# THE JOURNEY OF HACKING IN A NEW SERDE DATAFORMAT

# 2024-02-03 / FOSDEM '24 / RUST DEVROOM

## ABSTRACT (PARAPHRASED)

neo4rs use serde. I want to present the journey of building [that].

Introduction to serde

- Introduction to serde
- Introduction to neo4rs

- Introduction to serde
- Introduction to neo4rs
- An actual deep deep-dive

- Introduction to serde
- Introduction to neo4rs
- An actual deep deep-dive
- Discussion about how to pronounce serde

#### **ABOUT ME**

- Hi, I'm Paul
- @knutwalker[@hachyderm.io]

# 



- A Graph Database written in Not-Rust
- We have a Rust driver: neo4rs, written in pure Rust
- Developed under Neo4j Labs
- github.com/neo4j-labs/neo4rs

Neo4j drivers generally comunicate via Bolt

- Neo4j drivers generally comunicate via Bolt
- Bolt is a binary protocol

- Neo4j drivers generally comunicate via Bolt
- Bolt is a binary protocol
  - packstream for general data types

- Neo4j drivers generally comunicate via Bolt
- Bolt is a binary protocol
  - packstream for general data types
  - "binary JSON-ish"

- Neo4j drivers generally comunicate via Bolt
- Bolt is a binary protocol
  - packstream for general data types
  - "binary JSON-ish"
  - Domain-specific structs (~15)

#### **ABOUT NEO4RS**

Bolt structs as a BoltType enum

```
enum BoltType {
    Null,
    Integer(i64),
    String(String),
    List(Vec<BoltType>),
    Node(BoltNode),
    // ...
}
```

```
let event = node.get::<String>("event").unwrap();
let year = node.get::<u64>("year").unwrap();
```

```
let event = node.get::<String>("event").unwrap();
let year = node.get::<u64>("year").unwrap();
```

```
#[derive(serde::Deserialize)]
struct Session {
    event: String,
    year: u64,
}
```

```
let session = node.to::<Session>().unwrap();
```

- A framework for \_ser\_ializing and \_de\_serializing
   Rust data structures (serde.rs)
- Jon's decrusting video

data type

data format

data model

- data type
- #[derive(Serialize, Deserialize)]
- data format

data model

- data type
- #[derive(Serialize, Deserialize)]
- data format
- Serializer and Deserializer traits

data model

- data type
- #[derive(Serialize, Deserialize)]
- data format
- Serializer and Deserializer traits
- notice the r at the end of the names
- data model

- data type
- #[derive(Serialize, Deserialize)]
- data format
- Serializer and Deserializer traits
- notice the r at the end of the names
- data model
- mostly represented in the API only

example data format: JSON

- example data format: JSON
- data format implementation: serde\_json

In neo4rs, data is already available as BoltType

- In neo4rs, data is already available as BoltType
- Implement a data format for BoltType

- In neo4rs, data is already available as BoltType
- Implement a data format for BoltType
- Deserializeronly

- In neo4rs, data is already available as BoltType
- Implement a data format for BoltType
- Deserializeronly
- Maintain API compatibility, if possible

#### NODE

```
Node::Structure(
    id::Integer,
    labels::List<String>,
    properties::Dictionary,
)
```

#### NODE

```
pub struct BoltNode {
    pub id: BoltInteger,
    pub labels: BoltList,
    pub properties: BoltMap,
}
```

#### **EXAMPLE**

```
CREATE (n:Session { event: 'FOSDEM', year: 2024 })
RETURN n
```

```
1 CREATE (n
2 :Session
3  {
4         event: 'FOSDEM',
5         year: 2024
6    }
7 ) RETURN n
```

```
#[derive(Deserialize)]
struct Session {
   event: String,
   year: u64,
}
```

```
1 let session = row.get::<Session>("n").unwrap();
2
3 let node = row.get::<Node>("n").unwrap();
4 let session = node.to::<Session>().unwrap();
```

```
1 let session = row.get::<Session>("n").unwrap();
2
3 let node = row.get::<Node>("n").unwrap();
4 let session = node.to::<Session>().unwrap();
```

```
1 let session = row.get::<Session>("n").unwrap();
2
3 let node = row.get::<Node>("n").unwrap();
4 let session = node.to::<Session>().unwrap();
```

```
1 #[derive(Deserialize, Serialize)]
2 pub struct BoltNode {
3    pub id: BoltInteger,
4    pub labels: BoltList,
5    pub properties: BoltMap,
6 }
```

```
1 #[derive(Deserialize, Serialize)]
2 pub struct BoltNode {
3    pub id: BoltInteger,
4    pub labels: BoltList,
5    pub properties: BoltMap,
6 }
```

```
pub fn to<T: Deserialize>(self) -> Result<T, Error> {
    let value = serde_json::to_value(self)?;
    let result = serde_json::from_value(value)?;
    Ok(result)
}
```

```
pub fn to<T: Deserialize>(self) -> Result<T, Error> {
    let value = serde_json::to_value(self)?;
    let result = serde_json::from_value(value)?;
    Ok(result)
}
```

```
pub fn to<T: Deserialize>(self) -> Result<T, Error> {
    let value = serde_json::to_value(self)?;
    let result = serde_json::from_value(value)?;
    Ok(result)
}
```

```
pub fn to<T: Deserialize>(self) -> Result<T, Error> {
    let value = serde_json::to_value(self)?;

let result = serde_json::from_value(value)?;

Ok(result)
}
```

Are we done?

#### Are we done?

```
missing field `event`
missing field `year`
```

```
#[derive(Deserialize, Serialize)]
pub struct BoltNode {
   pub id: BoltInteger,
   pub labels: BoltList,
   pub properties: BoltMap,
}
```

```
#[derive(Deserialize)]
struct SessionNode {
   properties: Session
}
```

#### **SECOND ATTEMPT**

```
pub fn to<T: Deserialize>(self) -> Result<T, Error> {
    let value = serde_json::to_value(self.properties)?;
    let result = serde_json::from_value(value)?;
    Ok(result)
}
```

#### **SECOND ATTEMPT**

- Kinda works
- No way to get id and labels

#### THIRD ATTEMPT

```
1 #[derive(Deserialize)]
2 struct Session {
3    event: String,
4    year: u64,
5    id: u64,
6    labels: Vec<String>,
7 }
```

#### THIRD ATTEMPT

```
1 #[derive(Deserialize)]
2 struct Session {
3    event: String,
4    year: u64,
5    id: u64,
6    labels: Vec<String>,
7 }
```

```
1 impl Deserialize for Session {
2    fn deserialize<D>(deserializer: D)
3    -> Result<Self, D::Error>
4    where
5         D: Deserializer,
6    {
7         todo!()
8    }
9 }
```

```
1 impl Deserialize for Session {
2    fn deserialize<D>(deserializer: D)
3    -> Result<Self, D::Error>
4    where
5         D: Deserializer,
6    {
7         todo!()
8    }
9 }
```

```
1 impl Deserialize for Session {
2    fn deserialize<D>(deserializer: D)
3    -> Result<Self, D::Error>
4    where
5         D: Deserializer,
6    {
7         todo!()
8    }
9 }
```

```
1 impl Deserialize for Session {
2    fn deserialize<D>(deserializer: D)
3    -> Result<Self, D::Error>
4    where
5        D: Deserializer,
6    {
7        todo!()
8    }
9 }
```

```
1 struct SessionVisitor;
2
3 impl Visitor for SessionVisitor {
4    type Value = Session;
5
6    fn expecting(&self, formatter: &mut Formatter) -> fmt::Res
7         Formatter::write_str(formatter, "struct Session")
8    }
9 }
```

```
1 struct SessionVisitor;
2
3 impl Visitor for SessionVisitor {
4    type Value = Session;
5
6    fn expecting(&self, formatter: &mut Formatter) -> fmt::Res
7         Formatter::write_str(formatter, "struct Session")
8    }
9 }
```

```
1 struct SessionVisitor;
2
3 impl Visitor for SessionVisitor {
4    type Value = Session;
5
6    fn expecting(&self, formatter: &mut Formatter) -> fmt::Res
7         Formatter::write_str(formatter, "struct Session")
8    }
9 }
```

```
1 struct SessionVisitor;
2
3 impl Visitor for SessionVisitor {
4    type Value = Session;
5
6    fn expecting(&self, formatter: &mut Formatter) -> fmt::Res
7         Formatter::write_str(formatter, "struct Session")
8    }
9 }
```

```
1 impl Visitor for SessionVisitor {
2    fn visit_map<A>(self, mut map: A)
3    -> Result<Self::Value, A::Error>
4    where
5         A: MapAccess,
6    {
7         todo!()
8    }
9 }
```

```
1 impl Visitor for SessionVisitor {
2    fn visit_map<A>(self, mut map: A)
3    -> Result<Self::Value, A::Error>
4    where
5         A: MapAccess,
6    {
7         todo!()
8    }
9 }
```

```
1 impl Visitor for SessionVisitor {
2    fn visit_map<A>(self, mut map: A)
3    -> Result<Self::Value, A::Error>
4    where
5         A: MapAccess,
6    {
7         todo!()
8    }
9 }
```

```
1 fn visit_map<A: MapAccess>(self, mut map: A) -> Result<Self::\
2    let mut event: Option<String> = None;
3    let mut year: Option<u64> = None;
4    let mut id: Option<u64> = None;
5    let mut labels: Option<Vec<String>> = None;
6 }
```

```
fn visit map<A: MapAccess>(self, mut map: A) -> Result<Self::</pre>
       while let Some(key) = map.next key::<&str>()? {
           match key {
               "event" => event =
                   Some(map.next value::<String>()?),
                "year" => year = Some(map.next value::<u64>()?),
               "id" => id = Some(map.next value::<u64>()?),
               "labels" => labels = Some(map.next value::<Vec<St
               => todo!("unknown field"),
12 }
```

```
fn visit map<A: MapAccess>(self, mut map: A) -> Result<Self::</pre>
       while let Some(key) = map.next key::<&str>()? {
           match key {
               "event" => event =
                   Some(map.next value::<String>()?),
                "year" => year = Some(map.next value::<u64>()?),
               "id" => id = Some(map.next_value::<u64>()?),
               "labels" => labels = Some(map.next value::<Vec<St
               => todo!("unknown field"),
12 }
```

```
fn visit map<A: MapAccess>(self, mut map: A) -> Result<Self::</pre>
       while let Some(key) = map.next key::<&str>()? {
           match key {
               "event" => event =
                   Some(map.next value::<String>()?),
                "year" => year = Some(map.next value::<u64>()?),
               "id" => id = Some(map.next_value::<u64>()?),
               "labels" => labels = Some(map.next value::<Vec<St
               => todo!("unknown field"),
12 }
```

```
fn visit map<A: MapAccess>(self, mut map: A) -> Result<Self::</pre>
       while let Some(key) = map.next key::<&str>()? {
           match key {
               "event" => event =
                   Some(map.next value::<String>()?),
                "year" => year = Some(map.next_value::<u64>()?),
               "id" => id = Some(map.next value::<u64>()?),
               "labels" => labels = Some(map.next value::<Vec<St
               => todo!("unknown field"),
12 }
```

```
fn visit map<A: MapAccess>(self, mut map: A) -> Result<Self::</pre>
       while let Some(key) = map.next key::<&str>()? {
           match key {
               "event" => event =
                    Some(map.next value::<String>()?),
                "year" => year = Some(map.next_value::<u64>()?),
 6
                "id" => id = Some(map.next value::<u64>()?),
                "labels" => labels = Some(map.next value::<Vec<St
               => todo!("unknown field"),
12 }
```

```
1 struct BoltNodeDeserializer(BoltNode);
2
3 impl IntoDeserializer<DeError> for BoltNode {
4    type Deserializer = BoltNodeDeserializer;
5
6    fn into_deserializer(self) -> Self::Deserializer {
7        BoltNodeDeserializer::new(self)
8    }
9 }
```

```
1 struct BoltNodeDeserializer(BoltNode);
2
3 impl IntoDeserializer<DeError> for BoltNode {
4    type Deserializer = BoltNodeDeserializer;
5
6    fn into_deserializer(self) -> Self::Deserializer {
7        BoltNodeDeserializer::new(self)
8    }
9 }
```

```
impl Deserializer for BoltNodeDeserializer {
   type Error = DeError;

forward_to_deserialize_any! {
   bool i8 i16 i32 i64 i128 u8 u16 u32 u64 u128
   f32 f64 char str string bytes byte_buf option
   unit unit_struct seq tuple tuple_struct struct
   identifier enum map ignored_any newtype_struct
}
```

```
impl Deserializer for BoltNodeDeserializer {
   type Error = DeError;

forward_to_deserialize_any! {
   bool i8 i16 i32 i64 i128 u8 u16 u32 u64 u128
   f32 f64 char str string bytes byte_buf option
   unit unit_struct seq tuple tuple_struct struct
   identifier enum map ignored_any newtype_struct
}
```

```
1 fn to<T: Deserialize>(self) -> Result<T, DeError> {
2    T::deserialize(self.into_deserializer())
3 }
```

• We now have the same result as before

- We now have the same result as before
- Let's add id and labels

Kinda works

- Kinda works
- We do get id and labels, but:

- Kinda works
- We do get id and labels, but:
- We only map them by the names id and labels

- Kinda works
- We do get id and labels, but:
- We only map them by the names id and labels
- We could use something like \_\_\_id, but, meh

- Kinda works
- We do get id and labels, but:
- We only map them by the names id and labels
- We could use something like \_\_\_id, but, meh
- Let's try something else instead

```
struct Id(pub u64);
struct Labels(pub Vec<String>);
```

```
1 #[derive(Deserialize)]
2 struct Session {
3    event: String,
4    year: u64,
5    id: neo4rs::Id,
6    labels: neo4rs::Labels,
7 }
```

```
1 #[derive(Deserialize)]
2 struct Session {
3    event: String,
4    year: u64,
5    id: neo4rs::Id,
6    labels: neo4rs::Labels,
7 }
```

```
impl Deserializer for BoltNodeDeserializer {
       fn deserialize struct<V>(
           self,
           _name: &'static str,
 5
           fields: & 'static [& 'static str],
 6
           visitor: V,
       ) -> Result<V::Value, Self::Error>
     where
           V: Visitor,
           visitor.visit map(MapDeserializer::new(todo!()))
11
13 }
```

```
impl Deserializer for BoltNodeDeserializer {
       fn deserialize struct<V>(
           self,
           _name: &'static str,
           fields: &'static [&'static str],
           visitor: V,
     ) -> Result<V::Value, Self::Error>
     where
           V: Visitor,
           visitor.visit map(MapDeserializer::new(todo!()))
11
13 }
```

```
impl Deserializer for BoltNodeDeserializer {
       fn deserialize struct<V>(self, fields: &[&str], visitor:
           let property fields = self.0.properties
                .iter()
                .map(|(k, v)| (k, StructData::Property(v)));
 5
           let additional_fields = fields
                .iter()
                .copied()
                .filter(|f| !self.0.properties.contains key(*f))
11
                .map(|f| (f, StructData::Node(self.0)));
13 }
```

```
impl Deserializer for BoltNodeDeserializer {
       fn deserialize struct<V>(self, fields: &[&str], visitor:
           let property fields = self.0.properties
                .iter()
                .map(|(k, v)| (k, StructData::Property(v)));
           let additional_fields = fields
 8
                .iter()
                .copied()
                .filter(|f| !self.0.properties.contains key(*f))
                .map(|f| (f, StructData::Node(self.0)));
13 }
```

```
impl Deserializer for BoltNodeDeserializer {
       fn deserialize struct<V>(self, fields: &[&str], visitor:
           let property fields = self.0.properties
                .iter()
                .map(|(k, v)| (k, StructData::Property(v)));
           let additional_fields = fields
                .iter()
                .copied()
10
                .filter(|f| !self.0.properties.contains key(*f))
11
                .map(|f| (f, StructData::Node(self.0)));
13 }
```

```
impl Deserializer for BoltNodeDeserializer {
       fn deserialize struct<V>(self, fields: &[&str], visitor:
           let property fields = self.0.properties
                .iter()
                .map(|(k, v)| (k, StructData::Property(v)));
           let additional_fields = fields
                .iter()
                .copied()
                .filter(|f| !self.0.properties.contains key(*f))
                .map(|f| (f, StructData::Node(self.0)));
11
13 }
```

```
impl Deserializer for AdditionalNodeData {
       fn deserialize_newtype_struct<V>(
           self,
           name: &str,
          visitor: V
     ) -> Result<V::Value, Self::Error>
   where
          V: Visitor,
           todo!()
11
12 }
```

```
impl Deserializer for AdditionalNodeData {
       fn deserialize_newtype_struct<V>(
           self,
           name: &str,
          visitor: V
   ) -> Result<V::Value, Self::Error>
   where
          V: Visitor,
           todo!()
11
12 }
```

```
impl Deserializer for AdditionalNodeData {
       fn deserialize newtype struct<V: Visitor>(self, name: &st
           match name {
               "Id" => visitor.visit_newtype_struct(
                   self.data.id.into deserializer()
                "Labels" => visitor.visit newtype struct(
                   SeqDeserializer::new(
                        self.data.labels.iter()
                 => todo!()
15 }
```

```
impl Deserializer for AdditionalNodeData {
       fn deserialize newtype struct<V: Visitor>(self, name: &st
           match name {
                "Id" => visitor.visit_newtype_struct(
 4
 5
                    self.data.id.into deserializer()
                "Labels" => visitor.visit newtype struct(
                    SeqDeserializer::new(
                        self.data.labels.iter()
                 => todo!()
15 }
```

```
fn deserialize_newtype_struct<V: Visitor>(self, name: &st
           match name {
                "Id" => visitor.visit_newtype_struct(
                    self.data.id.into deserializer()
                "Labels" => visitor.visit_newtype_struct(
                    SeqDeserializer::new(
                        self.data.labels.iter()
10
11
                 => todo!()
15 }
15 }
```

Works

- Works
- There are still some downsides

- Works
- There are still some downsides
- #[serde(default)] doesn't work

# #[serde(default)]

# #[serde(default)]

# #[serde(default)]

```
impl Deserializer for BoltNodeDeserializer {
    fn deserialize_struct<V>(self, fields: &[&str], visitor: V
        let additional_fields = fields
        .iter()
        .copied()
        .filter(|f| !self.0.properties.contains_key(*f))
        .map(|f| (f, StructData::Node(self.0)));
}
```

- Works
- There are still some downsides
- #[serde(default)] doesn't work

- Works
- There are still some downsides
- #[serde(default)] doesn't work
- Workaround: Option<T>

# BoltType

Rinse and repeat for BoltType and its ~20 variants

# BoltType

- Rinse and repeat for BoltType and its ~20 variants
- Almost...

```
impl Deserializer for BoltTypeDeserializer {
       fn deserialize bytes<V>(self, visitor: V) -> Result<V::Va
       where
           V: Visitor,
           if let BoltType::Bytes(v) = self.value {
               visitor.visit bytes(v.clone())
           } else {
               self.unexpected(visitor)
12 }
```

```
impl Deserializer for BoltTypeDeserializer {
       fn deserialize bytes<V>(self, visitor: V) -> Result<V::Va</pre>
      where
          V: Visitor,
 6
           if let BoltType::Bytes(v) = self.value {
                visitor.visit bytes(v.clone())
           } else {
                self.unexpected(visitor)
12 }
```

```
impl Deserializer for BoltTypeDeserializer {
       fn deserialize bytes<V>(self, visitor: V) -> Result<V::Va</pre>
      where
          V: Visitor,
           if let BoltType::Bytes(v) = self.value {
               visitor.visit bytes(v.clone())
           } else {
               self.unexpected(visitor)
12 }
```

No, but out of time

- No, but out of time
- Here are more things to consider:

- No, but out of time <sup>(2)</sup>
- Here are more things to consider:
- Maintain existing API:

- No, but out of time <sup>(2)</sup>
- Here are more things to consider:
- Maintain existing API:
  - Have special cases for converting into chrono/time types

- No, but out of time
- Here are more things to consider:
- Maintain existing API:
  - Have special cases for converting into chrono/time types
    - is\_human\_readable

- No, but out of time <sup>(2)</sup>
- Here are more things to consider:
- Maintain existing API:
  - Have special cases for converting into chrono/time types
    - is\_human\_readable
    - different precision for timestamps

- No, but out of time
- Here are more things to consider:
- Maintain existing API:

- No, but out of time
- Here are more things to consider:
- Maintain existing API:
  - Need to deserialize into itself

- No, but out of time <sup>(2)</sup>
- Here are more things to consider:
- Maintain existing API:
  - Need to deserialize into itself
    - Custom Deserialize impl for BoltType

- No, but out of time
- Here are more things to consider:
- Maintain existing API:
  - Need to deserialize into itself
    - Custom Deserialize impl for BoltType
    - That impl is not really usable for other data formats

- No, but out of time
- Here are more things to consider:
- Allow for unexpected fields

- No, but out of time
- Here are more things to consider:
- Allow for unexpected fields
  - deserialize\_ignored\_any

- No, but out of time
- Here are more things to consider:
- Allow for zero-copy deserialization

- No, but out of time
- Here are more things to consider:
- Allow for zero-copy deserialization
  - Keep 'de lifetimes around

- No, but out of time <sup>(2)</sup>
- Here are more things to consider:
- Allow for zero-copy deserialization
  - Keep 'de lifetimes around
  - implement all the \_borrowed methods

## THANK YOU!

Q/A?

```
use
std::process::exit;
 exit(42);
```