



Geodaten Technisches Basiswissen

erarbeitet durch die Koordinierungsstelle GDI-NI beim Landesamt für
Geoinformation und Landentwicklung Niedersachsen (LGLN)



Landesamt für Geoinformation
und Landentwicklung Niedersachsen
Landesvermessung und Geobasisinformation



Niedersachsen

Geodaten – Technisches Basiswissen

Unterschiede zwischen Raster- und Vektordaten	2
Vektordaten im Detail	4
Layout bei Raster- und Vektordaten	4
Multi-Part-Features und ihre Besonderheiten	5
Qualitative Beurteilung von Raster- und Vektordaten	6

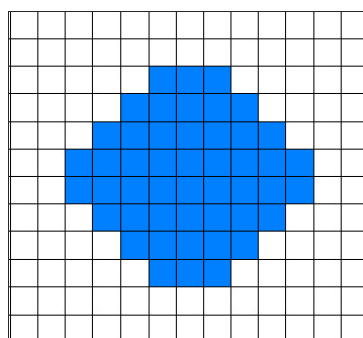
Geodaten sind Daten mit Raumbezug. Sie sind rechnerlesbare Informationen über Geländeform, Gegebenheiten und Gegenständen mit räumlichem Bezug zur Erdoberfläche. Geodaten bestehen aus Geometrie- und zugehörigen Sachdaten. Sie bilden die Grundlage für die Erstellung von Karten und Plänen. Technisch lassen sich Geodaten, wie auch andere Bilddaten in Raster- und Vektordaten unterteilen. Zudem gibt es Geodaten, die in alphanumerischer Form (Datenbank, Tabelle, Text usw.) vorliegen. Alphanumerische Geodaten können in der Regel in Raster- oder Vektordaten umgewandelt werden.

Unterschiede zwischen Raster- und Vektordaten

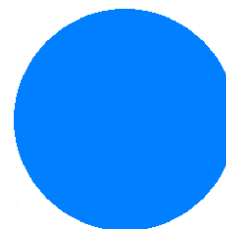
Durch ihre einfache Georeferenzierbarkeit werden Rasterdaten derzeit in der Regel als direkte Nachfolger ehemals analoger Karten und Pläne verwendet. Auch dort, wo es auf eine konkrete Ausgestaltung einer Karte oder eines Plans ankommt, werden Rasterdaten für das Drucken von Karten benutzt.

Vektordaten werden üblicherweise dort verwendet, wo Häufigkeitsverteilungen, Weglängen, Flächengrößen oder Grenzen ermittelt werden müssen. Da sie auf einfache Art und Weise mit Attributtabelle verknüpft werden können, bietet sich ihr Einsatz an, um Datenbankinhalte zu visualisieren. Zwar bieten sich viele Möglichkeiten der Ausgestaltung des Layouts von Vektordaten, diese sind aber für die Nutzer derzeit eher von niedriger Bedeutung.

Punkt (Rasterdaten)



Punkt (Vektordaten)



Rasterdaten bestehen aus in Reihen und Zeilen angeordneten Farbsegmenten. Diese Farbsegmente, die Pixel genannt werden, haben eine feste Kantenlänge. Wegen dieser festen Kantenlänge der Farbpunkte wirken Rasterbilder „pixelig“, wenn sie zu stark vergrößert werden. Am deutlichsten wird dieser Effekt, wenn ein Rasterbild in einer Zoomstufe angeschaut wird, die diejenige der 1:1 Darstellung überschreitet, womit

„Treppenstufen“ im Bild sichtbar werden. Rasterdaten sollten prinzipiell immer nur bis zu einer Darstellung von 1:1 angezeigt werden.

Rasterdaten können aus Gründen der Performance nur bis zu einer bestimmten Größe erzeugt oder genutzt werden, daher werden sie häufig gekachelt. Die einzelnen Kacheln können dann z. B. unter Verwendung eines Geoinformationssystems als nahtlos erscheinende Bilddaten angezeigt werden.

Vektordaten bestehen aus Vektoren, die sich abermals in Punktinformationen, Linieninformationen und Polygoninformationen untergliedern lassen. Dabei besteht ein Punkt aus mindestens einer Koordinate, die eine Angabe zu x und y umfasst. Optional kann eine Höhenangabe z zusätzlich aufgeführt werden. Eine Linie besteht durch die Verbindung von zwei Koordinaten. Ein Polygon besteht aus einer Reihe von Koordinaten, von denen entweder die letzte Koordinate gleich der ersten ist oder aber eine zusätzliche Information darüber vorliegt, bei welcher Koordinate die Linie, die das Polygon abbildet geschlossen werden muss.



Punkt-, Linien- und Polygonelemente in Vektordaten werden in der Regel als Objekte oder mit dem englischen Begriff Features bezeichnet.

Zusammenfassend besitzen Vektordaten einen geringeren Speicherbedarf als Rasterdaten. Es können Attributinformationen hinzugefügt werden, wodurch eine Sachdatenabfrage möglich wird. Auf die dargestellten Objekte (Punkt, Linie, Polygon) kann gezielt zugegriffen werden. Die Objekte besitzen eine hohe geometrische Genauigkeit. Der Nachteil von Vektordaten ist ihre komplexe Datenstruktur. Somit ist auch die Erfassung von Geometrie und Topologie (Bezug der Lage der Objekte zueinander) aufwändig.

Die Vorteile der Rasterdaten liegen vor allem in ihrer einfachen Datenstruktur. Die Daten können innerhalb kurzer Zeit erfasst werden, indem z.B. analoge Pläne oder Karten eingescannt werden. Rasterdaten sind für die Darstellung kontinuierlich verteilter Beobachtungen wie z.B. eine Temperaturverteilung gut geeignet. Die häufigste Verbreitung finden Rasterdaten in der Geodatenwelt jedoch in Form von Luft- und Satellitenbildern. Der Nachteil von Rasterdaten liegt in den verhältnismäßig ungenauen Grenzen der dargestellten Sachverhalte. Diese relative Ungenauigkeit ist direkt abhängig von der bei der Datenerstellung gewählten Bodenauflösung. Klassifizierungen auf der Basis von Rasterdaten beruhen auf einer Auswertung der Farbwerte der einzelnen Pixel. Nach der Datenerhebung kann die Qualität von Rasterdaten kaum verbessert werden. Auch ist der Prozess, um Rasterdaten in ein anderes Koordinatensystem zu transformieren im Vergleich zu Vektordaten verhältnismäßig aufwändig.

Vektordaten im Detail

Grundsätzlich werden Koordinaten bei Linien und Polygonen als Stützpunkte oder vertices bezeichnet. Stützpunkte, die zugleich eine Sonderfunktion übernehmen, wie das Schließen eines Polygons oder das Abzweigen mehrerer Linien werden als Knoten oder nodes bezeichnet. Einzelne Liniensegmente heißen in der Regel arcs.

Ein Punkt ist damit genau genommen lediglich ein Stützpunkt. Aus mindestens zwei Stützpunkten entsteht eine Linie. Ein Polygon wird aus einer Linie gebildet, die zusätzlich die Eigenschaft hat, dass sie an einem bestimmten Stützpunkt, dem Knoten oder node, geschlossen wurde. Durch die Mitgabe der Topologie-Eigenschaft wird aus der geschlossenen Linie eine Fläche.

Jede Fläche kann wieder in eine Linie aufgesplittet werden. Ein Polygon besitzt einen Schwerpunkt, was genutzt werden kann, um aus einem Polygon Punkte zu extrahieren. Dieses Verfahren ist jedoch bei denjenigen Polygonen fehleranfällig, deren Schwerpunkt außerhalb des eigentlichen Polygons liegt, da es dadurch zu Verfälschungen hinsichtlich der Lage kommen kann. Polygone und Linien können ebenfalls in Punkte aufgelöst werden. Die Stützpunkte bzw. vertices und bei Polygonen zusätzlich die Knoten bzw. nodes werden damit zu vollwertigen Punkten. Die Polygoneigenschaft in der Topologie, sowie die Liniensegmente bzw. arcs entfallen. Punkte zu Linien oder Polygonen zu konvertieren ist grundsätzlich schwierig, da eine nicht enthaltene Topologie-Information zusätzlich erzeugt werden muss, was in der Regel nur durch User-Interaktion erreicht werden kann.

Im Vektorformat wird eine Linie bzw. ein arc durch zwei Koordinatenpaare (mindestens Angabe von x und y) abgebildet. Vektordaten können daher prinzipiell in jeder Zoomstufe einwandfrei betrachtet werden.

Einschränkungen ergeben sich in der Praxis häufig durch die Anzahl der Stützpunkt für Linien oder Polygone, die einen Vektordatensatz aus der Ferne betrachtet schnell „klumpig“ erscheinen lassen. Beim starken Zoomen auf Vektordaten kann jedoch auch eine „Übergenaugigkeit“ erzielt werden. Obwohl z. B. eine Linie immer noch in voller Schärfe dargestellt wird und damit eine hohe Genauigkeit vorgetäuscht wird, ist bereits ein Maßstab der Darstellung erreicht, der denjenigen der ursprünglichen Datenerstellung überschreitet. Um Vektorinformationen zu nutzen, muss damit eine Metainformation bekannt sein, die den ursprünglichen Digitalisiermaßstab nennt. Nur bis zu diesem Maßstab darf eine Vektorinformation aus fachlicher Sicht zu Zwecken der Interpretation dargestellt werden.

Layout bei Raster- und Vektordaten

Während beim Rasterbild die einzelnen Farbpunkte abgespeichert werden, werden bei einem Vektorbild die einzelnen Stützpunkte in Form von Koordinatenpaaren abgespeichert. Bei Rasterdaten handelt es sich um Bilder von Geodaten, die bereits ein bestimmtes Layout besitzen. Bei Vektordaten handelt es sich um Geometrie- und Attributinformationen zu denen bei Bedarf ein bestimmtes Layout hinzugefügt werden kann. Rasterdaten sind damit nur sehr eingeschränkt hinsichtlich ihrer Darstellung veränderbar. So ist beim



Vorliegen einheitlicher Farbpaletten für alle Kacheln des Rasterdatensatzes ein Farbtasch (z. B. rot statt gelb) möglich. Vektordaten können hingegen frei gestaltet werden.

Bei Vektordaten, die in Form von Punkten, Linien oder Polygonen dargestellt werden, kann die Form und Farbe der Objekte je nach ausgewähltem Maßstab gesondert festgelegt werden. Hierzu wird eine Gestaltungsvorschrift für die Vektordaten für einen bestimmten Maßstabsbereich in der Ansicht angelegt. Vektordaten können damit, anders als Rasterdaten, beliebig gestaltet werden. Auch ist es möglich, mehrere Gestaltungsvorschriften für Vektordaten zu hinterlegen, so dass der eine Nutzer, das eine Layout angezeigt bekommt und ein anderer Nutzer ein davon abweichendes.

Multi-Part-Features und ihre Besonderheiten

In der Praxis existieren die drei Grundformen Punkt, Linie, Polygon für Vektordaten häufig auch als so genannte Multi-Part-Features hinter denen sich mehrteilige Objekte verbergen. Die Mehrteiligkeit resultiert vor allem aus der Sachinformation, die mit der oder den Geometrien verknüpft ist. Während bei den Grundformen jeweils eine Sachinformation mit einer Geometrie verknüpft ist, haben Multi-Part-Features eine n:1 Beziehung. Das bedeutet, an mehreren Punkten, mehreren Linien oder mehreren Polygonen hängt gleichzeitig nur eine einzige Sachinformation.

Bei Polygonen gibt es zusätzlich weitere Varianten. So kann an mehreren Polygonen eine Sachinformation hängen oder aber ein Polygon kann eine Inselfläche beinhalten, so dass das Polygon selbst streng genommen aus einer Abfolge von Polygonen besteht, die von einander zu subtrahieren sind. Die Inselfläche selbst kann ein weiteres Polygon darstellen, kann aber ebenso eine reine Insel ohne Sachinformation abbilden. Multi-Part-Features verursachen sehr häufig Probleme bei der Konvertierung der Vektordaten von einem Format in ein anderes. Sind bei dieser Konvertierung im Zielformat keine Multi-Part-Features vorgesehen, so muss das Multi-Part-Feature aufgesplittet werden. Dies kann folgendermaßen geschehen.

Jeder Sachinformation zu einem Punkt, einer Linie oder einem Polygon wird eine identische Geometrie zugewiesen. In diesem Falle liegt die Geometrie, die es nur einmal geben sollte, mehrfach vor (Redundanz), was zu falschen Berechnungen z. B. hinsichtlich der Flächengröße führen kann. Ein Informationsverlust liegt jedoch insgesamt nicht vor, so dass der Mangel durch spezielle Algorithmen bei der Weiterverarbeitung stark abgemildert bis ausgeglichen werden kann.

Der Sachinformation zu einem Punkt, einer Linie oder einem Polygon wird eine einzige der verknüpften Geometrien zugewiesen. Dieser Fall ist in der Praxis häufig schwerwiegend, da der Fehler nur über einen direkten Vergleich aller Objekte des konvertierten Datensatzes mit den Objekten des Originalformats eindeutig nachgewiesen werden kann, denn die „überzähligen“ Geometrien sind weggefallen, so dass zwischen der Geometrie und der Sachinformation wieder eine 1:1 Beziehung besteht. Nachgewiesen werden kann der Fehler beispielsweise mittels Größenvergleich der Geometrien, deren Anzahl, Länge oder Fläche hinsichtlich der Originaldaten abweicht. In der Regel kann auch dieser Mangel durch entsprechende Nachbearbeitung behoben werden, jedoch ist dafür ein Zugriff auf die Originaldaten erforderlich, um die



korrekte Lage der Geometrien zu ermitteln, so dass schlussendlich die Sachdateninformation redundant geführt werden kann.

Bei Polygonen mit Inselflächen führt die Konvertierung dazu, dass entweder der Inselfläche die selbe Sachinformation wie dem Polygon zugewiesen wird, oder aber, diejenige Geometrie, die die Inselfläche markiert, entfällt, wodurch sich die Flächengröße des (ehemals) umgebenden Polygons erhöht. Im einfachsten, aber gleichzeitig schwerwiegenden Fall, kann die Geometrie nicht konvertiert werden und fällt weg. Es ist dann im Einzelfall abzuwägen, welche Möglichkeiten für eine Lösung zur Verfügung stehen und ob eine angemessene Lösung herbeigeführt werden kann.

Als Fazit kann festgehalten werden, dass zu einer vollständigen Beschreibung von Geodaten auch eine Information gehören muss, ob in den Geodaten Multi-Part-Features enthalten sind. Werden in den Geodaten Multi-Part-Features verwendet, ist es für den Datennutzer zwingend erforderlich, eine durchgeführte Formatkonvertierung eingehend vorab auf potentielle Fehler hin zu überprüfen. Dies geschieht im Idealfall mittels einer Testdatei, die sämtliche relevante Fälle hinsichtlich der Multi-Part-Features abdeckt. Eine Serviceleistung des Datenanbieters könnte hier das Erstellen einer solchen Testdatei für die Kunden sein, die anhand der Datei überprüfen können, welche konkreten Probleme bei einer Konvertierung der Daten im Zusammenspiel mit ihrer Software auftauchen. Aus Sicht der GDI ist hier speziell der Bereich der WebFeatureServices / Download-Dienste betroffen, die Daten im Austauschformat gml anbieten.

Qualitative Beurteilung von Raster- und Vektordaten

Aus Sicht der vielfältigen Nutzbarkeit sind Vektorinformationen stets höher anzusiedeln als Rasterdaten. Auch ist der Speicherbedarf von Vektordaten deutlich geringer. Hauptunterschied zwischen Raster- und Vektordaten ist die Form der Georeferenzierung, die bei Rasterdaten durch eine mitgelieferte Eckpunkt-Datei mit Koordinaten geschieht, während Vektordaten ihre Georeferenzierung aus den Stützpunkten beziehen. Nebeneffekt dieses Unterschiedes ist, dass zwar Vektordaten in einer nahezu beliebigen Projektion dargestellt werden können, Rasterdaten jedoch insbesondere bei einem Wechsel von flächentreuer auf geographischer Projektion verzerrt (z. B. schräge Schriften, verzogene Signaturen) oder durch die Verzerrung gar unleserlich abgebildet werden, da ihr Layout direkt abhängig von der Ursprungsprojektion ist, in der die Erstellung erfolgte.

Zwar sind Konvertierungen von Vektor- zu Rasterdaten in die jeweils eine oder andere Richtung möglich, jedoch ist der Informationsgehalt der Rasterdaten grundsätzlich niedriger als der Informationsgehalt von Vektordaten.

Rasterdaten, die nachträglich vektorisiert werden, stehen im Nachgang als neue Geodaten im Vektorformat zur Verfügung. Der Vorgang kann jedoch nicht exakt wiederholt werden, da es keine vollständige automatische Vektorisierung gibt, die aus geodatentechnischer Sicht ausreichend oder fehlerfrei ist. So muss stets eine Nachbearbeitung per Hand erfolgen. Das gleiche gilt für eine Klassifizierung auf Basis der Farbpalette, um Attributinformationen aus Rasterdaten zu gewinnen. Zwar kann die automatische



Klassifizierung erneut durchgeführt werden, muss jedoch, um die in der Regel notwendige hohe Qualität zu erreichen, dennoch per Hand und mit Sachverstand nachbearbeitet werden.

Rasterdaten, die eine Vektorisierung erfahren, erfahren eine qualitative Verbesserung hin zu einem vollständig neuen Geodatensatz.

Hingegen wird bei der Umwandlung von Vektordaten in Rasterdaten auf ein festgefügtes Layout zurück gegriffen. Der Vorgang des Ausrasterns kann jederzeit in genau dieser Form erneut ausgeführt werden. Bei auf diese Art und Weise erstellten Rasterdaten handelt es sich daher um abgeleitete Daten auf einer niedrigeren Qualitätsstufe, die zudem einen hohen Speicherbedarf besitzen.

Für die Archivierung ist es unerlässlich, Rasterdaten als Rasterdaten in einem für die Archivierung geeigneten Format zu hinterlegen und Vektordaten als Vektordaten eines unabhängig lesbaren ASCII-Format zu konvertieren, sofern eine qualitativ vollständige Archivierung erfolgen soll. Wurden die Vektordaten aufwändig gestaltet, so kann zusätzlich eine Rasterdatei zu den Vektordaten hinterlegt werden, die das ursprünglich gelayoutete Bild, sowie eine Legende für das Layout zeigt. Aus dem ebenfalls zu archivierenden Metadatensatz können im Idealfall alle übrigen Informationen entnommen werden.