matériel ne sont pas du goût de l'archiviste. Ainsi pendant des années les archives avaient l'idée de faire une salle de lecture numérique.

Nous sommes au milieu du travail– aucune solution finale

Ce qui suit est une description de certaines des réflexions et des expériences acquises durant le travail de préparation et la réalisation d'autres projets à ce moment là. Je dois souligner que nous ne sommes pas à mi chemin, mais les autres projets ont déjà donné des raisons de reconsidérer le projet. Cela prend également du temps de scanner les cartes (nous avons numérisés environ 10.000 cartes et au cours d'un assez grand projet pour l'e-gouvernement ayant pour objet de numériser la plupart des cartes cadastrales, approximativement 40.000 cartes, qui seront numérisées d'ici un an).

Un ancien projet Internet

Il y a quelques années, lors d'un projet sous le patronage de la Communauté européenne nous avons collaboré avec une organisation suédoise soeur (Lantmäteriet) et l'université de Greifswald afin de mettre un certain nombre de cartes sur internet (notre part était de 3.000 cartes). Le critère de recherche ici, cependant, était la paroisse, et aussi toutes les cartes liées à un village ou à une paroisse. Ainsi elles pouvaient facilement être gardées sur une base alphanumérique (comté-paroisse-village). La page est maintenant fermée, mais ce fut un grand succès, et les autres associés ont à ce jour incluent

leur partie dans leur page d'accueil. A bien des égards cela a été enrichissant, non seulement au niveau des flux et techniques de production, mais aussi en particulier du fait que les utilisateurs posaient de nombreuses questions concernant les cartes. Une chose a heureusement été prise en compte : on s'attendait à ce que des généalogistes utilisent la partie danoise, il était nécessaire d'avoir des options de recherche du Danemark tel qu'il est maintenant aussi bien que la situation de la paroisse à l'époque des cartes. Beaucoup de généalogistes ont localisés leurs ancêtres grâce aux recensements ou aux registres paroissiaux– cela reflétait l'ancienne situation. Il était alors important d'inclure les noms historiques des paroisses afin de s'assurer qu'ils pourraient trouver les bonnes cartes, même si la paroisse n'existait plus.

Les possibilités d'internet pour les géomètres

D'autres types d'expériences ont été obtenus par la numérisation, l'indexation et le géocodage des relevés. Au Danemark, l'activité de géomètre est une profession libérale, mais tous les changements doivent être approuvés par le Service national du Cadastre, et tous les relevés depuis 1906 peuvent être trouvés ici. Pour une région du Danemark, tous les relevés peuvent être consultés sur internet, et pour le Danemark dans l'ensemble tout relevé approuvé depuis peut aussi être consulté ici maintenant. Un projet de numérisation de tous les relevés remontant au moins jusqu'en 1985 (plusieurs centaines de milliers), d'indexation et de géocodage dans le but de les mettre sur l'Internet vient d'être approuvé. Je reviendrai plus tard sur la définition du géocodage, mais elle reflète le problème que ni le numéro du cas ni du lot de l'identificateur n'est suffisant, car l'identificateur du lot peut changer au cours du temps. Cette solution pourrait également être une solution pour le problème du changement des frontières juridictionnelles au cours du temps (par exemple les paroisses) et les problèmes de la sorte.

Nous ne sommes pas seuls

Avant que je ne passe à l'aspect technique, je soulignerai qu'au Danemark d'autres établissements sont intéressés par le travail des références spatiales. Les Archives nationales et la Bibliothèque royale possèdent d'importantes cartothèques. Aux Archives nationales, le besoin de recherche des juridictions actuelles et historiques– les comtés, les évêchés, et ainsi de suite sera ajouté. Comme le matériel des archives est fait et enregistré sur la base des juridictions, elles sont la clef de recherche la plus importante. D'autres établissements sont également importants ; par exemple l'Institut de recherche des noms (Name Research Institute) de l'université de Copenhague a fait un grand registre des noms de lieux médiévaux et de ressources et des transactions (la plupart du temps des actes) reliées à ceux-ci.

Les collections mentionnées ci-dessus ajoutent aux complexités, mais d'autre part il est important d'inclure ces derniers dans les réflexions sur les solutions. L'utilisateur moyen au Danemark aurait des difficultés à comprendre pourquoi les mêmes cartes sont traitées d'une manière au Cadastre national, d'une autre aux archives nationales et encore d'une troisième manière à la Bibliothèque royale. Il ne comprendrait guère pourquoi ceci devrait affecter la disponibilité de la carte ou le choix des métadonnées concernant la carte. Et les volumes combinés des cartes et des informations s'ajouteraient au bénéfice de l'utilisateur.

Le centre d'attention principal - l'utilisateur

Avant d'élaborer un système, il est important de se tourner vers l'utilisateur final: Que sait-il, que veut-il vraiment ? Et à nouveau il peut y avoir de nombreuses réponses. Nous en conviendrons tous, il y a

besoin d'un moteur de recherche des cartes et de toute autre information spatiale du/des établissement(s). En outre nous pouvons convenir que l'utilisateur a besoin de métadonnées pour les cartes et l'information, bien que nous puissions être en désaccord sur l'importance des méta-données. Mais un grand nombre d'utilisateurs pourrait mieux être servi si nous avions quelques services, par exemple des cartes sélectionnées avec une plus grande fonctionnalité. Sur la page d'accueil du Service national du Cadastre du Danemark on trouve des nouvelles et d'anciennes cartes rendues en continuité, à laquelle est associée la récente base de données d'adresse du Danemark. Ainsi vous pouvez mettre une adresse d'aujourd'hui et voir à quoi ressemble l'endroit aujourd'hui, à quoi il ressemblait dans les années 30 ou les années 1890. Ou - à un autre endroit de la page d'accueil- vous pouvez à nouveau mettre une adresse et obtenir une carte de celle-ci avec un repère sur l'endroit exact (voir fig. 1) Sur la carte il y a un "fil" que vous pouvez tirer vers la gauche ou la droite. A droite c'est une carte des années 1890, à gauche c'est une carte d'aujourd'hui. Ainsi vous pouvez faire défiler dans les deux sens entre la partie récente et la partie historique. Le dernier "service" a été l'une des parties les plus populaires de la page d'accueil.



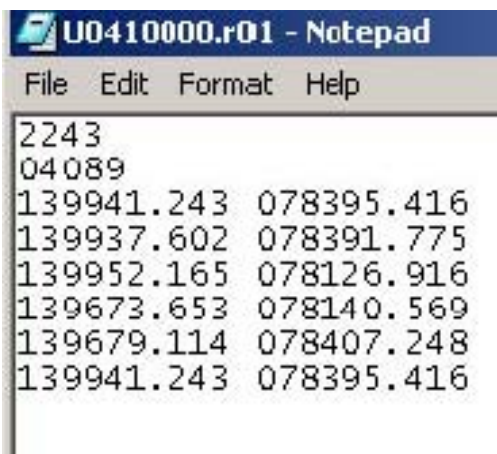
Fig. 1. Page d'accueil du Service national du Cadastre (www.kms.dk). Malheureusement cette partie est uniquement en danois. La recherche est faite soit en choisissant un nom de lieu, soit en écrivant le nom d'une rue, en choisissant la municipalité et le numéro de l'adresse. Le repère indique l'adresse que vous avez choisie. À gauche de la bande de roulement se trouve une carte récente, à droite une carte de 1898. En tirant le fil vous pouvez faire défiler dans les deux sens.

Techniques - Géocodage

En passant à l'aspect technique, je voudrais commencer par souligner les définitions que j'utilise dans cet article : indexer une carte c'est donner une description alphanumérique afin de la trouver dans les bases de données ordinaires, par exemple les noms de paroisse, les noms de cartes. Le géocodage c'est donner à la carte une référence spatiale sous forme de coordonnées. Ces coordonnées sont reliées à un système fixé à l'avance. L'indication de coordonnées ou d'un secteur de points indiquant des coordonnées peut ainsi servir de base à une recherche de cartes appropriées.

Le géocodage d'une image de fichier raster d'une carte peut être obtenu de différentes manières. Une manière est de préciser les coordonnées de quelques pixels dans l'image. De cette façon les autres pixels obtiennent aussi automatiquement des coordonnées. Habituellement cela est fait en associant les points d'une carte moderne géocodée avec les points de la carte plus ancienne. Dans certains cas, par exemple les cartes topographiques plus récentes, vous pouvez connaître l'endroit exact des coins de la carte et de cette façon obtenir une telle exactitude qu'après le processus vous pouvez avoir des cartes en continuité, vous pouvez avoir des programmes d'affichage manipulant les fichiers séparés comme s'ils constituaient une seule carte.

Au cours de ces types de géocodage vous pouvez également avoir une matrice de caractères, qui est un petit fichier texte indiquant un polygone autour de la carte.



```
U0410000.r01 - Notepad
File Edit Format Help
2243
04089
139941.243 078395.416
139937.602 078391.775
139952.165 078126.916
139673.653 078140.569
139679.114 078407.248
139941.243 078395.416
```

Fig. 2. Une matrice de caractères indiquant un pentagramme. Les deux premières rangées concernent l'identifiant ; les colonnes suivantes sont les coordonnées des points du pentagramme. Il y a six entrées, mais les premiers et les derniers sont identiques.

Voici une autre manière de géocoder. Il y a le pour et le contre concernant les deux manières.

Le géocodage par le codage de pixels

- prend du temps, particulièrement si cela doit être très précis
- peut être difficile avec des cartes plus anciennes étant donné

- qu'elles ont été réduites – pas toujours de manière proportionnelle
- qu'il peut être difficile de trouver les points correspondants – j'ai même vu des bornes routières qui ont été déplacées au cours du temps
- peut également créer une matrice de caractères dans un fichier texte séparé
- donne la possibilité d'utiliser le fichier raster géocodé si vous avez le logiciel capable de traiter le géocodage
- donne la possibilité de stocker le géocodage ainsi que des métadonnées dans le fichier raster, permettant ainsi de copier l'information dans le fichier pour être utilisé en dehors de l'établissement (voir la fig. 5)
- augmente peu la taille du fichier, mais il pourrait toutefois être énorme tout comme les fichiers raster tendent à l'être
- à moins que vous fassiez quelques coupures, le géocodage des pixels concernera le fichier raster entier, donc le cadre blanc de la carte sera également traité en tant qu'élément de la carte. De cette façon vous obtiendrez plus « d'occurrences » que nécessaire lors de la recherche.

Le géocodage par les matrices de caractères

- est plus simple et dans certains cas peut être automatisé
- le fichier créé est très petit, habituellement < 1 KB
- si deux cartes couvrent la même zone, le fichier texte contenant la matrice de caractères créée pour la première la carte peut être réutilisé pour la deuxième carte
- ne change pas le fichier raster, ainsi c'est toujours seulement un fichier raster sans géocodage ou métadonnées
- des matrices de caractères peuvent être créées pour des cartes modernes vectorielles. Les cartes vectorielles ne sont pas constituées de données maillées, mais sont basées sur des bases de données et se composent de points, de lignes et des surfaces sous formes de nombreux thèmes, parmi d'autres souvent des juridictions.

En gros, les matrices de caractères sont bonnes pour la recherche sur internet, alors que si vous voulez faire des services, l'autre type de géocodage est plus approprié.

Dans cet article il y a un exemple de matrice de caractères (voir la fig. 2). C'est très simple - d'abord quelques codes donnant l'identifiant et le genre de groupe. Ensuite il y a quelques ensembles de coordonnées, et vous verrez que les premiers et les derniers couples sont identiques. Les coordonnées sont les coins d'un polygone.

Lors de l'apparition sur l'écran de l'établissement, vous pouvez choisir si ces matrices de caractères seront mises au dessus d'une carte ordinaire ou être cachées sous elle. Sur internet, cependant, c'est la décision du webmestre. Normalement, le nombre de polygones sera trop grand et produira seulement du « bruit » sur l'écran. Dans certains cas, cependant, le nombre d' « occurrences » dans une zone est si petit qu'il vaut mieux voir s'il n'y en a pas du tout. En outre, comme dans le cas des relevés (voir fig. 3), le géomètre peut souvent dire quel est le bon document juste en examinant les formes. Concernant les relevés, vous cliquez simplement sur le polygone, et l'image maillée du document apparaît. S'il y avait beaucoup de cartes couvrant un certain secteur, il serait préférable d'avoir un menu avec toutes les cartes pertinentes décrites par des métadonnées pour ensuite faire votre sélection à partir du menu. C'est ce qui se produit dans un cas mentionné par la suite (voir fig. 4),

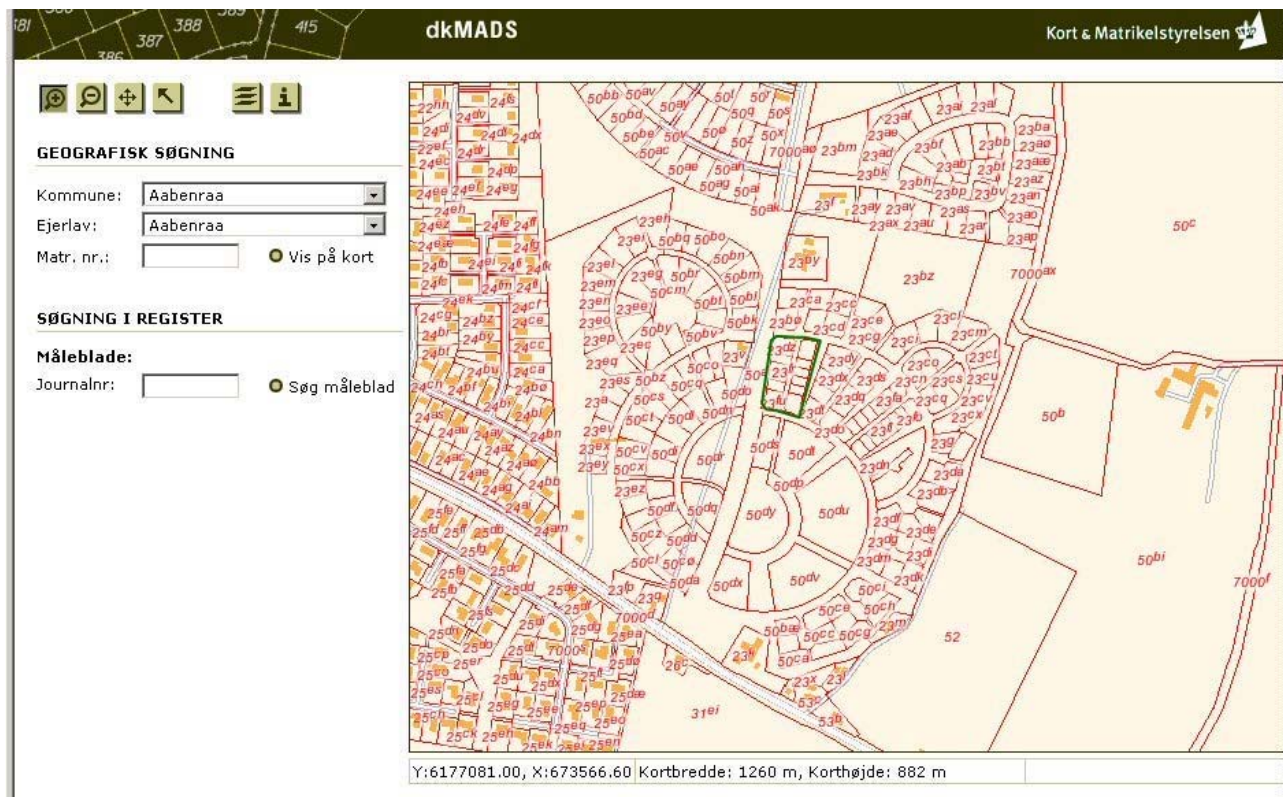


Fig. 3. Impression d'écran du Service national du Cadastre – recherche de relevés récents. A gauche c'est la manière habituelle – si vous êtes géomètre – par numéro de cas, l'identifiant de la parcelle, etc. A droite c'est une carte. Il y a plusieurs couches de cartes, parce que vous pouvez avoir besoin de commencer par la totalité du Danemark et puis faire un zoom sur cette couche. C'est la carte cadastrale récente avec les bâtiments ajoutés pour une meilleure orientation. Vous avez les frontières des lots et des numéros cadastraux.

Ceci est une carte ordinaire maillée qui a été géocodée sous forme de pixels. Au milieu de la carte on a une boîte noire, une matrice de caractères indiquant un lien au relevé. En cliquant sur la flèche sur le côté supérieur gauche et ensuite sur la matrice de caractères vous obtiendrez le relevé numérisé dans une fenêtre contextuelle. Comme il n'y a que quelques relevés dans la base de données, il est nécessaire que vous puissiez immédiatement voir s'il y a une occurrence ou non.

La transformation de vieilles cartes en cartes vectorielles pourrait être une autre solution.

Une carte vectorielle a été utilisée par un étudiant pour une carte historique vectorielle de juridictions (villages, paroisses et comtés), permettant alors de créer des matrices de caractères pour des juridictions historiques. L'utilisation d'une carte vectorielle moderne avec les nouvelles frontières et d'une carte maillée avec les anciennes frontières et les éléments de 1820 l'a fait. La dernière était géocodée et employée comme papier peint sous les nouvelles frontières. Celles-ci ont alors été ajustées à la forme des vieilles frontières et une carte vectorielle des juridictions de 1820 a ainsi été créée. Chaque petit secteur de cette carte peut alors être transformé en matrice de caractères et être employé comme élément de géocodage d'autres cartes. Ou bien, dans le cas de l'index des noms de lieux de l'Institut des noms à l'université de Copenhague : le registre est établi autour des paroisses telles qu'elles étaient

autour de 1950 – alors tous les noms de lieux appartenant à une paroisse donnée pourraient être liés à la matrice de caractères pour cette paroisse aujourd'hui, donnant l'occasion à la fois de voir quels noms de lieu appartiennent à une paroisse donnée et à la fois d'une autre manière montrer le lieu approximatif du nom de lieu.

La recherche

Si vous faites votre recherche par le moyen de coordonnées, vous pouvez avoir plusieurs options selon le logiciel utilisé. Un outil commun est un point – qui est habituel pour la recherche d'adresses. Dans les systèmes locaux ceci pourrait également être étendu à un cercle avec un rayon défini par l'utilisateur. Peut-être le plus commun sur internet est un rectangle, les côtés allant au nord-sud et est-ouest. Il n'est peut être pas toujours le plus efficace, mais il est facile à manipuler pour l'utilisateur, particulièrement s'il peut lui-même dessiner le rectangle avec le curseur.

Une autre manière de limiter le nombre d'occurrences c'est naturellement par des critères tels que des thèmes ou le temps. Les problèmes de cette sorte sont si bien connus que je m'abstiendrai de les traiter. Une source d'inspiration peut être <http://eu-geoportal.jrc.it/gos>. C'est la page d'accueil d'INSPIRE (infrastructure pour l'information spatiale en Europe), une initiative lancée par la Commission européenne visant entre autres à déclencher une structure de données spatiale européenne (voir fig. 4).

The screenshot shows the INSPIRE European Geo-Portal search interface. It is organized into three main columns: 'Where?', 'What?', and 'When?'.
 - **Where?:** Features a map of Europe with country labels (e.g., NORWAY, SWEDEN, GERMANY, FRANCE, ITALY, GREECE, TURKEY). It includes a search box 'FIND A PLACE ..', a 'locate your area here' section with map navigation icons, and a checkbox 'Data must fall completely inside area'. Buttons for 'LAUNCH MAP VIEWER TO DRAW BOX' and 'GET BOX ALREADY DRAWN ON VIEWER' are present.
 - **What?:** Contains search criteria: 'Keywords' (text input), 'Match' (radio buttons for 'Exact phrase', 'Any word', 'All words'), 'Data Category' (dropdown menu), 'Provider Level' (dropdown menu), 'Scale/Coverage' (dropdown menu), and 'Type/Format' (dropdown menu).
 - **When?:** Includes 'Anytime' (selected radio button), 'Data for time period (yyyyymmdd)' with 'From:' and 'To:' input fields, and 'Data updated recently (yyyyymmdd)' with an 'After:' input field.
 At the bottom, there is a 'Sort results by:' dropdown menu set to 'Content type/format', and 'SEARCH' and 'RESET' buttons.

Fig. 4. La page de recherche d'INSPIRE. Comme elle est ouverte pour de nombreux types cartes, il est important de rétrécir le champ. Vers la gauche il y a la possibilité de limiter le secteur de recherche – vous pouvez zoomer sur la carte.

L'importance des standards

Les travaux de coordination d'INSPIRE, et cet aspect doivent être bien soulignés. Nous avons dans le domaine des métadonnées des normes telles que le Dublin Core et l'ISO 19.115 : L'information géographique - métadonnées. Au Service national du Cadastre nous utilisons l'ISO et gardons à l'esprit la possibilité d'utiliser le Dublin Core. Ces normes, cependant, ne couvrent pas bien de nombreuses cartes du Service national du Cadastre, car il a été difficile d'adapter les cartes cadastrales à ces normes. Dans le cas des photographies aériennes il y a besoin de plus de métadonnées, par exemple des données au sujet de l'appareil photo. En plus des données de la carte elle-même il y a un besoin de données au sujet du scannage et du traitement des fichiers du downsampling au type de géocodage. Ces données sont nécessaires si le public va utiliser le fichier. Enfin il est nécessaire de suivre les normes des systèmes géographiques de référence, et naturellement le Service national du Cadastre est à la tête dans le domaine danois et doit suivre les accords internationaux. Nous utilisons le standard européen UTM/ETRS89 (au Danemark la plupart du temps appelé UTM/EUREF89).

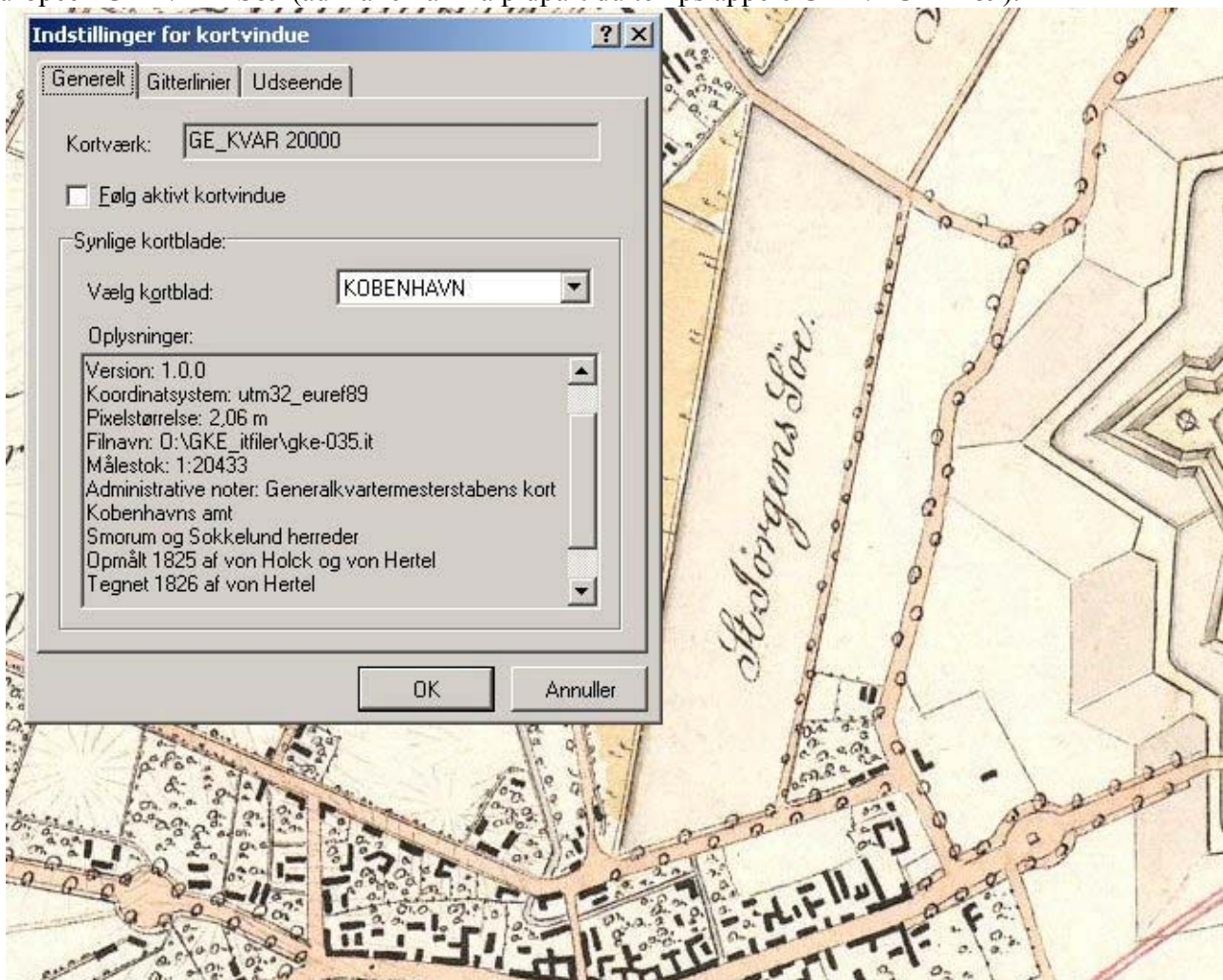


Fig. 5. Un exemple d'information stockée dans un fichier raster geo-codé. Avec ce logiciel, vous faites juste un clic-droit sur la carte, vous choisissez les « propriétés » et alors vous obtenez une fenêtre contextuelle indiquant certaines des métadonnées de la carte. Cependant, cela prend un certain temps de les écrire tous dans le fichier, vous ne devez donc pas être trop difficile. Mais entre autres il donne le nom des séries de carte, le nom de cette carte, les années de production, et l'échelle. Il indique également dans quel système de référence il a été géocodé. Autrement il ne pourrait pas être de grande valeur pour l'utilisateur.

Conclusion

Ceci constitue quelques uns des nombreux détails. D'une certaine façon il reflète la situation de nos archives en ce moment. Il me reflète moi, historien/archiviste, qui doit essayer d'apprendre encore plus au sujet des techniques numériques. Mais il reflète également le fait qu'il y a de nombreux chemins à suivre, de nombreuses décisions à prendre, si vous voulez aller dans cette direction. Je suis tout à fait certain que c'est la bonne direction, car nous avons beaucoup de cartes qui ne peuvent pas être trouvées par titre – et toutes les photos aériennes. Il n'y a pas longtemps je discutais avec un expert en numérisation de mon département sur la manière de faire ceci; il a juste dit, note une liste de besoins et ils peuvent être satisfaits. Naturellement les coûts pour mener à bien le travail pourraient être énormes, mais il est vrai dans une certaine mesure qu'il vaut mieux avoir les idées et ne pas être limité par les aspects techniques.

Je proposerais, si vous pensez poursuivre, que vous devriez au moins

- décrire les utilisateurs finaux - ce qu'ils veulent, ce qu'ils sont capables de faire
- décrire le matériel - quel sont les problèmes
- penser également aux services, pas seulement à un moteur de recherche
- faire que les utilisateurs aient seulement besoin de regarder, ou ils voudraient également obtenir les fichiers
- si vous avez beaucoup de matériel, pensez à le faire en modules à développer. – mais d'une manière telle que le tout s'assemble à la fin
- pensez aux normes des méta-données et des systèmes de référence - vous pourriez devoir coopérer avec d'autres, et de toutes façons le logiciel pour transformer des données d'un système de référence à l'autre sera plus facilement à portée de la main
- penser à la réutilisation de l'information, par exemple des matrices de caractères ou des données des cartes vectorielles
- se renseigner sur les pratiques d'autres établissements

J'ai beaucoup de chance d'être dans un établissement qui est techniquement très avancé. Beaucoup d'établissements culturels patrimoniaux ne sont peut-être pas aussi bien dotés dans ce domaine, ainsi la coopération des établissements peut être le vrai mot clé.